Apples to Apples: Differences in Viewer Retention When Longer Content is Chopped into Smaller Bites

Jennifer S. Kay
Computer Science Department
Rowan University
Glassboro, NJ, USA
kay@rowan.edu

Colin Lambe
Computer Science Department
Rowan University
Glassboro, NJ, USA
lambec37@students.rowan.edu

Tyler J. Nolan*
Glassboro, NJ, USA
tylernolan94@gmail.com

Thomas M. Grello*
Raytheon
Atlantic City, USA
tommgrello@gmail.com

Anthony Breitzman
Computer Science Department
Rowan University
Glassboro, NJ, USA
breitzman@rowan.edu

ABSTRACT

Numerous studies have concluded that viewer retention decreases as video length increases. However, we are not aware of any prior work in which a set of longer MOOC (Massive Open Online Course) videos are compared with the same content split into multiple shorter videos. We are fortunate to be in the unique position to have two separate MOOCs that teach essentially the same content using two different platforms (the LEGO Mindstorms NXT and EV3 robots). In our NXT MOOC, videos are quite long, with over 20% of the videos having a running time of more than ten minutes. The EV3 MOOC has very similar content; EV3 MOOC scripts were written by modifying NXT scripts as appropriate. However, many of the EV3 lessons were split into two or three shorter videos in place of a single longer one. NXT videos that are very close in terms of both content and duration to EV3 videos have similar average percentage viewed. This suggests that the two populations watching the videos are similar and that we have a promising setup for analyzing the relationship between NXT lessons whose EV3 counterparts consist of multiple shorter videos. We present an analysis of our data, along with various interpretations some, but not all, of which support the "shorter videos are better" hypothesis.

L@S '19, June 24–25, 2019, Chicago, IL, USA

ACM 978-1-4503-6804-9/19/06…$15.00
DOI: 10.1145/3330430.3333617

Author Keywords

Video analysis; MOOC; viewer engagement; online education; in-video dropout;

INTRODUCTION

Numerous studies have concluded that viewer retention decreases as video length increases. However, we are not aware of any prior work in which a single longer MOOC (Massive Open Online Course) video was compared with the same content split into multiple shorter videos. We are fortunate to be in the unique position to have two separate MOOCs that teach essentially the same content using two different platforms (the LEGO Mindstorms NXT and EV3 robots). In our NXT MOOC, videos are quite long, with over 20% of the videos having a running time of more than ten minutes. The EV3 MOOC has very similar content; EV3 MOOC scripts were written by modifying NXT scripts as appropriate. However, many of the EV3 lessons were split into two or three shorter videos in place of a single longer one. This paper reports on a study comparing viewer retention between videos from the two MOOCs.

RELATED WORK

Type and Style of Video Content

The goal of most video producers is, presumably, for viewers to watch 100% of each video that they produce. As one might expect, there are many factors beyond simply just video length. A study of online advertising that used eye tracking and facial expression to better understand viewers concluded that levels of "surprise and joy" in an advertisement can affect viewer retention [9]. Authors of a Biology MOOC were surprised[1] to find that their unscripted videos of a faculty lecture were more engaging than their

[1] But perhaps not joyful!
*Work done while at Rowan University

carefully scripted and fully animated "deep dives" [8]. Seeing the instructor, particularly close up and making good eye contact with the audience was also found to improve viewer retention in a study of videos from edX Computer Science, Statistics, and Chemistry MOOCs [3]. This study also found other ways that a more human touch affects retention, from incorporating handwriting to enthusiastic instruction [3].

Video Duration
Regardless of recommendations about video content, most studies of viewer engagement include a conclusion that shorter videos are better. Their conclusions do, however, differ on their definition of "short." The online advertising study had 58 test subjects and only considered the first 30 seconds of a video, since there was too little watch data for videos watched beyond the 30 second point [9]. Wistia, a site that hosts videos for business websites, considered 1.3 billion video plays and concluded that viewer retention is fairly steady for the first two minutes, followed by an exponential drop off until the six minute mark. Between six and twelve minutes there was little drop off [1]. This is in contrast to an earlier study of Wistia data that found videos of between two and ten minutes in length had little drop off [7]. In the context of online learning, Guo et. al. recommended (based on 6.9 million video plays) that longer content should be split into videos of at most six minutes [3], while Thornton et. al. recommend that videos be at most 10 minutes [8].

