



## How Do English Language Learners Interact with Different Content Types in MOOC Videos?

Judith Uchidiuno<sup>1</sup> · Ken Koedinger<sup>1</sup> ·  
Jessica Hammer<sup>1</sup> · Evelyn Yarzebinski<sup>1</sup> ·  
Amy Ogan<sup>1</sup>

Published online: 21 November 2017

© International Artificial Intelligence in Education Society 2017

**Abstract** English Language Learners (ELLs) are a substantial portion of the students who enroll in MOOCs. In order to fulfill the promise of MOOCs – i.e., making higher education accessible to everyone with an internet connection – appropriate interventions should be offered to students who struggle with the language of course content. Through the analysis of clickstream log data gathered from two MOOC courses deployed on Coursera, Introduction to Psychology and Statistical Thermodynamics, we show that compared to native English speakers, ELL students have distinct behavioral patterns in how they engage with MOOC content including increased interaction with content that contains text, increased seeking away from content without visual support, and decreased video play rates. These patterns are expressed differently in response to different types of course content and domains. Our findings not only suggest more fine-grained methods for automatically identifying students who need language interventions, but also have further implications for the design of language support interventions and MOOC videos.

---

✉ Judith Uchidiuno  
jio@andrew.cmu.edu

Ken Koedinger  
koedinger@andrew.cmu.edu

Jessica Hammer  
hammerj@andrew.cmu.edu

Evelyn Yarzebinski  
eey2@andrew.cmu.edu

Amy Ogan  
aao@andrew.cmu.edu

<sup>1</sup> Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA

**Keywords** MOOCs · English language learners · MOOC behavioral analysis · ELL student identification · Language support interventions

## Introduction

Imagine a world where every student, regardless of economic or cultural background, has equal access to high quality education. This vision is not yet a reality, but Massively Open Online Courses (MOOCs) aim toward it. MOOCs provide access to college-level courses at scale, with some courses enrolling over 100,000 students (Online Course Report 2015), and encourage learners to connect and collaborate with others worldwide through discussion boards and other site features. While MOOCs hope to support all students' learning, they are currently not supporting all students equally. Language is one major challenge for MOOCs. A significant percentage of enrollees originate from non-English speaking countries (DeBoer et al. 2013), but a majority of MOOCs are deployed in English (MOOC List 2015). English Language Learners (ELLs) must therefore not only overcome the challenge of learning the course content, but also the challenge of learning it in a non-native language.

There are some efforts to make MOOCs accessible to ELL students by translating MOOCs to local languages (Coursera 2013; Khan Academy 2015). However, this strategy is expensive, and may not be scalable given the diversity of languages spoken by enrollees. It is also not an appropriate solution for every ELL student. Many ELL students deliberately seek out MOOCs delivered in English for a variety of reasons, including improving their career prospects, connecting to other English language speakers, and preparing for geographic mobility (Uchidiuno et al. 2016b). These students need language support interventions that help them accomplish their goals *within* English-language MOOCs.

Designing and deploying effective language support interventions requires that researchers are able to accurately identify students who need these interventions. While many interventions support ELL students without harming the performance of English-native students, others may be distracting to native speakers (Kim and Chang 2010; Silverman and Hines 2009). Prior research studies identify likely ELL students using demographic data and other proxies such as IP address (DeBoer et al. 2013; Guo and Reinecke 2014; Seaton et al. 2014b). Even with proxies that show better success identifying ELL students such as the default language of their web browsers (Uchidiuno et al. 2016b), these methods may only indicate students' native or preferred language, and do not give any specific insights to how supportive interventions might be designed, or if the students identified actually need language support.

During a MOOC, students spend the majority of their time watching and engaging with educational videos (Seaton et al. 2014a). Research studies have shown that the analysis of video clickstream logs in MOOCs can show distinct behavioral patterns and guide intervention design (Guo et al. 2014; Kim et al. 2014b; Kovacs 2016). However, to our knowledge, there are no studies that use these logs to identify ELL students struggling in MOOCs. Additionally, MOOC videos incorporate multiple types of content, such as an instructor speaking over text slides or conducting a demo. ELL students may find certain types of content particularly difficult (Chang and Read 2006; Renandya and Farrell 2010), but no studies have yet examined their behavioral logs

based on video content type. We therefore go beyond simply identifying these students to characterizing their behavioral interactions with different kinds of MOOC content.

To fill this gap, we analyzed clickstream logs from two different MOOCs – an Introduction to Psychology MOOC taken by 13,887 students that includes 47 videos, and a Statistical Thermodynamics MOOC taken by 2971 students that includes 26 videos. The Psychology MOOC videos featured the instructor lecturing over a series of text-based instructional slides. The Thermodynamics course, on the other hand, incorporated many different instructional strategies, including the instructor speaking without slides, displaying equations and charts, and conducting hands-on experiments. By selecting courses with different structures, we can determine to what extent ELL students' behavior is similar across MOOCs, and how it may interact with different kinds of MOOC content.

Using this data, we identified differences in clickstream behavior patterns between students *classified as* ELL (and henceforth referred to simply as ELL) or native-level English, based on browser language. We also examined how those patterns differed based on the video content type, such as text-based slides, equations, or figures; the domain of the content (physics or psychology); and the language difficulty of the content as determined by words per minute and the Coh-Metrix L2 readability score (Crossley et al. 2011; McNamara and Graesser 2012). These patterns reveal three underlying strategies that these ELL students use to manage challenging course content. Our research makes the following contributions to the Learning at Scale and AIED communities:

- We identify clickstream behavior patterns associated with ELL students who struggle to understand course video that reveal specific coping strategies.
- We characterize how similar coping strategies are expressed by different ELL clickstream patterns based on the content type displayed, the content domain, and language difficulty.
- We discuss the implication of these insights on the design of language support interventions and the production of video content in MOOCs.

