THIS DOCUMENT CONTAINS THE AUTHOR'S ACCEPTED MANUSCRIPT. THE PUBLISHED MANUSCRIPT (ACCEPTED FOR PUBLICATION IN COMPUTERS AND EDUCATION ON JUNE 13, 2020) IS AVAILABLE FROM ELSEVIER, DOI:
10.1016/j.compedu.2020.103961

Exploring Student and Teacher Usage Patterns Associated with Student Attrition in an Open Educational Resource-Supported Online Learning Platform

Dongho Kim
Sungkyunkwan University

Yongseok Lee
University of Florida

Walter L. Leite
University of Florida

A. Corinne Huggins-Manley
University of Florida

Corresponding author: Dongho Kim; 51206 Hoam Hall, Sungkyunkwan University, Seoul, South Korea 32611; +82-2-760-0534; dkim3@skku.edu

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C160004 to the University of Florida. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Abstract

Online learning platforms integrating open educational resources (OERs) are increasingly adopted in secondary education as supplemental resources for teaching and learning. However, students report difficulties sustaining their engagement because of the self-paced nature of OER-supported learning environments. We noted that little attention has been paid to factors related to student perseverance and attrition in the learning environment. Little knowledge about these factors prevents discussion on how to promote OERs as pedagogical tools that complement the formal school curriculum. To address this research gap, we analyzed student- and teacher-level usage data, and demographic information. The purpose was to explore student- and teacher-level factors associated with the duration of student usage in Algebra Nation, an OER-supported online learning platform adopted in many secondary schools. The results revealed that at the student level, student engagement with video lectures, self-assessment, social tools, and additional videos relevant to solved test items significantly predicted student usage duration. At the teacher level, teachers' use of teacher resources was positively associated with student usage duration. An additional analysis of student and teacher total usage time indicated that, compared with heavy users, light and medium users were more likely to discontinue their engagement in the long term if their teachers did not use the platform. Based on our findings, we provide recommendations for promoting student engagement with OER-supported online learning in secondary education contexts.

Introduction

Open educational resources (OERs) are educational materials publicly available under an open license that permits their use and re-use by any user (Mishra, 2017). OERs are increasingly integrated in secondary educational settings as supplemental resources for teaching and learning. Thanks to open licenses that allow for open access and adaptation, OERs are emerging as next-generation pedagogical tools (Pérez-Paredes, Ordoñana Guillamón, & Aguado Jiménez, 2018).

Accordingly, a growing number of online learning platforms integrate OERs to offer teachers and students access to high-quality materials at no cost (McShane, 2017). The OER-supported online learning platform Algebra Nation has been adopted in secondary schools in seven states across the US. Educators, students, and parents are granted access to free online resources for learning Algebra, which are actively used in and outside the classroom (Leite, Cetin-Berber, Huggins-Manley, Collier, & Beal, 2019). Other non-profit online learning platforms such as Khan Academy, OpenStax, and OER Commons also provide various OERs including open lectures, lesson plans, and quizzes for secondary teachers and students. We are now witnessing the pedagogical opportunities these platforms provide in secondary education.

Despite the pedagogical benefits of OER-supported online learning platforms, students have difficulties sustaining their consistent engagement with the platforms because they are expected to complete a large portion of learning tasks and assignments outside the classroom in a self-paced manner. For example, viewing lecture videos and completing the required online quizzes after class time as directed by their teachers requires students to be self-regulated learners. According to Borup and Stevens (2016), inconsistent student engagement and high attrition rates are critical issues to address for the successful integration of OERs into formal education contexts. The pedagogical advantages we could reap from OERs have not been maximized, as shown by remarkably high student attrition and discontinuation (Burd, Smith, & Reisman, 2015). In fact, student attrition in OER-supported learning platforms integrated in formal settings may unfold in different ways. Some students begin learning by opening resources on the platforms but fail to maintain their engagement when facing difficult problems (Marple, Jaquet, Laudone, Sewell, & Liepmann, 2019). Other students demonstrate consistently low engagement with the platforms and stop using them without further exploration (Muir, 2014). The reason for students' inconsistent engagement and attrition may be attributed to various student- and teacher-level factors related to the adoption of the platforms in classrooms (Cargile & Harkness, 2015). Therefore, we argue that the pedagogical innovations brought by OER-supported online learning platforms to education are being disrupted by a lack of understanding of how educators and students use OERs. In the era of open pedagogy, more efforts should be made to understand how to successfully utilize non-formal educational resources in formal education.

Previous research has been limited to exploring the opportunities and challenges of OERs for formal education. As such, little attention has focused on which factors may lead to student attrition in OER-supported learning environments. Knowing these factors would provide valuable insights into how to promote OERs as pedagogical tools that complement the formal school curriculum. To address the current research gap, we analyzed multimodal data on student academic performance, demographic information, and usage traces in Algebra Nation, an OER-supported online learning platform. Specifically, we explored the association of teacher and student usage patterns of Algebra Nation, as represented by activity logs in the system, with student usage duration. We focused on student usage duration rather than learning outcomes because low student engagement and attrition have been identified as barriers to using OERs to complement the formal school curriculum (Brahimi & Sarirete, 2015). To our knowledge, no prior study has adopted a data-driven approach to predict student usage duration in OER-supported learning environments while considering both teacher- and student-level usage patterns. This exploratory study paves the path for future research on interventions and refined OERs that may increase student engagement with and duration on OER-supported online learning platforms. In addition, our findings can inform teacher practices when imbedding OERs into teaching. Finally, this study is expected to provide direction for implementing OERs in secondary education settings.