Partial Videos
Wistia's earlier study had one additional result which is of particular relevance to our own work. They compared two specific marketing videos: Video A was 30 seconds in length. Video B ran for a total of 90 seconds, but the first 30 seconds of Video B were identical to Video A. Surprisingly, viewer retention was very different at the 30 second mark for the two videos – with Video A having double the number of viewers when it finished (at 30 seconds) as Video B did at the 30 second point! [7]

MOOC Participants Have Differing Goals
Wilkowski et. al. [11] pointed out that not everyone has the same goals when they begin a MOOC. Some students are "completers," whose goal is to pass the course and earn a certificate of completion. But there are also "observers," who may just want to have a brief peek to see what the course is like, as well as "casual learners" who are not intent on completing all of the material, but rather wish to learn one or two things. Finally, there are "no-shows" who sign up but never actually log on to the course.

THE NXT AND EV3 MOOCS
Prior to our robotics MOOCs we had been offering face-to-face NXT workshops that ran for 3 full days. In 2013, the NXT MOOC [6] served as a first foray into the world of MOOCs. Having chosen the Course Builder [2] online education platform for our course, the decision was made to make the format of our course similar to that of another Course Builder MOOC that we liked the look of – Power Searching with Google [4]. Thus, our MOOC was split into five "**weeks**," each week consisted of a set of "**lessons**," and each lesson had a video followed by either some self-test questions or a robot programming project. In addition, we offered an "extra help" section with a set of optional troubleshooting videos that users could choose to watch, which we have not included in the data presented in this paper.

At the recommendation of colleagues in the film department, we came up with carefully written scripts for our existing slides. The hands on demonstrations were roughly outlined but unscripted. We thought of our videos as being relatively short, but as shown in Table 1, only about half of our videos were less than six minutes. The average duration across all videos for the NXT MOOC is roughly 6.3 minutes.

NXT Videos			
Duration (minutes)	**<6**	**6-10**	**>10**
# of Videos	20	10	9
Approximate Percentage (Out of n=39)	51%	26%	23%

Table 1. The NXT MOOC has only 39 videos. Almost half of them are longer than six minutes, and almost one quarter are longer than ten minutes.

Shortly after we had decided to create the NXT MOOC, but before we had begun filming, LEGO announced their intention to release a new Mindstorms Robot, the EV3. We will be forever grateful to our program manager who pointed out that it would take some time for most NXT owners to replace their NXTs with EV3s, and that we should proceed with the NXT MOOC.

The EV3 robot is a major improvement over the NXT, with faster hardware and an improved graphical programming language, and once we had completed our NXT MOOC we started thinking about the creation of a second MOOC covering the EV3. Despite the differences between the two platforms, the underlying programming concepts that we wanted to teach were essentially the same. Thus, when creating the scripts and slides for our EV3 MOOC, [5], we began with those of the NXT MOOC. On many slides we did not need to do much more than replace pictures of NXT hardware with pictures of EV3 hardware and replace screenshots of NXT software with EV3 software; many scripts retained large portions of virtually identical content, though the software walkthroughs did, of course, differ.

We started work on the EV3 MOOC in 2014, shortly after the first Learning @ Scale conference and as a result of the "shorter is better" result from [3], we chose to split several of the longer videos (and thus, lessons) into multiple "**parts**," explicitly calling them "part 1," "part 2," "part 3," and so on. Table 2 shows that while the new EV3 MOOC does still have a handful of videos that exceed ten minutes in duration, we also have many more videos that are less than ten minutes long. The mean video duration is just over 4.5 minutes. As of April 2019, the NXT MOOC has had approximately 9000 participants, and the EV3 MOOC has had approximately 6500.

METHOD

Video Selection

For this study, we reviewed the scripts from the NXT and EV3 MOOCs. As noted above, the content differences between the two MOOCs are not so much related to the pedagogical content as to the differences in implementation using the older NXT and newer EV3 robots and software.