## Background

### Demographics and ELL Participation in MOOCs

There are no official estimates of the number of ELL students who participate in MOOCs, but there is data that suggests that their participation is significant. Guo et al. analyzed MOOC data from 140,546 students from 196 countries and showed that the countries with the most certificate-earning students include Russia, Spain, and India; students from these countries made up almost 28% of the student population (Guo and Reinecke 2014). Similar studies show that participation in MOOCs is especially pronounced in Brazil, Russia, India, China, and South Africa (Dillahunt et al. 2014a, b; Snyder and Writer 2013). While country of origin may not accurately predict ELL status, as 80% of the students enrolled from these countries are from the wealthiest 6% of the population and the vast majority of MOOC participants already have a college degree (Dillahunt et al. 2014a, b; Snyder and Writer 2013), other

methods also confirm that ELL participation in MOOCs is significant. In the Introduction to Psychology course analyzed in this research study, 34% of the 13,887 students were categorized as ELL using their browser language as a proxy for identification (Uchidiuno et al. 2016a). Using the same analysis, approximately 32% of the 2971 students in the Statistical Thermodynamics course were categorized as ELL. Finally, preliminary analysis of a single Conversational English MOOC specifically targeted towards ELL students shows that over 60,500 students have enrolled since February of 2016. Taken together, this data shows that the numbers of ELL students in MOOCs are significant, and calls for research to identify whether these students have unique approaches, strategies, and needs while learning with MOOCs, and furthermore to understand how to support those needs.

### **Difficulties of Learning in a Foreign Language**

ELL students face considerably greater difficulty comprehending spoken English language than native English speakers, especially in situations where they cannot interact with the speaker (Chang and Read 2006; Renandya and Farrell 2010). A prevalent problem affecting listening comprehension for ELL students is speech rate (Renandya and Farrell 2010). While researchers have tried to address this issue by slowing down multimedia listening materials in second language classrooms, they have seen little success in improving comprehension (Derwing and Munro 2001; Hayati 2010). One reason may be that native speakers of a language subconsciously modify, drop, or add sounds, and additionally blend their words together, making it difficult for second language learners to recognize words and distinguish their boundaries (Renandya and Farrell 2010). This phenomenon is evident even among different dialects of English; research studies show that changing the English dialect of instructional materials to the local dialect of students brings about positive learning outcomes (Cutrell et al. 2013; Finkelstein et al. 2013). Additionally, ELL students particularly can feel apprehensive about speech because “it cannot be touched and held the way written text can” (Bacon 1989). Such difficulties in processing voice input, which are not present in reading, have been shown to hurt learning in ELL classrooms (Mayer et al. 2003). Because instructional videos in MOOCs also rely on voice input, they are a likely site of struggle for ELL students.

### **Identification of ELL Students and Their Needs**

Research studies focused on identifying ELL students’ skills in classroom settings have typically relied on students’ self-assessment of their speaking, writing, reading, and listening skills. A comprehensive literature review conducted by Ross et al. (Ross 1998) shows that students’ self-assessment in these individual skills are highly correlated with their placement scores in that area. However, the high attrition rate of MOOCs makes them a poor context for self-assessment. For example, only 24% of students filled out the demographic survey associated with the Introduction to Psychology course examined in this paper. In addition, advanced MOOCs contain technical terms that may not be reflected in a self-assessment of one’s general language skills.

Given these confounding factors, researchers have looked to other metrics to identify ELL students in MOOCs. One of the most popular methods of identification is the use

of students' IP address and demographic information (DeBoer et al. 2013; Guo and Reinecke 2014; Seaton et al. 2014b), while Uchidiuno et al. (2016a, b), showed that students' web browser language preferences are more predictive of their interactions in MOOCs than their IP addresses are. Browser language is the method we choose to classify learners in this study. However, while these metrics may be readily available to researchers, they do not provide insights to specific learning areas where ELLs struggle. In order to identify what scaffolds or interventions might best support these students (and to offer language support interventions only to students who actually need it), researchers must find better ways to identify behaviors that not only indicate struggle, but also provide insights on how to improve the design of MOOCs to meet ELL students' needs.

### **Inferring Behavioral Patterns from MOOC Clickstream Data**

A popular method of understanding student behavior in MOOCs is by analyzing video interaction clickstream logs. Logs are kept separately for each video, and contain timestamped information on when students play, pause, change the play rate of the video, skip to other parts of the video, or encounter an error. Guo et al. analyzed clickstream logs from 6.9 million video watching sessions and found that shorter and informal talking videos are more engaging and lead to lower dropout rates compared to high quality pre-recorded classroom lectures (Guo et al. 2014) (Kim et al. 2014b). Kim et al. (2014a) analyzed video clickstream data to identify areas of confusion in MOOC videos. This analysis then guided the design of a video player that highlighted areas of confusion on the video timeline, which helped students pay more attention to these areas and improved how they navigated. Finally, Uchidiuno et al. (2016a) analyzed clickstream data from a Psychology MOOC and found that students' web browser language predicted behavioral differences in video activity that were consistent with the ELL literature. These studies show that low-level clickstream data is invaluable for inferring a concrete understanding of students' needs in MOOCs, and can improve how MOOCs are designed.

### **Limitations to Clickstream Log Data Provided by the Coursera Platform**

Prior to discussing our methodology and reviewing our findings, it is important to mention the limitations of the dataset we gathered, which provides context for the types of analyses that were not possible to include due to the availability of data from the platform. First, for both courses, there was no indicator of students' use of closed captions. The platform has recently been modified to include this data, but cannot retroactively be recovered for courses completed prior to the change. Also, Coursera's opt-in demographic survey was removed in March 2015, meaning that students who created a new account on Coursera after that date (i.e., all students in the Thermodynamics MOOC) were not given the survey. As a result, we do not have demographic information for the students enrolled in that course. For the Psychology MOOC, only a quarter of the enrolled students volunteered to participate in the demographic survey (see Table 1). Finally, at the time the courses were completed, there was no log event to capture when a student reached the end of a video. A student who watched a video from beginning to end, without any other interactions or clicks, generates a single

**Table 1** Demographic information for Psych MOOC students – Metric for ELLs is “primary browser language is not English”

		ELL (33% of student population)	English
Profession	Student	45%	39%
	Industry Professional	13%	16%
	Research Scientist	5%	3%
	Academic/Professor	7%	4%
	Hobbyist	19%	19%
	None of the Above	11%	19%
Education	High School	10%	13%
	Bachelors	16%	17%
	Masters	15%	11%
	PhD	3%	2%
	Other	4%	8%
	Not Indicated	52%	49%
Age	Under 18	2%	4%
	18–29	60%	51%
	30–39	22%	21%
	40–49	10%	11%
	50–59	4%	8%
	60+	1%	3%
Gender	Female	63%	63%
	Male	37%	37%

‘play’ event logged at time 0.00. There are no other logs to indicate whether they watched the full video, abandoned the video by closing their browser, navigated to another page etc. As a result, it is impossible to determine with precision how much of the videos the students actually watched without making very generalized assumptions that make the findings difficult to interpret. Given these limitations, we focus our analysis and discussion on interaction logs that are captured based on the students’ interactions with the video player, rather than speculating on student behavior where the data was unavailable.