Related work

Integration of open educational resources in formal secondary education

Butcher and Wilson-Strydom (2008) define OERs as educational resources publicly available for free use and adaptation. In response to the global trend that emphasizes open access to educational content to ensure equity in education, OERs have contributed to the decentralization of education (Kim, Lee, & Park, 2019). High-quality educational resources and services previously provided to only a limited number of students enrolled in a few prestigious institutions have become more accessible via open learning platforms such as massive open online courses (MOOC) (Hew & Cheung, 2014). Learners benefit from participating in such non-formal learning opportunities for personal development (Kim, 2018). The opportunity for OER-supported non-formal education allows learners to continue pursuing their studies out of the formal school curriculum.

Another notable trend is the integration of online learning platforms that incorporate OERs (e.g., Algebra Nation) into formal secondary education (Brahimi & Sarirete, 2015). That is, while formal curricula are implemented in schools, the platforms can provide added support for teachers and students in various ways. For example, OERs can serve as a primary source of content learning prior to in-class lectures and activities (Brahimi & Sarirete, 2015). In addition to the benefit of OERs in enhancing flipped classrooms, teachers can use OERs provided by various platforms to improve the quality of their teaching materials and student assessment methods. Through exploring OERs provided by other experts or peer teachers, teachers can also ensure they meet appropriate pedagogical standards (Kim, 2018). Furthermore, the adaptability of OERs supported by open licenses enables teachers to reuse and remix existing teaching resources (e.g., workbooks) (Wiley, Bliss, & McEwen, 2014).

Prior studies reported the positive outcomes of OERs for teaching practices in formal education. Wiley, Hilton III, Ellington, and Hall (2012), for example, focused on the impact of digital open textbooks in terms of cost savings and flexible teaching. Their study, which involved 20 secondary school teachers and their 3,900 students, showed that using open textbooks greatly reduced costs compared to adopting traditional textbooks without compromising the quality of student learning. The authors highlighted the potential outcomes of open digital textbooks as a new form of curriculum material that can be flexibly adapted to enhance teaching and learning in formal education contexts. Beyond cost savings, Weller, De Los Arcos, Farrow, Pitt, and McAndrew (2015) examined the comprehensive impacts of OERs adopted in various formal education sectors including K–12 and post-secondary institutions. Using a mixed-methods approach, they found that the use of OERs in formal settings improved student performance and satisfaction. In addition to these positive impacts, OERs were also associated with long-term outcomes such as educators' reflection, student retention, and equity in education. These positive long-term outcomes were confirmed in Ngimwa and Wilson's (2012) finding that educators recognized OERs as contributing to the overall education systems in African countries, while fulfilling the specific needs of teachers and students in different schools.

On the other hand, students use OERs to refine their understanding of what they learned in class. For example, they review additional materials about a new topic and take quizzes (Brahimi & Sarirete, 2015) outside the classroom. There is ample evidence that OERs can motivate students to take initiative for their own learning (Zheng, Rosson, Shih, & Carroll, 2015). Another notable advantage is that OERs allow students to study at their own pace (Friesen & Wihak, 2013). For example, slow learners can use OERs to review what they missed during class. Likewise, OERs can be used for high achievers willing to further their learning beyond class. To enable students to benefit from the flexibility of OERs, a growing number of secondary school districts have started encouraging the adoption of OER platforms to strengthen their formal curricula (Wiley et al., 2012).

Empirical studies revealed the positive correlation between students' use of OERs, their learning experiences, and achievement. Heo and Choi (2014) reported a case study in which students studied with free educational videos in a seventh-grade math class. They found that students who watched more videos achieved higher scores on the mid-term exam. Wiley, Webb, Weston, and Tonks (2017) analyzed performance data collected from students who completed school assignments that required them to create and share OERs. The authors found that students who participated in a self-creation OER activity received significantly higher grades than those who did not participate in the activity. The authors also highlighted the need to integrate OERs into the formal school curriculum to be able to offer student-centered activities in classrooms.

Beyond the abovementioned correlational studies, Robinson, Fischer, Wiley, and Hilton III (2014) revealed the causal effect of OERs on student achievement through a quasi-experimental design approach based on propensity score matching techniques. Based on data on 43 teachers and their 4,183 students, their findings indicated that students who used open digital textbooks in high school science classes achieved better grades than those who did not use them. Noteworthy is that open digital textbooks were successfully integrated into the existing curricula at the participating schools through adaptions made by teachers.

Student learning strategies in OER-supported learning environments

The successful use of OER-supported online learning platforms cannot be assumed, because this new type of learning environment is inherently unstructured and uncontrolled compared to traditional settings where teachers can directly provide their students with pedagogical support. The high attrition rates reported in open learning environments such as MOOCs are attributed to the frustrations students experience as independent learners (Moreno-Marcos et al., 2019). It has been shown that because of the self-directed nature of the open learning environment, students who learn with OERs often fail to sustain their engagement (Engle, Mankoff, & Carbrey, 2015). Without immediate guidance from instructors, students are likely to be confused about how to accomplish learning tasks (Kim et al., 2019). According to Engle et al. (2015), students' successful use of OERs depends on their ability to direct their own learning; however, the self-directed learning strategies required in open learning environments may differ from those needed in traditional classroom settings. For example, Kim et al. (2019) noted that many learners are concerned about the quality of OERs. Furthermore, their social strategies to interact with other online learners were positively associated with their continued use of OERs. As learners tend to consider OERs supplementary resources (Hu, Li, Li, & Huang, 2015), it is important to examine how they use them in conjunction with other primary resources to enhance their learning.