EV3 Videos			
Duration (minutes)	**<6**	**6-10**	**>10**
# of Videos	61	15	6
Approximate Percentage (Out of n=82)	74%	18%	7%

Table 2. The EV3 MOOC's videos have similar content to the NXT MOOC's videos. However, the EV3 MOOC has more than twice as many videos, the majority of which are shorter than 6 minutes.

Corresponding lessons between the two courses were given one of three labels:

- **One-to-one**: Lessons whose videos are essentially identical in both content and duration between the two courses.
- **One-to-many**: Lessons in which one NXT video is essentially identical in terms of content to multiple EV3 videos.
- **Other**: All remaining lessons.

We determined that there were eight sets of one-to-one videos (each of which contained a single NXT video and its single EV3 counterpart) as well as seven one-to-many sets of videos (in which the NXT versions contained a single video, but the EV3 versions contained between two and four shorter videos).

Data Collection

We retrieved summary data for each of the one-to-one and one-to-many lessons using YouTube's basic video analytics tools [10]. The data for these videos are summarized in Tables 3 and 4, where:

- **Length** refers to the run time of an individual video.
- **Number of Views** is the number of individual views of the video.
- **Watch Time** is the sum of all of the individual views' durations. For example, if 8 people watch 100% of a 2 minute video, total watch time would be 16 minutes.
- **Average Percentage Viewed** is computed as follows:

$$Average\ \%\ Viewed = \frac{Video\ Watch\ Time}{\#Views \times Video\ Length}$$

NXT MOOC					EV3 MOOC				
Week / Lesson	Length (minutes)	Number of Views	Watch Time (minutes)	Average Percentage Viewed	Week / Lesson	Length (minutes)	Number of Views	Watch Time (minutes)	Average Percentage Viewed
1 / 6	3.7	1996	5100	69.1 %	1 / 6	3.68	1402	3574	69.3 %
2 / 1	1.63	2132	2440	70.2 %	2 / 1	1.9	1864	2605	73.6 %
2 / 2	2.25	1419	2622	82.1 %	2 / 2	2.43	1383	2819	83.9 %
2 / 7	4.82	1196	4204	72.9 %	2 / 7	5.87	1271	5848	78.4 %
3 / 4	5.73	1104	4655	73.6 %	3 / 4	6.27	1079	4934	72.9 %
3 / 10	10.73	667	4892	68.4 %	3 / 9	11.18	582	4313	66.3 %
4 / 6	6.37	601	2691	70.3 %	4 / 4	7.87	589	3133	67.6 %
5 / 6	2.05	250	368	71.8 %	5 / 6	2.75	184	339	67.0 %

Table 3. The one-to-one video data. Each row of data shows the information for a pair of videos, one from the NXT MOOC and a corresponding one from the EV3 MOOC, that cover essentially the same content.

Analysis

While our data represents thousands of video views, YouTube has aggregated the individual views into a total watch time. As a result, our sample size is actually very small, corresponding to the rows on Tables 3 and 4, and so we have chosen to use Mann-Whitney U test for this analysis. As discussed in the future work section, we hope in the future to be able to collect a larger data set for a more robust result. Nevertheless, our small sample sizes do not necessarily invalidate our results – as a non-parametric test Mann-Whitney does not have the same restrictions as a parametric t-test.

Similarity of the Two Participant Populations

MOOC participants are self-selected, and there are certainly plausible arguments as to why the two populations might differ in their behavior. For example, NXT owners who upgrade to EV3 robots may join the EV3 MOOC but flit in and out of videos, watching segments of some videos until they feel they understand the mapping between the two systems, skipping other videos altogether.

The presence of the one-to-one videos allows us to investigate whether there is a statistically significant difference between NXT viewer behavior and EV3 viewer behavior. It is useful to note that our one-to-one data set includes at least one video from each week.

The results of Mann-Whitney on the Average Percentage Viewed data from the NXT and EV3 MOOCS are shown in Table 5, which suggest that there is no statistically significant (<0.05) or near-significant (<0.1) difference between the NXT and EV3 viewers viewing habits.

DISCUSSION: ARE SHORTER VIDEOS BETTER?