## Methods

We analyzed clickstream data for two MOOCs deployed on Coursera, focusing on interactions with video. To select our MOOC dataset, we first identified eight courses with available clickstream data; from that set, we selected two courses that represented different disciplines (psychology and physics) and that had student browser language recorded for each interaction. The first course selected was an Introduction to Psychology course with 47 videos, which was offered on Coursera starting March 25, 2013. This course has been studied previously in (Koedinger et al. 2015; Wang et al. n.d.). We

analyzed data from each interaction with video by all 13,887 enrolled students. Of the total number of students enrolled, approximately 33% were categorized as ELL based on their browser language preferences (Uchidiuno et al. 2016a). The top five English-speaking countries represented in the data were United States (31%), India (14%), Canada (5%), United Kingdom (5%), and Singapore (3%), while the top five countries categorized as ELL were Brazil (9%), Russia (7%), Greece (7%), China (6%), and Spain (6%).

We also analyzed data from a Statistical Thermodynamics course with 26 videos deployed on Coursera from April 1 to June 30, 2015. We analyzed data from all 2971 participants, of whom 32% were categorized as ELL based on browser language preference (Uchidiuno et al. 2016a). Given that this browser language metric is a proxy for inferring language ability, students who are likely ELL in these courses are referred to as ELL for the sake of simplicity.

### Course Content and Structure

The format of the Psychology MOOC, across the 47 videos, was typically text-based slides with verbal narration. Figure 1 shows a screenshot of this typical structure, in which the instructor lectured while the video displayed his image and text slides. However, one video was comprised entirely of the instructor performing a demonstration and was the only demo in the entire course; this was an outlier from the typical structure of videos in this MOOC, therefore, we removed this video from our analysis. The Statistical Thermodynamics course, however, was structured differently. Each video typically included a ‘Learning Objectives’ slide at the beginning of the course, and a ‘Summary’ slide at the end of the video. Only 11% of the videos consisted of the instructor lecturing over text-based slides. Instead, the vast majority of each video varied between the instructor talking without any text or figures for support, explaining equations, showing figures and charts, and conducting hands-on demonstrations.

In order to understand how student behavior varied across different kinds of content, we coded the videos for content type (see Table 2). Two researchers independently coded 5% of the videos. In an initial coding pass, they achieved 93% agreement on the codes; they discussed and iterated on the codes until 100% agreement was reached. Coding of the remaining videos was split among the coders.



**Fig. 1** Typical structure of “Introduction to Psychology” MOOC (left) vs “Statistical Thermodynamics” MOOC



**Table 2** Video content type codes – content percentage may sum to greater or less than 100% as multiple content types may be displayed simultaneously in the same time interval, and non-instructional content such as video introductions and long pauses were not coded

Code	Meaning	Stat. Therm. MOOC	Psych MOOC
Talking	Instructor talking, no support materials displayed	45.59%	7.13%
Equation	An equation was displayed	28.48%	–
Text	Text content was displayed	10.53%	69.19%
Figure	Picture or image was displayed	5.94%	11.29%
Chart	Graph or table was displayed	3.52%	8.60%
Demo	Hands-on experiment	1.29%	–

As the instructors spoke for the entire length of each video, codes primarily reflected visual differences. For example, “Talking” reflected the instructor speaking with no visual supporting materials, while “Text” meant that a text slide was displayed. If more than one visual was displayed, e.g. an equation and a chart, the video was annotated as both kinds of content. Timestamps were logged for the beginning and end of each section of video.

### Variables of Interest

For each MOOC, we computed a number of video interaction behavioral variables (see Table 3). While most of these computations were straightforward, seeking required some additional analysis. Coursera logs only indicate where students seek *to*. If a student presses ‘play’ at the 2 s mark (listening at a normal play rate), watches until 22 s, and seeks back to 15 s, the clickstream logs only indicate the ‘play’ event at 2 s, and ‘seek’ at 15 s. In order to determine where the students sought *from*, as well as whether the seek was backward or forward, the difference in clock time (adjusted by the play rate) was compared to the seek time – for the example given above, the student’s watch time of 20 s on the clock is added to the ‘play’ time at 2 s to determine that the student was at the 22 s mark before they sought, and compared to the 15 s ‘seek’ point to determine that the student sought backward 7 s.

**Table 3** Behavioral variables of interaction with a video

Play	Play button is pressed on content type
Pause	Pause button is pressed on content type
Seek Back From	Student seeks back to earlier point in video from content type
Seek Forward From	Student seeks forward to later point in video from content type
Seek Back To	Student seeks back from later point in video to content type
Seek Forward To	Student seeks forward from earlier point in video to content type
Slow Down	Number of times ‘reduce play rate’ button was clicked
Speed Up	Number of times ‘increase play rate’ button was clicked
Play Rate	Average play rate for the video



Using timestamps, we connected interaction events to what content type was present in the video at the time. For example, for a seek event, we identified the video codes associated with the start point and end point, allowing us to create seeking sequences, such as a student seeking back from Text to Equation, Text to Text, etc.

## Analytic Methods

For both courses, we ran a cumulative odds ordinal logistic regression with proportional odds to determine whether the frequency of each video behavior (play, pause, speed up, slow down, seek forward and seek backward) during each content type (text, charts, figures, talking, demos, and equations), increased or reduced the probability of a student being classified as ELL. We ran a different logistic regression model for each content type, such as text and charts (6 models total), to enable us to compare the behavior for that content type across all videos, rather than limit the statistical inference to only those set of videos that contained all 6 content types. In the logistic regression models for each content type, the video ID was included as a random effect, because our exploratory data analysis revealed that students' behavior appeared to have some dependency on the specific content of the video being watched. There were several videos that showed significant effects for either reducing or increasing the overall intercept. However, in this research study, we are not investigating holistic features of videos that affect student interaction. Therefore, although their effect is included in the regression models to reveal the video-independent differences between ELL and English students, their estimates are not included in the regression tables in this paper.

## Findings

### ELL Students Interact Differently with Content by Domain

As shown in Table 4, ELL students showed broad interaction patterns that differed from those of English natives. For example, each one-unit increase in the average play rate across a full video decreases the log odds and probability of being categorized as ELL for both courses. Figure 2 shows the average play rate between English and ELL students. However, the expression of other common interaction patterns differed depending on the domain (see Fig. 3).