In response to the need for instructional support tailored to OER-supported learning environments, key online behaviors associated with students' perseverance and successful learning have been examined. Schaeffer and Konetes (2010) highlighted student disengagement and attrition in secondary online courses as a critical issue amidst the rapid growth of online learning in the education sector. Their review of previous literature on student engagement in online programs revealed that successful online students have goal-oriented learning strategies and a willingness to take risks. Prior studies reported consistent findings regarding formal learning in secondary schools. For example, Light and Pierson (2014) observed 25 math lessons using OERs from Khan Academy and interviewed participating teachers and students to identify how the OERs integrated into classroom contexts supported students. They confirmed the benefits of OER-supported online learning platforms as a tool to support students' self-regulated learning. They highlighted that Khan Academy, the platform used in the study, had many features such as a monitoring tool and adaptable assessment materials. Furthermore, students engaged in math using self-regulated learning strategies such as planning, monitoring, and self-evaluation. Similarly, Kim et al. (2019) related students' self-regulated learning strategies such as utilizing extra online resources to their future intentions to use OERs. Wong et al. (2020) verified the effect of OERs on students' self-regulated learning in elementary education. They found that i-Classroom, a mobile learning platform utilizing OERs, helped students develop key self-regulated learning factors such as learning motivation, learning management, and self-monitoring skills.

The current study

We are now witnessing the integration of OER-supported online learning platforms in formal school curricula. The new educational phenomenon is rapidly blurring the boundaries between formal and non-formal education. However, engaging students in the online portion of learning remains a challenge. Empirical studies consistently reveal high attrition rates and low student engagement in online learning environments. Evidence supports the need to examine the student activity and learning patterns associated with student perseverance in OER-supported online learning platforms. However, as shown in the results reviewed earlier, studies have only recently focused on factors associated with these high attrition rates. In response to the need for large-scale empirical studies on student perseverance in OER-supported online learning environments, we extend the previous literature by exploring predictors of student attrition. Knowledge of these predictors is important because they can enable interventions to reduce student dropout. In this study, we address the research question, "Which student- and teacher-level usage patterns are associated with student usage duration in Algebra Nation, an OER-supported online learning platform, after controlling for other factors?"

Methods

Research context

The research context of this study was Algebra Nation, an online tutoring platform designed to support teachers and students in teaching and learning algebra. Launched in Florida as an online learning platform modularized by OERs for Algebra 1 in 2012, the platform has been expanded to other states. The data used in this study were collected from teachers and students who used the system between August 2017 and July 2018 in Florida. Algebra Nation provides teachers and students with OERs that can be adapted for their own teaching and learning purposes. For example, based on their needs, teachers and students can adapt the practice test items and worksheets attached to each video lecture. Furthermore, students can sign up and use Algebra Nation regardless of whether their teachers use it.

In Algebra Nation, students have access to lecture videos in each module and practice quizzes to check their learning progress. After taking the ten-question quizzes, students are prompted to watch solution videos or review lecture videos on the main topic covered in the quiz. Despite the recommended sequence of learning, students can watch videos without taking quizzes, or answer quizzes without watching videos. Furthermore, Algebra Nation contains a Student Wall, which is an online space that allows students to ask study experts assigned to the space questions. Students can also see experts' answers to other students' questions.

For teacher users, the Teacher Area in Algebra Nation contains a dashboard they can use to download relevant teaching materials (e.g., worksheets) and monitor students' progress. Teacher Wall is another space in the Teacher Area. Separate from the Student Wall, it provides a space for teachers to share ideas with other teachers.

Participants

From an initial dataset that contained information on all 107,104 students enrolled in Algebra Nation, we excluded 10,586 students who had used the platform for less than one day, since the number of days of use was employed as a dependent time variable in our survival model. The number of students was further reduced to 77,300 after removing students and teachers whose log records in the assessment pages (i.e., Test Yourself) were not available. We also excluded 924 students whose usage time was 3 standard deviations away from the mean. The same outlier deletion mechanism was applied to 34 teachers and their 1,599 students. We deleted outliers because extreme cases observed in online learning systems are often caused by technology failure (e.g., unstable Internet connection) or unintentional user behavior such as leaving a browser open after learning (Kovanovic et al., 2015). We also confirmed with the Algebra Nation development team that as in other online learning platforms, extreme log records could be generated by unexpected errors or unintended use in the Algebra Nation system. For the same reason, we removed cases that had missing values, rather than performing imputation to avoid estimating biased values, because there were not many missing values. After listwise deletion, we included 74,056 students and 1,858 teachers in the final dataset.

Measures

The dependent variable was the duration of student usage of Algebra Nation. This was coded in days, since the first usage and censoring variable was a binary indicator where "1" indicated that the student stopped using Algebra Nation or the study time ran out, and "0" indicated otherwise. The student-level usage predictors included eight variables that consisted of counts of student activity related to videos, quizzes, wall posts, and total usage. The six teacher-level usage predictors consisted of counts of student activities in the Teacher Area of Algebra Nation.

This study focused on the association between student and teacher usage patterns in an OER-supported online learning environment and students' duration (time spent) in the environment. However, we assumed that the prediction of student usage duration is influenced by other factors such as the school curriculum, which is shaped by school district policies; teacher initiatives; and student social economic status. Therefore, we included in the survival model control variables such as school district affiliation, student race, free lunch status, and teacher usage duration of Algebra Nation to more precisely and reliably estimate the predictive power of student and teacher usage factors. We used school district as a variable rather than school because some schools had only one participating teacher. We also considered the number of days between the beginning of the fall semester and the first day each student started using Algebra Nation. All student demographic information was obtained from the Florida Department of Education and linked with their log traces using their unique identifiers collected in Algebra Nation. Note that we could not capture teachers' practices in the classroom in terms of how they adaptively used the platform. For example, some teachers may have taught their students using materials printed from the online materials. It was also impossible to track specific features in the Teacher Area; therefore, we only analyzed teachers' access to the area as a whole. Table 1 presents the definitions and descriptions of each variable.