The results of performing Mann-Whitney on the one-to-many data from Table 4 are shown in Table 6 and show a significant (<0.05) difference in average watch percentage

NXT MOOC						EV3 MOOC				
Week / Lesson	Length (minutes)	Number of Views	Watch Time (minutes)	Average Percentage Viewed		Week / Lesson / Part	Length (minutes)	Number of Views	Watch time (minutes)	Average Percentage Viewed
1 / 3	9.13	2809	14100	55.0 %		1 / 3 / 1	2.8	2067	4177	72.2 %
						1 / 3 / 2	2.87	1730	3640	73.3 %
						1 / 3 / 3	4.27	1575	4381	65.1 %
1 / 8	5	1936	7042	72.7 %		1 / 8 / 1	1.15	1428	1339	81.5 %
						1 / 8 / 2	7.83	1559	8417	69.0 %
						1 / 8 / 3	0.5	1258	558	88.7 %
2 / 3	5.97	1742	7135	68.6 %		2 / 3 / 1	3.35	1618	4082	75.3 %
						2 / 3 / 2	4.32	1497	4682	72.4 %
3 / 1	8.37	2240	8930	47.6 %		3 / 1 / 1	2.78	2064	4056	70.7 %
						3 / 1 / 2	4.8	1421	4974	72.9 %
						3 / 1 / 3	1.7	1222	1704	82.0 %
3 / 8	7.53	1217	6202	67.7 %		3 / 6 / 1	6.35	1312	6295	75.6 %
						3 / 6 / 2	5.45	1052	4341	75.7 %
3 / 12	10.17	681	4365	63.0 %		3 / 11 / 1	5.27	620	2118	64.8 %
						3 / 11 / 2	6.75	535	2287	63.3 %
						3 / 11 / 3	7	537	2299	61.2 %
						3 / 11 / 4	5.05	473	1711	71.6 %
4 / 7	11.05	753	4647	55.8 %		4 / 5 / 1	0.98	560	454	82.7 %
						4 / 5 / 2	2.35	587	1092	79.2 %
						4 / 5 / 3	8.97	653	3646	62.2 %

Table 4. The one-to-many video data. A single NXT video on the left covers the same material as is covered in multiple EV3 videos on the right. Each video on the left corresponds to the set of videos on the right starting with the EV3 video in its row and continuing through the others in contiguous rows below with the same shading.

Ranks

	MOOC	N	Mean Rank	Sum of Ranks
Avg % viewed	NXT	8	8.88	71.00
	EV3	8	8.13	65.00
	Total	16		

Test Statistics[a]

	Avg % viewed
Mann-Whitney U	29.000
Wilcoxon W	65.000
Z	-.315
Asymp. Sig. (2-tailed)	.753
Exact Sig. [2*(1-tailed Sig.)]	.798[b]

a. Grouping Variable: MOOC

b. Not corrected for ties.

Table 5. Mann-Whitney U Test on one-to-one data

between the average watch % of the NXT and EV3 videos.

Our initial reaction to this result was that it was a strong confirmation of what we had anticipated would be the case, that shorter videos have better viewer retention. However, a closer look at the data in Table 4 makes us question its validity. For example, consider the first one-to-many set in Table 4. There is quite a big variation between the number of views in each of the sub-parts of EV3 Lesson 1/3: 1/3/1 had 2067 views, 1/3/2 had 1730 views, and 1/3/3 had a mere 1575 views.

It is impossible to know how to interpret these numbers without knowing more about the individual viewers. Here we consider two alternative interpretations.

Possible Interpretation 1: Loss of Viewership over Time

It might be the case that all 2067 viewers initially planned on watching all 3 of the videos, but changed their minds as they progressed through them – some dropping out part way through lesson 1/3/1, others continuing on to 1/3/2 but not making it all the way through and quit without starting to watch 1/3/3. In this case, it's not really fair to look at the individual percentages viewed for each of the shorter videos. Instead, we should assume that each of the videos had 2067 viewers, and that some of those viewers watched zero seconds of lesson 1/3/2 and/or 1/3/3.