In the Thermodynamics MOOC, ELL students reduced the speech rate to listen to the content at a slower pace: every one-unit increase in the slow down count (number of times reduce speed button was clicked) significantly increased the likelihood of being categorized as ELL (particularly for 'Text', 'Figures', and 'Equations'). In the Psychology MOOC, ELL students did slow down some content types such as 'Talking' and 'Text'. More commonly, though, every one-unit increase in the pause count across content types in the Psychology course significantly increased the log odds and the probability of being categorized as an ELL student.

This difference in strategies between course types, especially as they interact with text content, may relate to different purposes that 'Text' serves in both courses. In the

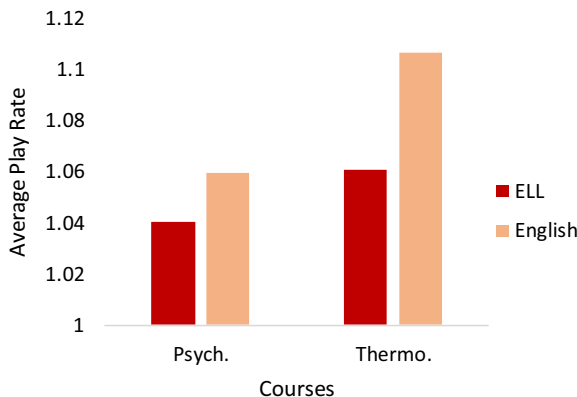
**Table 4** Logistic regression results summary on the video behaviors/content type on the probability of being categorized as ELL

		Psychology			Thermodynamics		
		Estimate	Pr(> z )	sig	Estimate	Pr(> z )	sig
Text	English (Intercept)	-0.712789	<2e-16	***	-1.16281	5.97E-16	***
	Play	-0.012341	0.089794	.	-0.05069	0.2978	
	Pause	0.036714	1.63E-07	***	0.0625	0.1989	
	Seek Back From	-0.010742	0.430722		-0.02116	0.51	
	Seek Forward From	-0.015503	0.472654		-0.03773	0.458	
	Seek Back To	0.003389	0.804601		0.02465	0.4477	
	Seek Forward To	0.020637	0.338705		0.05727	0.238	
	Slow Down	0.083446	8.63E-05	***	0.2575	0.0118	*
	Speed Up	-0.078578	0.00012	***	-0.06991	0.4646	
Talking	English (Intercept)	-0.704774	<2e-16	***	-1.14157	3.36E-15	***
	Play	0.006573	0.749386		0.02871	0.41477	
	Pause	0.05863	0.002031	**	-0.03639	0.29641	
	Seek Back From	0.016352	0.532178		0.06224	0.00874	**
	Seek Forward From	0.111759	3.49E-05	***	0.11585	0.00161	**
	Seek Back To	-0.039965	0.132454		-0.06416	0.00853	**
	Seek Forward To	-0.06081	0.030004	*	-0.10387	0.00801	**
	Slow Down	0.057648	0.096838	.	0.07129	0.1698	
	Speed Up	-0.056824	0.132279		-0.07757	0.15618	
Charts	English (Intercept)	-0.73652	<2e-16	***	-1.179661	5.58E-16	***
	Play	0.004817	0.884264		-0.146517	0.1166	
	Pause	0.08767	0.004485	**	0.147692	0.1144	
	Seek Back From	-0.027807	0.278737		-0.040324	0.5045	
	Seek Forward From	0.013072	0.766214		-0.009976	0.9194	
	Seek Back To	0.010048	0.696965		0.04917	0.427	
	Seek Forward To	-0.001429	0.973645		0.065742	0.4845	
	Slow Down	0.026414	0.697832		0.22796	0.2965	
	Speed Up	0.030877	0.660861		-0.203582	0.2936	
Figures	English (Intercept)	-0.7417904	<2e-16	***	-1.15745	9.28E-16	***
	Play	0.0016726	0.9386		-0.07016	0.32128	
	Pause	0.0452209	0.03139	*	0.05334	0.44568	
	Seek Back From	0.0207058	0.34942		0.01479	0.77751	
	Seek Forward From	-0.011752	0.71402		0.0234	0.78471	
	Seek Back To	-0.0367997	0.10113		-0.01476	0.78764	
	Seek Forward To	0.0223611	0.481		0.01236	0.88567	
	Slow Down	0.0633846	0.2183		0.39943	0.00455	**
	Speed Up	-0.086302	0.104		-0.12693	0.46899	
Demo	English (Intercept)	-	-		-1.157064	8.91E-16	***
	Play	-	-		0.041574	0.8437	
	Pause	-	-		0.066573	0.7425	

**Table 4** (continued)

		Psychology			Thermodynamics		
		Estimate	Pr(> z )	sig	Estimate	Pr(> z )	sig
Equations	Seek Back From	–	–		–0.249494	0.117	
	Seek Forward From	–	–		0.177995	0.1252	
	Seek Back To	–	–		0.17869	0.2049	
	Seek Forward To	–	–		–0.172927	0.2474	
	Slow Down	–	–		0.164049	0.2529	
	Speed Up	–	–		–0.136164	0.4719	
	English (Intercept)				–0.839302	<2e-16	***
	Play	–	–		0.034768	0.383	
	Pause	–	–		–0.03452	0.3835	
	Seek Back From	–	–		–0.007512	0.7448	
	Seek Forward From	–	–		–0.013458	0.7243	
	Seek Back To	–	–		0.005486	0.812	
	Seek Forward To	–	–		0.033337	0.3797	
	Slow Down	–	–		0.155154	0.0453	*
	Speed Up	–	–		–1.16281	5.97E-16	***
Average play rate	English (Intercept)	–8.82	0.903		–0.038263	0.8389	
	Average Play Rate	–0.74	<2e-16	***	–1.02429	<2e-16	***

Psych MOOC, almost 70% of the course time is spent looking at text information, unlike the 10.5% in the Thermodynamics MOOC. Figure 4 shows the difference in the nature of the content that is disseminated through text in both courses. In the Psych MOOC, the vast majority of information that students must comprehend and are tested on is communicated via text, unlike in the Thermodynamics MOOC, where text only appears as ‘Learning Objectives’ at the beginning of each video, and ‘Summary’ at the end of each video.