Table 1
Variables for the survival model

Category	Name	Description	Possible Range / Value
Demographic	Grade	Grade	Year 6–10
	Free lunch	Free lunch	No = 0, Yes = 1
	Race	Race	Dummy with White as reference group (Possible groups: Hispanic/Latino, White, Black, Asian, Pacific Islander, Multiracial, American Indian/Alaskan Native)
	Gender	Gender	Female = 0, Male = 1
Control variable	Prior math	Prior Florida Standards Assessment mathematics score (: FSA Math)	260–393 points
	Teacher duration	Teacher usage duration	0–365 days
	Student start date	Days between the first day of student use of Algebra Nation and August 1	0–354 days
Student-level predictors	Student_Lecture	Number of video lectures student watched	0–640 videos
	Student_Pausing	Number of times student used the pause button on the video player	0–322,249 times
	Student_Seeking	Number of times student used the seek button on the video player	0–33,154 times
	Student_Test	Number of test items student solved	0–3,949 items
	Student_Test solution video	Number of test solution videos student watched after taking tests	0–1,757 videos
	Student_Relevant lecture review	Number of video lectures relevant to solved test items student reviewed after taking tests	0–164 videos

	Student_Wall	Number of times student accessed Student Wall	0–3,121 times
	Student_Total usage time	Total amount of time student spent on Algebra Nation	0.35–66.54 hours
Teacher-level predictors	Teacher_Lecture	Number of video lectures teacher watched	0–393 videos
	Teacher_Test	Number of test items teacher solved	0–2,640 items
	Teacher_Test solution video	Number of test solution videos teacher watched after taking tests	0–82 videos
	Teacher_Relevant lecture	Number of video lectures relevant to solved test items teacher reviewed after taking tests	0–11 videos
	Teacher_Area	Number of times teacher accessed Teacher Area	1–169 times
	Teacher_Total usage time	Total amount of time teacher spent on Algebra Nation	0–37.63 hours

Data analysis

To analyze the multilevel data collected from 74,056 students nested within 1,858 teachers, we conducted multilevel parametric survival analyses based on the Weibull distribution. Similar to other survival models based on different distributions, the dependent variable for our model was a right-censored response (Crowther, 2017). We tracked participating students' and teachers' usage of Algebra Nation from August 1 to July 30. As participating students began using Algebra Nation at different times throughout the calendar year, we controlled for the time (days) between the first day of possible use (August 1). This assisted our analysis because student usage duration in Algebra Nation could be influenced by when they or their teachers started using the platform. For example, some teachers might have discouraged students from using it until they finished the first chapter of the course textbook, while others might have encouraged students to use the platform whenever they wanted. Furthermore, students may have started using it regardless of their teachers' preferences (if any) prior to the academic year. Thus, we used the entire year as a timeframe rather than the academic year, controlled for start time as mentioned, and controlled for teacher usage duration (days).

Survival models use the occurrence of an event as a binary variable and the time-to-event as a dependent variable (Hox, 2010). In this study, our long-format data indicated when students stopped using Algebra Nation (i.e., occurrence of an event, coded as "1"), and until the event occurred, the event variable was recorded as "0." Our survival analysis used the occurrence of an event as a target event and usage duration as a dependent variable. Therefore, the results provided the hazard function indicating the likelihood that participating students would stop using Algebra Nation. We used a Weibull distribution-based model because it has been shown to yield unbiased estimations with rare occurrences of the target event. Methods relying on the standard logit distribution do not properly capture the error distribution (Asmussen, Binswanger, & Højgaard, 2000). Overall, a proportional hazard function using the Weibull distribution helps to obtain unbiased estimates of the risk of occurrence of a target event.

We tested three hierarchical multilevel parametric survival models (Austin, 2017). First, an unconditional model was tested to obtain an intraclass correlation (ICC) representing the amount of variance in the outcome variable explained by teacher-level variance. Second, the student-level predictor model (i.e., random-intercept model; Snijders & Bosker, 2012) was tested to examine student-level predictors while controlling for other covariates. For the final model, we added a random slope of student total usage, and the interaction between student and teacher total usage time on hazard probability. We controlled for teacher total usage time as a predictor at the teacher level, reflecting our hypothesis that teachers' use of Algebra Nation had varying impacts on the baseline risk. We performed group-mean centering for all student-level predictors and grand-mean centering for all teacher-level predictors to distinguish individual from contextual-level relationships (Algina & Swaminathan, 2011).

To examine the model fit of the three hierarchical models, we used the likelihood ratio χ^2 test, Akaike information criterion (AIC), and Bayesian information criterion (BIC) fit indices to indicate the relative fit between models. For interpretation purposes, we rescaled the continuous predictors dividing all usage variables by ten because the base unit of analysis of each variable with a large range (e.g., total usage time) could result in unnoticeable changes in the hazard ratio. In this study, the hazard ratio for each subject indicated the probability that students stopped using Algebra Nation within the one-year timeframe (Hox, 2010). The proportional hazards final mixed-effects model that used a linear link function for specifying the relationship was defined as:

$$\log[h_{\mathrm{ij}}(\mathrm{t})] = \log[h_0(\mathrm{t})] + \beta_{1j}x_{1ij} + \gamma_1 Z_{1j} + \delta_1(x_{1ij} \times Z_{1j}) + \sum_{\mathrm{k}=2}^{\mathrm{K}} \beta_{kj}\, x_{kij} + \sum_{p=2}^{P} \gamma_p\, z_{pj} + \mu_{0j} + x_{1ij}\mu_{1j}\,, \quad (1)$$

where $h_0(\mathrm{t})$ is the baseline hazard function of the Weibull distribution, x_{1ij} is student total usage time, β_{1j} indicates the degree of change in the hazard probability per increase in student total usage time, and Z_{1j} represents teachers' Algebra Nation usage time. The interaction term coefficient (δ_1) represents the extent that teacher total usage time moderates the relationship between the degree of student total usage time and usage duration. At the student level, controlled covariates denoted by x_{kij} include student individual background variables and learning activities such as the number of watched videos and completed tests. At the teacher level, covariates are denoted by z_{pj}, where p is the number of teacher-level predictors except for the teacher total usage time. School districts were dummy-coded and included in the model as fixed effects (Allison, 2009) to control for district-level effects. We also assumed that the two random effect factors followed the multivariate normal distribution such that

$$\mu_{0j}, \mu_{1j} \sim MVN(0, \Sigma), \Sigma = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}. \quad (2)$$