If we believe Interpretation 1 is correct, then we should change our computation on the EV3 side so that we get a single percentage computed based on all of the videos in a set as follows:

Average % Viewed =

$$\frac{\sum Watch\ Times\ of\ subvideos}{Max(Views\ of\ subvideos) \times \sum Lengths\ of\ subvideos}$$

Ranks

	MOOC	N	Mean Rank	Sum of Ranks
Avg % viewed	NXT	7	7.14	50.00
	EV3	20	16.40	328.00
	Total	27		

Test Statistics[a]

	Avg % viewed
Mann-Whitney U	22.000
Wilcoxon W	50.000
Z	-2.656
Asymp. Sig. (2-tailed)	.008
Exact Sig. [2*(1-tailed Sig.)]	.006[b]

a. Grouping Variable: MOOC

b. Not corrected for ties.

Table 6. Mann-Whitney U Test on one-to-many data

Ranks

	MOOC	N	Mean Rank	Sum of Ranks
Avg % viewed	NXT	7	6.57	46.00
	EV3	7	8.43	59.00
	Total	14		

Test Statistics[a]

	Avg % viewed
Mann-Whitney U	18.000
Wilcoxon W	46.000
Z	-.831
Asymp. Sig. (2-tailed)	.406
Exact Sig. [2*(1-tailed Sig.)]	.456[b]

a. Grouping Variable: MOOC

b. Not corrected for ties.

Table 7. Mann-Whitney U Test on one-to-many data in which the many are "merged"

The result of running a Mann-Whitney U Test on these data are shown in Table 7 and suggests that there is no statistically significant (<0.05) or near-significant (<0.1) difference between the NXT and EV3 viewers viewing habits if we assume that this interpretation is correct.

Possible Interpretation 2: Different Types of Viewers

The previous interpretation sounds good, until you notice that the second and final data sets do not follow the same drop-off pattern. The second data set (1/8/1, 1/8/2, and 1/8/3) show more views of the middle video.

Why did this happen? Well, it seems easier to understand when we consider the titles of the three videos.[2]

- Week 1 Lesson 8 Video 1 EV3: Programming Preview - Should You Watch this Preview?
- Week 1 Lesson 8 Video 2 EV3: Programming Preview - The Preview
- Week 1 Lesson 8 Video 3 EV3: Programming Preview - Coming up next week

Perhaps viewers thought: Why waste time deciding whether you should watch a preview – just watch it! Similar conclusions might be made based on the titles of the final set of videos:

- Week 4 Lesson 5 Video 1 EV3: Ultrasonic Sensor - Introduction
- Week 4 Lesson 5 Video 2 EV3: Ultrasonic Sensor - Understanding How the Sensor Works
- Week 4 Lesson 5 Video 3 EV3: Ultrasonic Sensor - Using the Sensor

Perhaps this is simply a case of MOOC participants "skipping over the boring bits."

(Anecdotally) Confirming the Past

It feels to us as though the examples above may be a subset of Wilkowski et. al.'s "completers" [11] who simply want to pass but don't care to learn material "if it's not going to be on the test." There is also the possibility that some students (correctly) interpreted that those videos that got fewer views consisted of less ad lib time with the professor and more highly produced and scripted videos. This would be consistent with Thornton et. al.'s findings. [8]

FUTURE DIRECTIONS

While our data are exciting, we also recognize that our sample size is quite small. We are currently investigating different approaches to getting more granular data from YouTube. We can certainly get daily (and possibly even finer data) on our videos, and are hopeful we may be able to find a way to collect data in terms of viewers rather than in terms of time.

ACKNOWLEDGEMENTS

Many thanks to Patrick Thomas, Ph.D. of 1790 Analytics for his help with the data analysis. The NXT and EV3 MOOCs were made possible thanks to generous grants from *Google for Education*.

REFERENCES

1. Ezra Fishman. July 2016. How Long Should Your Next Video Be? *Retrieved January 2019 from* https://wistia.com/learn/marketing/optimal-video-length

2. Google | Open Online Education. *Retrieved February 2019 from* https://edu.google.com/openonline/course-builder/index.html

3. 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, USA, 41-50. DOI: https://doi-org/10.1145/2556325.2566239.