**Fig. 2** Average play rate between English and ELL students on both courses

	TEXT		TALKING		CHARTS		FIGURES		DEMOS		EQUATIONS	
	Psych.	Therm.	Psych.	Therm.	Psych.	Therm.	Psych.	Therm.	Psych.	Therm.	Psych.	Therm.
Pause	ELLs +	=	ELLs +	=	ELLs +	=	ELLs +	=	NA	=	NA	=
Seek Back From	=	=	=	ELLs +	=	=	=	=	NA	=	NA	=
Seek Forward From	=	=	ELLs +	ELLs +	=	=	=	=	NA	=	NA	=
Seek Back To	=	=	=	ELLs -	=	=	=	=	NA	=	NA	=
Seek Forward To	=	=	ELLs -	ELLs -	=	=	=	=	NA	=	NA	=
Slow Down	ELLs +	ELLs +	ELLs +	=	=	=	ELLs +	=	NA	=	NA	ELLs +
Speed Up	ELLs -	=	=	=	=	=	=	=	NA	=	NA	ELLs -

Fig. 3 Summary of regression results table grouped by behavior, content type, and course

### ELL Behavior with Specific Content Types

#### Interaction with Visual Aids: Charts, Figures, and Equations

The most typical ELL behavioral differences seen across domains (pauses for Psychology, slow down for Thermodynamics) were generally present in content types that included visual aids, such as charts, figures, and equations. The results of the logistic regression estimates (Table 4) show that in the Psych MOOC, every one-unit increase in the pause count on both charts and figures significantly increases the log odds and probability of being categorized as ELL: ELL students press pause on both figures and charts more often. In the Thermodynamics MOOC, every one-unit increase in the slow down count on Figures significantly increases the log odds and probability of being categorized as ELL. However, this behavior is not consistent with interaction specifically with charts in the Thermodynamics MOOC, where no differences were observed. Equations were only present in the Thermodynamics course, and, as with other visual aids, we found that ELL students slowed down significantly more and sped up much less than non-ELL students.

Examples of charts are shown in Fig. 5 and figures in Fig. 6. ELL students in general spent more time on Figures in both courses even though the subject matter of the courses is different. We reviewed the Figures in both courses. In the Psychology MOOC, some images are iconic; however, we found many other images, similar to Fig. 6, that rely heavily on English language skills for interpretation. ELL students may be interacting with these figures similar to the way they are interacting with text

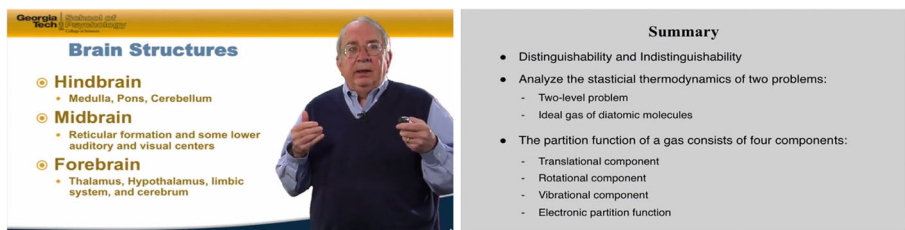


Fig. 4 Example of text content from Psych (left) and Thermodynamics (right)

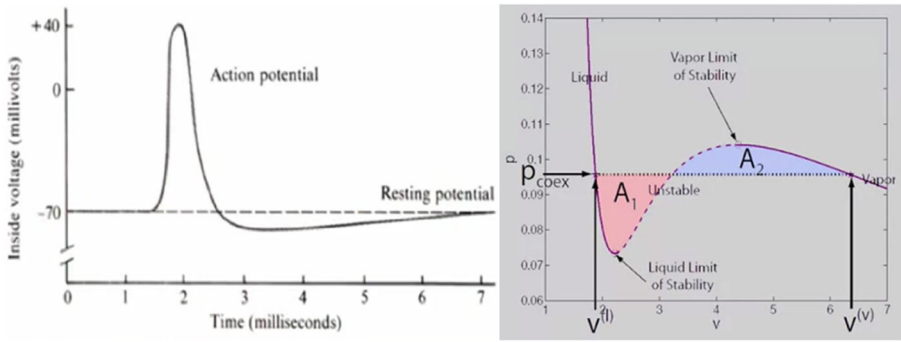


Fig. 5 Example of a chart from Psych (left) and Thermodynamics (right)

information. We reviewed all the Figures on the Psych MOOC and found that 41% of them contain text information that may cause an ELL speaker to take extra time to process the information.

*Seeking Behavior in the Absence of Visual Aids*

In both courses, sections were coded as “Talking” when the instructor was talking without any visual aids (figures, charts, equations, or text) in the background. Unlike these other content types where the instructor is talking with text or images as the background information, ELL students have no alternative channels to help them comprehend the course content, if they have trouble understanding the language.

We found that that ELLs are generally seeking away from “Talking” sections, and are less likely to seek towards it. In both MOOCs, every one-unit increase in seeking forward, and in the Thermodynamics MOOC, additionally every one-unit increase in seeking backward, from “Talking” sections to other content types increases the log odds and probability of being categorized as ELL. In both MOOCs, every one-unit increase in seeking forward and seeking backward to “Talking” sections decreases the log odds and probability of being categorized as ELL. If the ELL students enrolled in the course have trouble comprehending spoken language (without supporting information), this finding supports the notion that they are supporting their understanding of “Talking” sections with other content types that better support their learning needs.

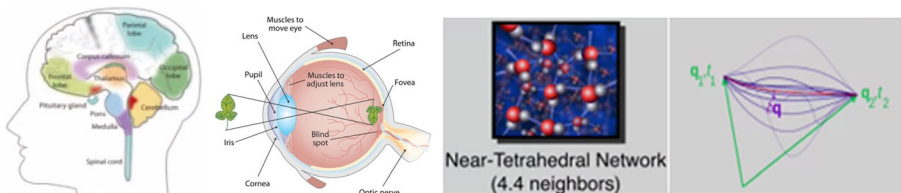


Fig. 6 Examples of figures from the Psychology (left) and Thermodynamics (right) MOOC

*Effects of Language Difficulty and Speed on Behavior with Text*

To better understand why ELL students are seeking away from the instructor talking without support, but pausing more on text content, we ran a linear regression model to determine how the difficulty (using the Coh-Metrix L2 Reading Index (Crossley et al. 2011; McNamara and Graesser 2012)) and the speed (using number of words per minute – WPM) of each video predicts specific prominent behaviors that ELL students exhibit. In particular, we looked at play rate and pauses in the Psych MOOC, and play rate and slowdown in the Thermodynamics course, which were the more prominent differences between ELLs and non-ELLs.