Results

Descriptive statistics

Table 2 shows the descriptive statistics for the demographic variables. The correlation table in Appendix A also shows descriptive statistics for all variables other than the demographic variables.

Table 2.
Descriptive statistics of demographic variables

	Variable	N	Mean	S.D.	Min	Max
Student demographic	Grade	74,056	8.22	0.70	1.00	5.00
	Free lunch	74,056	0.51	0.50	0.00	1.00
	Hispanic/Latino	74,056	0.32	0.47	0.00	1.00
	American Indian / Alaskan Native	74,056	0.03	0.18	0.00	1.00
	Asian	74,056	0.06	0.23	0.00	1.00
	Black	74,056	0.21	0.41	0.00	1.00
	Pacific Islander	74,056	0.01	0.09	0.00	1.00
	White	74,056	0.74	0.44	0.00	1.00
	Multiracial	74,056	0.03	0.18	0.00	1.00
	Gender (female = 0, male = 1)	74,056	0.48	0.50	0.00	1.00

Multilevel survival analysis

Table 3 provides the results of the three hierarchical multilevel survival analyses. For the unconditional model, ICC was estimated at .083[1], indicating that 8.3% of the variance of student usage duration of Algebra Nation can be explained by teacher-level variance, and 91.7% of the variance was at the student level. The grand-mean of the hazard ratio in the unconditional model was statistically significant (hazard ratio (HR) = 0.002, $p < 0.00$).

Because the total usage time variables represented the sum of time spent on other activities at the student and teacher level, we tested another model using only student and total usage time, excluding the other student- and teacher-level usage variables, to determine the degree to which the other behavior variables reduced the weights of the total usage time variables. The results indicated that the coefficient of the teacher total usage time variable was significant (HR: 0.76, $p = .00$). Similarly, the student total usage time variable showed a higher coefficient (HR: 0.35, $p = .00$). However, we chose to include all other usage variables in the final model because the teacher total usage time variable was primarily used as a control variable, and including other usage variables did not cause critical multicollinearity issues such as the reversed direction of the coefficients of predictors. For the analysis of the tolerance and variance inflation factor (VIF), all variables had a VIF of less than 10 and tolerance of more than 0.1, indicating acceptable multicollinearity (Hair, Anderson, Tatham, & Black, 1995). In fact, the number of videos watched by students and teachers were the only variables strongly correlated with the total usage time variables (see Appendix A).

Of the three hierarchical models, the final model best fit the data, as indicated by the smallest AIC and BIC, and the statistically significant improvement over the other two models according to the likelihood ratio chi-square test. Specifically, the student-level predictor model had a better fit than the unconditional model, as indicated by the likelihood ratio chi-square test ($\chi^2(86) = 55268.67$, $P < 0.00$), but the final model yielded a better fit than the student-level predictor model ($\chi^2(9) = 1277.48$, $P < 0.00$).

[1] As we use the multilevel proportional hazard (PH) survival analysis with a Weibull distribution, the variance of the level-1 (student) residual was calculated by the equation, $\frac{\pi^2}{(6\times p^2)}$ where p is the ancillary parameter ("ln$p$" in Table 3) of the Weibull distribution (Rodríguez, 2010). The calculation of the ICCs used the information of the calculated variance of level-1 residuals and variance of level-2 (teacher) random intercepts (μ_{0j}) in the unconditional model.

Table 3.