4. Power Searching with Google. *Retrieved February 2019 from* https://coursebuilder.withgoogle.com/

5. Educational Robots for Absolute Beginners – EV3 *retrieved February 2019 from* https://cs4hsev3robots.appspot.com/

6. Educational Robots for Absolute Beginners – NXT *retrieved February 2019 from* https://cs4hsrobots.appspot.com/

7. Chris Savage. December 2009. Does length matter? It does for video! *Retrieved February 2019 from* http://wistia.com/blog/does-length-matter-it-does-for-video.

8. Sera Thornton, Ceri Riley, and Mary Ellen Wiltrout. 2017. Criteria for Video Engagement in a Biology MOOC. *In Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale (L@S '17)*. ACM, New York, NY, USA, 291-294. DOI: https://doi.org/10.1145/3051457.3054007

9. Thales Teixeira, Michel Wedel, and Rik Pieters. "Emotion-Induced Engagement in Internet Video Advertisements." *Journal of Marketing Research* 49, no. 2 (2012): 144-59.

10. Watch time in YouTube Analytics. *Retrieved February 2019 from* https://support.google.com/youtube/answer/6299733

11. Julia Wilkowski, Amit Deutsch, and Daniel M. Russell. "Student skill and goal achievement in the mapping with google MOOC." In *Proceedings of the first ACM conference on Learning@ scale conference*, pp. 3-10. ACM, 2014.

[2] A full list of all of the titles of the videos in tables 3 and 4 can be found in Table 8 below.

One – to – One Video Titles	
EV3	**NXT**
Wk1 L6: Peanut Butter and Jelly Redux	Wk1 L6: Hardware, Software, PB&J - PB&J Reduxe
Wk2 L1: Introduction	Wk2 L1: Setup Overview
Wk2 L2: Battery Installation and Charging	Wk2 L2 EV3 Battery InstalL& Charging
Wk2 L7: Connecting the Robot to Your Computer	Wk2 L7: Connecting the Robot to the Computer
Wk3 L4: A Program Using Display & Sound Blocks	Wk3 L4: Sleepy Night: A Program Using Display & Sound Blocks
Wk3 L10: Potential Pitfalls 2	Wk 3 L9: More Pitfalls
Wk4 L6: Advanced Move Blocks	Wk 4 L4: Advanced Move Blocks
Wk5 L6: Wrapping Up & Next Steps	Wk 5 L6: Wrap up & Next Steps

One – to – Many Video Titles	
EV3	**NXT**
W1 L3: Teaching with Robots	W1 L3 P1: Teaching with Robots - Part 1: Introduction
	W1 L3 P2: Teaching with Robots: Part 2: Janet Moss
	W1 L3 P3: Teaching with Robots - Part 3: Teacher Perspective
W2 L3: Building your first robot	W2 L3 P1: Building 1st Robot: LEGO Robots
	W2 L3 P2: Building 1st Robot - Understanding LEGO Directions
W3 L1: More About the NXT Brick	W3 L1 P1: More about the EV3 Brick - Brick Buttons Overview
	W3 L1 P2: More about the EV3 Brick - More Button Details
	W3 L1 P3: More about the EV3 Brick - Exploring the buttons
W1 L8: Optional Preview of Robot Programming	W1 L8 P1: Programming Preview- Should You Watch this Preview?
	W1 L8 P2: Programming Preview- The Preview
	W1 L8 P3: Programming Preview- Coming up next week
W3 L8: Dancing Robots!	W3 L6 P1: Dancing Robots - Movement Example
	W3 L6 P2: Dancing Robots - The Move Steering Block
W3 L12: (Optional) More on Display Blocks	W3 L11 P1: More Display Block Details - Making Your Own Graphics
	W3 L11 P2: More Display Block Details - Simple Animations
	W3 L11 P3: More Display Block Details - Displaying Text
W4 L7: Using the Ultrasonic Sensor	W4 L5 P1: Ultrasonic Sensor - Introduction
	W4 L5 P2: Ultrasonic Sensor - Understanding How the Sensor Works
	W4 L5 P2: Ultrasonic Sensor - Using the Sensor

Table 8. A listing of the titles of all of the videos presented in this paper.