The Coh-Metrix L2 Reading Index is a metric that assigns a difficulty score to a reading passage (for our purposes, the video transcripts), and has been shown to rate text – including transcripts – more accurately than other popular readability scores (Crossley et al. 2011). The metric scores reading passages based on the word characteristics, sentence characteristics, and discourse relationships between ideas in text (McNamara and Graesser 2012). We also calculated the WPM for each of the videos in both MOOCs. Table 5 shows descriptive statistics for both courses.

For the CohMetrix variable, a higher score corresponds to an easier transcript. In the Psych MOOC, we found that the lower the CohMetrix score (indicating a more difficult transcript), the greater the number of pauses. Also, we found that the more difficult the transcript, the slower the play rate was; ELLs were significantly more likely to exhibit those pausing and slowing behaviors than non-ELL students (Table 6).

These results are as expected, as students may be more likely to listen slower, or perform more dictionary lookups for language that is hard for them to understand. In the Thermodynamics MOOCs, on the other hand, the CohMetrix score had no effect on the number of times students slowed down, or on the average play rate. This is possibly a result of the students being more familiar with the difficult words (e.g. viscosity) present in the course, as they may have learned them as part of their prior domain knowledge (regardless of their language of instruction).

For the WPM variable, a higher score corresponds to the instructor speaking at a faster rate. In the Psych MOOC, we found that the higher the WPM score, the slower the play rate was overall, as would be expected. However, the higher the WPM score, the fewer the number of pauses by students (especially ELLs). This result was contrary to our expectations, and calls for a need to supplement these findings with qualitative observational data to interpret the meaning as it seems counterintuitive. In the Thermodynamics MOOC, we found that the higher the WPM, the greater the number of

**Table 5** Summary of Coh-Metrix L2 readability scores (above; higher score indicates easier passage) and WPM (below) for both Thermodynamics and Psych MOOCs

	<i>mean</i>	<i>sd</i>	<i>min</i>	<i>max</i>
Psych.	20.88	2.75	14.11	26.08
Thermo.	16.06	2.64	11.69	21.59
	<i>mean</i>	<i>sd</i>	<i>min</i>	<i>max</i>
Psych.	155.02	20.60	121	243.83
Thermo.	132.17	10.45	116.72	149.46

**Table 6** Linear regression results summary for the effect of the transcript complexity and WPM on significant ELL student behaviors

	Parameter	Estimate	Pr(> t )	sig
Psych pause on text	(Intercept)	4.554682	< 2e-16	***
	CohMetrixScore	-0.0566892	4.68E-11	***
	WPM	-0.0131122	< 2e-16	***
	ELL	3.3282021	< 2e-16	***
	CohMetrixScore:ELL	-0.0616936	5.28E-05	***
	WPM:ELL	-0.0072946	1.08E-05	***
Psych avg. play rate	(Intercept)	1.07E + 00	<2e - 16	***
	CohMetrixScore	4.53E-04	0.0118	*
	WPM	-1.92E-04	<2e-16	***
	ELL	-1.84E-02	0.014	*
	CohMetrixScore:ELL	1.91E-04	0.5449	
	WPM:ELL	-2.33E-05	0.5211	
Therm slow down on text	(Intercept)	-0.1103596	0.0101	*
	CohMetrixScore	0.0015242	0.2365	
	WPM	0.0007662	0.0232	*
	ELL	-0.0130975	0.8704	
	CohMetrixScore:ELL	0.0046479	0.0555	.
	WPM:ELL	-0.0003669	0.5611	
Therm avg. play rate	(Intercept)	1.3072031	< 2e-16	***
	CohMetrixScore	-0.0012325	0.252	
	WPM	-0.0014357	3.58E-07	***
	ELL	-0.0801022	0.232	
	CohMetrixScore:ELL	0.0007764	0.702	
	WPM:ELL	0.0001827	0.729	

times that students slowed down the video, and consequently, the lower the play rate of the video. One possible hypothesis for future investigation is that ELL students may especially get overwhelmed by the content as the instructor speaks faster, and rather than pause the video, they are more likely to adjust to the faster speech rate by slowing it down.

### *Consistency in Behaviors with Demonstrations*

Demonstrations were only present in the Thermodynamics MOOC. Our results show that ELL students behave similarly to non-ELL students during demonstrations. Given that these hand-on experiments are conducted in physical space and present visually observable phenomena that are ephemeral in time, it is not surprising that ELL students may be able to comprehend the information as well as non-ELL students without relying on any auditory information.



## Variation in Behavior for Course-Completers vs. Non-completers

While knowing these behavioral differences such as pausing exist is important, this does not give us an indication of whether enacting them helps or hinders students in engaging with the course. To determine whether these behaviors differed for those who persisted through the courses, we ran a comparison “Text” and “Talking” regression model on only the students who completed the courses. In the Psych MOOC, students who took the final exam were considered course completers. However, we could not perform this same analysis on the Thermodynamics MOOC as less than 0.6% of the students completed the course.

The results (see Table 7) show that ELLs who completed the course paused even more than the general ELL population on text-based content, which supports the theory that they pause the videos to use other language aids such as dictionaries to help with their comprehension. For the “Talking” only sections, course completers also paused more, similar to their behavior on text content, but did not seek away from it like the general ELL population. This may be an indication that students who persist through the course understand spoken language enough that they treat it like text content. More importantly, this may indicate that ELLs who exhibit this seeking behavior in sections without visual aids are the ones who are in dire need of language interventions, and are unlikely to complete the course without assistance.