Results of the three hierarchical multilevel survival models

		Unconditional Model		Students-level Predictor Model		Final Model	
	Variable	HR (SE)	[95% CI]	HR (SE)	[95% CI]	HR (SE)	[95% CI]
	Intercept	0.00***	[0.00,0.00]	3.62e-06**	[2.53e-06,5.20e-06]	2.78e-06**	[1.94e-06,4.00e-06]
Demographic	Grade	-	-	1.07(0.00)**	[1.05,1.09]	1.08(0.01)**	[1.06,1.10]
	Free lunch	-	-	1.00(0.00)	[0.98,1.02]	1.00(0.00)	[0.98,1.01]
	Hispanic	-	-	1.00(0.01)	[0.98,1.02]	0.99(0.01)	[0.97,1.01]
	American Indian / Alaskan Native	-	-	1.01(0.03)	[0.95,1.08]	1.02(0.03)	[0.95,1.08]
	Asian	-	-	0.96(0.03)	[0.91,1.03]	0.98(0.03)	[0.92,1.04]
	Black	-	-	0.93(0.03)*	[0.89,0.99]	0.94(0.03)*	[0.89,0.99]
	Pacific Islander	-	-	1.03(0.04)	[0.94,1.13]	1.04(0.05)	[0.94,1.14]
	Multiracial	-	-	1.06(0.04)	[0.98,1.13]	1.05(0.03)	[0.98,1.13]
	Gender	-	-	1.00(0.00)	[0.99,1.02]	1.01(0.00)	[0.99,1.02]
Control variable	Prior math	-	-	1.00(0.00)**	[1.00,1.01]	1.01(0.00)**	[1.00,1.02]
	Teacher duration	-	-	-	-	0.98(0.00)**	[0.98,0.99]
	Student start date	-	-	1.02(0.00)**	[1.02,1.02]	1.02(0.00)**	[1.02,1.02]
Student-level predictors	Student_Lecture	-	-	0.98(0.00)**	[0.97,0.98]	0.98(0.00)**	[0.98,0.99]
	Student_Pausing	-	-	1.00(0.00)	[1.00,1.00]	1.00(0.00)	[1.00,1.00]
	Student_Seeking	-	-	0.99(0.00)**	[0.99,1.00]	0.99(0.00)*	[1.00,1.00]
	Student_Test	-	-	0.99(0.00)	[0.99,1.00]	0.99(0.00)*	[0.99,1.00]
	Student_Test solution video	-	-	0.99(0.00)	[0.99,1.00]	1.00(0.00)	[0.99,1.00]
	Student_Relevant lecture review	-	-	0.92(0.02)**	[0.87,0.96]	0.94(0.02)*	[0.90,0.99]
	Student_Wall	-	-	0.98(0.00)**	[0.97,0.98]	0.98(0.00)**	[0.97,0.98]
	Student_Total usage time (A)	-	-	0.83(0.02)**	[0.77,0.88]	0.45(0.02)**	[0.41,0.50]
Teacher-level predictors	Teacher_Lecture	-	-	-	-	1.01(0.01)	[0.99,1.04]
	Teacher_Test	-	-	-	-	1.00(0.00)	[1.00,1.01]
	Teacher_Test solution video	-	-	-	-	1.06(0.06)	[0.93,1.19]
	Teacher_Relevant lecture review	-	-	-	-	0.99(0.44)	[0.42,2.37]
	Teacher_Area	-	-	-	-	0.98(0.00)**	[0.98,0.99]
	Teacher_Total usage time (B)	-	-	-	-	0.79(0.12)	[0.60,1.06]
Interaction	(A) × (B)	-	-	-	-	1.26(0.00)**	[1.13,1.39]
	ln$p$	0.529		0.940		0.956	
	Var(μ_{0j})	0.539		1.845		1.069	
	Var(μ_{1j})	-		-		1.774	
	Deviance (df)	-433448.6 (3)		-405814.3 (88)		-405175.6 (97)	

$\chi^2(df)$	-	55268.62(85)**	1277.50(9)**
AIC	866903.3	811804.7	810545.2
BIC	866930.9	812615.4	811438.8

*Note. Lecture videos: Test items: Test solution videos *: p < 0.05, **: p<0.01*

In the final model, six student-level predictors of the hazard ratio were statistically significant: Student_Lecture (HR: 0.98, $p < .00$), Student_Seeking (HR: 0.99, $p < .00$), Student_Test (HR: 0.99, $p = .00$), Student_Relevant lecture review (HR: 0.94, $p = .02$), Student_Wall (HR: 0.98, $p < .00$), and Student_Total usage time (HR: 0.45, $p < .00$). For each predictor, students with higher values on the predictor were associated with less risk of discontinuing their use of Algebra Nation. The results can be interpreted as follows, bearing in mind that the following statements are based on statistically controlling other variables in the final model. For students who watched ten more videos, used the video seeking button ten more times, and accessed the Student Wall ten more times, the probability of leaving Algebra Nation decreased by 1 to 2 percentage points compared to students who did not do these actions. Furthermore, students who reviewed ten more videos relevant to solved test items were associated with a 6% decrease in hazard ratio, while those spending 10 more hours in Algebra Nation were associated with a 55% decrease in hazard ratio. At the teacher level, only the number of times they accessed the Teacher Area was a significant predictor of the hazard ratio (HR: 0.98, $p = .00$). Here, teachers who accessed the Teacher Area 10 more times were associated with a 2% lower hazard ratio.

The interaction between student and teacher total usage time was statistically significant (HR: 1.26, $p < 0.00$). For interpretation purposes, Table 4 shows the original coefficients from the final model of two predictors and their interaction, rather than the hazard ratios as shown in Table 3. Furthermore, in Figure 1, we discretized student total usage time such that an equal number of students (24,687) were assigned to each of the three groups (i.e., heavy, medium, and light users) representing different relatively defined levels of student total usage time. The original continuous distribution of student total usage time was highly right-skewed. After discretization, the light user group was defined as students who spent 0–1.46 hours per week in Algebra Nation, the medium user group spent 1.46–4.29 hours per week, and the heavy user group spent 4.29–67.89 hours per week using the platform. This discretization approach allowed the interaction graph (Figure 1) to show how the predicted hazard ratio differed for the three groups of students according to teacher total usage time.

Table 4.
Transform hazard ratio to coefficients (student, teacher total usage time, and their interaction)

Final Model				
Variable	Coefficient	Hazard ratio	SE	[95% CI]
Student total usage time	-0.79***	0.45	0.05	[-0.89, -0.69]
Teacher total usage time	-0.22	0.79	0.15	[-0.51, .059]
Student total usage time×Teacher total usage time	0.23***	1.26	0.05	[0.13, 0.33]

For students whose teacher total usage time was 0 (no usage), a 10-hour increase in student total usage time with all other variables held constant was associated with a hazard ratio equal to *exp* (-0.79×1) = 0.45. For students whose teacher total usage time was 3 hours, a 10-hour increase in student total usage time was associated with a hazard ratio equal to *exp* (-0.79×1 + 0.23×3) = 0.90.