## Discussion and Implications

Our results demonstrate distinct differences in how ELL students and native English-speaking students engage with MOOC content. These differences can be used both to adaptively identify students who may be struggling with the language content, and to provide adaptive support where it is needed most. Our findings suggest that ELL learners deploy a range of strategies to address the challenges of learning English-language content, and these strategies may be a strong indicator of where support is

**Table 7** Summary of logistic regression of video behaviors/content type, probability of being categorized as ELL for students who completed the course

	TEXT			TALKING		
	Estimate	Pr(> z )	sig	Estimate	Pr(> z )	sig
English (Intercept)	-0.645734	< 2e-16	***	-0.6423428	< 2e-16	***
Play	-0.024006	0.025325	*	-0.0991143	0.031309	*
Pause	0.049468	1.34E-06	***	0.1638037	1.71E-04	***
Seek Back From	-0.042004	0.093396		-0.0365706	0.428908	
Seek Forward From	-0.01802	0.621811		0.0734143	0.106087	
Seek Back To	0.030861	0.217825		-0.004321	0.927158	
Seek Forward To	0.017824	0.624238		-0.0515696	0.298521	
Slow Down	0.049495	0.166346		-0.0406816	0.446647	
Speed Up	-0.076823	0.02683	*	-0.0478358	0.381124	

needed. For example, ELL students listen to course videos at a slower play rate. While slowing down the pace of listening materials is a popular strategy used by ELL students and their instructors in classroom contexts, there is little to no evidence of the effectiveness of this strategy to improve listening comprehension for ELL students (Derwing and Munro 2001; Hayati 2010). Similarly, ELL learners seek away from video segments that include instructor speech unaccompanied by textual or visual aids. We hypothesize that they are looking for alternate ways of understanding the instructor, which indicates a commitment to learning the material. However, the same material may not be presented elsewhere in the video.

We observe that ELL students behave differently across different video content types, in ways that suggest differences in their learning strategies. Based on these behavioral differences, we propose three analytic categories for video content in MOOCs, as follows: 1) narration with no visual supports, 2) visual supports that are language-dependent, such as text or text-heavy figures, charts, and equations, and 3) visual supports that are language-independent, such as demos. Each offers different challenges in supporting ELLs. For content types that are language-independent, ELLs appear to behave similarly to English natives; therefore, such content types may not require extra consideration for ELLs. ELLs may even be able to use these language-agnostic content types as additional cues to enhance their comprehension. However, further research is needed to determine whether they *benefit equally* as well as *behaving similarly* to English native speakers.

When language-dependent visual supports are provided, ELLs may have an increased cognitive load as they must both read and hear information in their non-native language. Additionally, if they are using transcripts to support their comprehension, they now must split their attention between two sources of text. Moreover, as textual content is often a summary of the instructor's points rather than a verbatim transcript, hearing and reading different words at the same time may cause difficulty for ELLs' comprehension. ELLs who interact with content in this category may require additional support to equalize their MOOC experience.

Finally, ELLs engage least with narration that has no visual supports. Narrated content may reduce the *split-attention effect* and avoid cognitive overload, according to significant prior research (Chandler and Sweller 1992; Mayer and Moreno 2003; Plass et al. 2003). Instructional design recommendations based on this research therefore include “presenting words as narration,” so that “the words are processed in the verbal channel” (Mayer and Moreno 2003). Our findings suggest that these recommendations may benefit English natives, but negatively affect ELL students, who avoid video segments of this type and therefore are unlikely to master content delivered without visual support.

These insights into ELL learner behavior can be adopted by MOOC course designers. For example, course designers might ensure that every point made without visual support is backed up with visual aids elsewhere, ideally ones that are not language-dependent. However, they also illustrate the need for adaptive design of video interactions on MOOC platforms. Verbal-only video content may be most effective for English native speakers because of the split-attention effect, but least effective for ELL students who avoid watching it. Rather than removing content that is beneficial to English natives, ELL students could be directed to alternate video segments or even to other types of learning activities.

## Limitations

There are several limitations in our research study. First, our findings highlight the behavioral differences between ELLs and native English speakers; however, we do not address the impact of these behavioral differences on students learning or performance in MOOCs. Therefore, our design recommendations may improve ELLs experience and reduce their cognitive load in MOOCs, but we cannot hypothesize on how they affect the students' learning outcomes. We acknowledge that although browser language may be more indicative of students' language ability (compared to IP addresses), it is still a language proxy, and may not capture all the students that may benefit from language support interventions. There may also be other content types in MOOCs that are not represented in the courses we analyzed, therefore, there is an opportunity to perform similar types of analysis on additional content types to reveal more nuances in the ways that ELLs interact with MOOCs. Finally, the content domain and structural differences in the two courses we analyzed make it more difficult to compare the behavioral differences even across similar content types in both courses. Analyzing these clickstream behaviors among courses that are similar in domain and structure may help to reinforce the content types that would most greatly benefit from language support interventions.

## Conclusion

In this paper, we have used clickstream data from two different MOOCs to identify behavioral differences between ELL students and native English speakers in their interactions with course video. Additionally, we have shown that ELL students use coherent strategies to respond to video content. They slow the play rate for the entire video, they pause the video or slow down speech rates to have more time to look at text (text-based slides or text-heavy figures), and they seek through the video to find supporting content when no text is available. Our data suggests that ELL students can be accurately identified using a combination of annotated videos and clickstream data, which would be more precise than proxies such as IP address or browser language. Additionally, we provide recommendations for designing course videos to support ELL student strategies, such as including adaptivity for the way content is presented in MOOCs. With a combination of targeted interventions and better course design, we envision a future where ELL students are better supported in MOOCs than they are today.

## References

- Bacon, S. M. (1989). Listening for real in the foreign-language Classroom1. *Foreign Language Annals*, 22(6), 543–550.
- Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62(2), 233–246.
- Chang, A. C.-S., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly*, 40(2), 375–397. <https://doi.org/10.2307/40264527>.