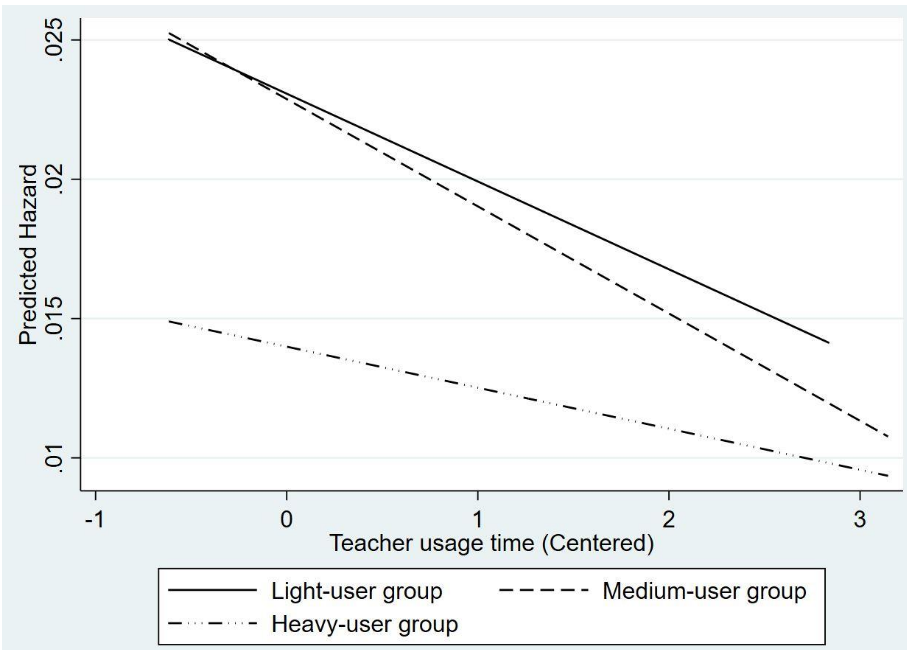

Figure 1. Interaction of student and teacher usage time on predicted hazard

The interpretation of Table 4 corresponds to the predicted hazard trajectories of the three student groups shown in Figure 1. In Figure 1, the y-axis represents the predicted hazard obtained from the final model and the x-axis indicates the centered teacher total usage time. The plot shows that across all student groups, the predicted hazard was less for students with teachers with higher total usage time. However, the relationship between teacher usage time and predicted hazard was stronger for the light and medium user student groups than for the heavy user student group. Specifically, the estimated slopes for the light (β_{low}: -0.0031 , $p<0.00$) and medium user groups (β_{middle}: -0.0038 , $p<0.00$) were similar and steeper than that for the heavy user group (β_{high}: -0.0014 , $p<0.00$). Noticeably, the gap in the predicted hazard between the heavy user group and the other two groups narrowed as teacher total usage time increased.

Discussion

Before discussing the results, we note two things. First, it was impossible to measure and control for all teacher practice and student experience variables that might relate to student duration in Algebra Nation. For example, some teachers may have forced their students to use the platform during class time, and some students' patterns of engagement in Algebra Nation may reflect unmeasured variables such as student access to technology and motivation toward academics. Therefore, we are careful to not draw causal inferences about the relationships between the predictor variables and duration in Algebra Nation. Second, although some demographic variables were statistically significantly associated with student usage duration, they served merely as control variables while answering our main research question, and thus are not further interpreted in this discussion section. Furthermore, we did not have sufficient information to capture meaningful dynamics introduced by different demographic groups, teachers, schools, and districts, and thus cannot meaningfully interpret these variables in relation to the model results. For example, students who identified themselves as African Americans used Algebra Nation longer than did students in other ethnic groups, but this could result from various factors at the student, teacher, school, and district levels, rather than theoretically explainable factors. Readers may find these results interesting in their own readings, but our discussion focuses on findings from usage patterns based on theoretical and empirical evidence provided in previous research, with all demographic factors serving only as controls.

The number of video lectures the students watched was positively associated with their usage duration of Algebra Nation, a finding congruent with those of prior studies in which variables related to video viewing were treated as a proxy of student engagement in online courses. For example, Heo and Choi (2014) implemented the flipped learning model in a seventh-grade math class using OERs such as lecture videos, quizzes, and exercise sheets to identify factors associated with student achievement. Of three learning-related variables including the number of lecture videos students watched, total learning time, and task completion rates, number of lectures was correlated with students' test scores. Kim, Yoon, Jo, & Branch (2018) claimed that watching video lectures was related to self-regulated learning skills in self-paced asynchronous online learning environments because it directly indicated the extent to which students engaged with course content. In recent forms of OER-supported online courses (e.g., Khan Academy), video lectures are a primary source for content learning in flipped learning in secondary education and typically followed by course activities building on content delivered in lectures (Pursel et al., 2016). Therefore, only students who successfully complete lecture videos are expected to proceed with subsequent online activities.

Our finding for student video seeking corroborates the importance of students' interaction with lecture videos. In this study, students who more frequently used the seek button while watching the lecture videos were more likely to remain in Algebra Nation when other factors were controlled. Video seeking behavior is considered related to strategic learning. Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, Morales, and Munoz-Gama (2018) noted that online students tend to superficially watch videos without reflecting on where they should focus. Online students do not necessarily learn from video lectures when they merely pass time while the videos are playing (Waldrop, 2013). Given the critical role of strategic focus in knowledge acquisition (Maldonado-Mahauad et al., 2018), students who used the seek button on the video player more frequently strategically focused their attention on content important to them than did those who just watched the videos without strategic thinking. In addition to the number of watched videos, information that displays how students watch them may indicate to teachers which students are more engaged with course material.