- Coursera (2013). A new partnership to bring Coursera to the hundreds of millions of Chinese learners across the world. Retrieved October 20, 2015, from <https://blog.coursera.org/a-new-partnership-to-bring-coursera-to-the/>.
- Crossley, S. A., Allen, D. B., & McNamara, D. S. (2011). Text readability and intuitive simplification: a comparison of readability formulas. *Reading in a Foreign Language*, 23(1), 86.
- Cutrell, E., Bala, S., Bansal, C., Cross, A., Datha, N., John, A., ... et al. (2013). *Massively empowered classroom: Enhancing technical education in india*. Microsoft Research Technical Report. Retrieved from [http://research-srv.microsoft.com/pubs/207167/MEC-TechReport\\_v3.pdf](http://research-srv.microsoft.com/pubs/207167/MEC-TechReport_v3.pdf).
- DeBoer, J., Stump, G. S., Seaton, D., & Breslow, L. (2013). Diversity in MOOC students' backgrounds and behaviors in relationship to performance in 6.002 x. In *Proceedings of the Sixth Learning International Networks Consortium Conference*. Retrieved from <http://tll.mit.edu/sites/default/files/library/LINC%20%2713.pdf>.
- Derwing, T., & Munro, M. J. (2001). What speaking rates do non-native listeners prefer? *Applied Linguistics*, 22(3), 324–337. <https://doi.org/10.1093/applin/22.3.324>.
- Dillahunt, T., Chen, B., & Teasley, S. (2014a). Model thinking: demographics and performance of mooc students unable to afford a formal education. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 145–146). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=2567851>.
- Dillahunt, T. R., Wang, B. Z., & Teasley, S. (2014b). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *The International Review of Research in Open and Distributed Learning*, 15(5). Retrieved from <http://www.irrodl.org/index.php/irrodl/article/view/1841>.
- Finkelstein, S., Yarzebinski, E., Vaughn, C., Ogan, A., & Cassell, J. (2013). The effects of culturally congruent educational technologies on student achievement. In *Artificial intelligence in education* (pp. 493–502). Springer. Retrieved from [http://link.springer.com/chapter/10.1007/978-3-642-39112-5\\_50](http://link.springer.com/chapter/10.1007/978-3-642-39112-5_50).
- Guo, P. J., & Reinecke, K. (2014). Demographic differences in how students navigate through MOOCs. *L@S 2014, March 4–5, 2014*. Retrieved from [http://pgbovine.net/publications/edX-MOOC-demographics-and-navigation\\_LAS-2014.pdf](http://pgbovine.net/publications/edX-MOOC-demographics-and-navigation_LAS-2014.pdf).
- 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. doi:<https://doi.org/10.1145/2556325.2566239>.
- Hayati, A. (2010). The effect of speech rate on listening comprehension of EFL learners. *Creative Education*, 01(02), 107–114. <https://doi.org/10.4236/ce.2010.12016>.
- Khan Academy (2015). Carlos Slim's Foundation puts its weight behind Khan Academy. Retrieved October 20, 2015, from <http://www.forbes.com/sites/kerryadolan/2013/01/14/carlos-slims-foundation-puts-its-weight-behind-khan-academy/>.
- Kim, S., & Chang, M. (2010). Computer games for the math achievement of diverse students. *Educational Technology & Society*, 13(3), 224–232.
- Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014a). Understanding in-video dropouts and interaction peaks in online lecture videos. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 31–40). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=2566237>.
- Kim, J., Guo, P. J., Cai, C. J., Li, S.-W. D., Gajos, K. Z., & Miller, R. C. (2014b). Data-driven interaction techniques for improving navigation of educational videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology* (pp. 563–572). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=2647389>.
- Koedinger, K. R., Kim, J., Jia, J. Z., McLaughlin, E. A., & Bier, N. L. (2015). Learning is not a spectator sport: Doing is better than watching for learning from a MOOC. In *Proceedings of the second (2015) ACM conference on learning @ scale* (pp. 111–120). New York: ACM. doi:<https://doi.org/10.1145/2724660.2724681>.
- Kovacs, G. (2016). Effects of in-video quizzes on MOOC lecture viewing. In *Proceedings of the third (2016) ACM conference on Learning@ Scale* (pp. 31–40). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=2876041>.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52.
- Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: role of speaker's voice. *Journal of Educational Psychology*, 95(2), 419.
- McNamara, D. S., & Graesser, A. C. (2012). Coh-Metrix: An automated tool for theoretical and applied natural language processing. *Applied natural language processing and content analysis: identification, investigation, and resolution*. Hershey: IGI Global. Retrieved from [ftp://129.219.222.66/Publish/pdf/McNamara\\_Graesser\\_Coh-Metrix.pdf](ftp://129.219.222.66/Publish/pdf/McNamara_Graesser_Coh-Metrix.pdf).

- MOOC List (2015). Find MOOCs by languages. Retrieved October 23, 2015, from <https://www.mooc-list.com/languages?static=true>.
- Online Course Report (2015). The 50 most popular MOOCs of all time. Retrieved April 4, 2017, from <http://www.onlinecoursereport.com/the-50-most-popular-moocs-of-all-time/>.
- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior*, *19*(2), 221–243.
- Renandya, W. A., & Farrell, T. S. (2010). Teacher, the tape is too fast! Extensive listening in ELT. *ELT Journal*, *65*(1):52–59.
- Ross, S. (1998). Self-assessment in second language testing: a meta-analysis and analysis of experiential factors. *Language Testing*, *15*(1), 1–20.
- Seaton, D. T., Nesterko, S., Mullaney, T., Reich, J., Ho, A., & Chuang, I. (2014a). Characterizing video use in the catalogue of MITx MOOCs. *Proceedings of the European MOOC Stakeholders Summit*. Lausanne: PAU Education, 140–146.
- Seaton, D. T., Bergner, Y., Chuang, I., Mitros, P., & Pritchard, D. E. (2014b). Who does what in a massive open online course? *Communications of the ACM*, *57*(4), 58–65. <https://doi.org/10.1145/2500876>.
- Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology*, *101*(2), 305.
- Snyder, S., & Writer, I. S. (2013). Study shows who's taking open online college courses. Retrieved October 14, 2015, from [http://articles.philly.com/2013-11-22/news/44330667\\_1\\_moocs-massive-open-online-courses-california-based-online-education-company](http://articles.philly.com/2013-11-22/news/44330667_1_moocs-massive-open-online-courses-california-based-online-education-company).
- Uchidiuno, J., Ogan, A., Yarzebinski, E., & Hammer, J. (2016a). Understanding ESL students' motivations to increase MOOC accessibility. In *Proceedings of the third (2016) ACM conference on Learning@ Scale* (pp. 169–172). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=2893398>.
- Uchidiuno, J., Ogan, A., Koedinger, K. R., Yarzebinski, E., & Hammer, J. (2016b). Browser language preferences as a metric for identifying ESL Speakers in MOOCs. In *Proceedings of the third (2016) ACM conference on Learning@ Scale* (pp. 277–280). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=2893433>.
- Wang, X., Yang, D., Wen, M., Koedinger, K., & Rosé, C. P. (n.d.). Investigating how student's cognitive behavior in MOOC discussion forums affect learning gains. Retrieved from [http://www.educationaldatamining.org/EDM2015/uploads/papers/paper\\_89.pdf](http://www.educationaldatamining.org/EDM2015/uploads/papers/paper_89.pdf).