The number of solved test items and reviewed videos relevant to these solved test items predicted students' usage duration when other variables were controlled for in our model. This finding is consistent with empirical evidence of prior studies on the relationship between student cognitive engagement and achievement. Maldonado-Mahauad et al. (2018) indicated that students who actively participate in online quizzes after watching lecture videos were likely to be aware of what they should study to improve their performance and achievement. In contrast, students with low cognitive engagement in online learning contexts tend to skip quizzes. Ziegelmeier and Topaz (2015) examined how high school students used quizzes to monitor their own understanding of class topics during flipped learning. Interviews conducted with students who participated in flipped learning revealed that they recognized quizzes as an important tool to "keep up-to-date with course material" (p. 854). Similarly, González-Gómez, Jeong, and Rodríguez (2016) showed that high school students who participated in the flipped classroom using an OER-supported online learning platform perceived online quizzes as helping them achieve learning objectives. Furthermore, students who carefully reviewed the quizzes before class concentrated on the challenging part of the problems addressed during the online video lectures without uncertainty about class topics.

For students, reviewing lecture videos that covered the test items they solved was associated with student usage duration beyond what can be explained in duration by the other variables in the model. Reviewing lecture videos is related to student cognitive engagement, because it indicates that students invested extra time in refining their understanding of the course topics addressed in the test items (Sun, 2014). Interestingly, watching solution videos particular to each solved test item did not predict student usage duration in our study. While checking solutions to solved test items can still be considered cognitive engagement, it requires less time and effort than reviewing entire lecture videos, which was a significant predictor in the study. We argue that this is because the length of each solution video in Algebra Nation ranges from 1–5 minutes, while the lecture videos are approximately 6–40 minutes long. In self-paced OER-supported learning environments, student quiz engagement can be a strong indicator of student engagement and achievement because it stems from students' strong motivation to learn without teachers forcing them to do so. We suggest paying attention to whether students invest extra time and effort in activities following main content learning in OER-supported online learning contexts. For example, instructors can regularly check whether their students use extra learning materials (e.g., watching relevant videos) to determine whether they are cognitively engaged.

Our finding may indicate the role of social tools in promoting students' use of OER-supported online learning platforms, which can be experimentally evaluated in future research. This study revealed that students' use of the Algebra Wall, where they can ask questions or see others' questions and responses, predicted their usage duration. Being isolated from teachers and peer learners, students who use OER-supported online learning platforms often find it challenging to establish communication with their peers and experts. Social tools can make students feel connected to other people, and this psychological easiness increases their emotional engagement with online communities. According to the community of inquiry framework, the social presence that students feel can be enhanced by their participation in social interactions such as online forums or annotation tools (So, 2009). Even though OER-supported online learning platforms are used as supplemental resources in formal school curricula, when initiated by teachers, students may feel isolated as they have to wait until the following class to receive help from their teachers. Social features such as discussion forums are often integrated in OER-supported online learning platforms such as Khan Academy, and have been acknowledged as supporting intellectual interactions between students and teachers outside the classroom (Light & Pierson, 2014).

One limitation of examining frequency of views is that is does not indicate that students fully engaged in seeking help.. Nevertheless, even observing social communication among other students (e.g., reading other students' questions and comments) can lead to knowledge development (Dennen, 2008). As such, so-called lurkers are known to cultivate their cognitive engagement by merely viewing others' posts and discussions (Dennen, 2008). This marginal participation still helps them develop their knowledge in the course. Therefore, it may be important to promote student engagement in social activities when students are isolated in online learning contexts.

At the teacher level, teachers' use of the Teacher Area was predictive of student usage duration. In Algebra Nation, the Teacher Area allows teachers to monitor their students' progress by checking their number of video watches and questions answered correctly and incorrectly, download instructional resources, and communicate with other teachers by posting in the teacher wall. We could not obtain fine-grained data on how frequently the teachers used each of the features in the area. However, we infer that engaging more in some combination of monitoring their students, utilizing extra teaching resources, and communicating with other teachers indicates that these teachers had higher use the OER during class time, which in turn indicates that students spent more time in the platform under the supervision of the teacher. Our finding coincides with that of Faber, Luyten, and Visscher (2017), who found the positive impact of teacher dashboards that support teachers' monitoring of their students on student motivation and achievement in mathematics classes. Compared to classroom contexts, online learning makes it difficult for teachers to intervene in their students' learning. Teacher support tools such as a learning analytics dashboard supported by student usage data and on-going performance have been found to mitigate the issue of control (Martinez-Maldonado, Clayphan, Yacef, & Kay, 2014). Furthermore, teacher social tools can engage teachers in the online community of practice (Tseng & Kuo, 2014). Many teachers find it difficult to fully benefit from using OERs (Misra, 2014). Teacher involvement in teacher communities for sharing ideas and help-seeking can assist them in using OERs in a pedagogically appropriate way and resolves potential issues arising from integrating OERs into classroom teaching. Based on our finding regarding the use of the Teacher Area, we suggest that future research experimentally evaluate the effects of teacher spaces in OER-supported online learning platforms on student attrition in such platforms..

We conclude by discussing the interaction of student and teacher usage time, which provides new insight into integrating OERs in formal education. Controlling for all other covariates, an increase in teacher usage time was associated with a decrease in predicted hazard, as expected, meaning that students whose teachers spent more time in Algebra Nation were also likely to spend more time in the platform. However, this relationship was less prevalent for students who were heavy users (relatively defined) than for light or moderate users. Here, the gap between these three student groups in their risk of attrition decreased when they were associated with teachers who spent more time in Algebra Nation. These findings can be explained within the self-regulated learning framework. Successfully self-regulated students may be less affected by teacher usage and likely to persist in completing online activities, whether or not they are required to do so. In contrast, students who minimally use OERs are likely to discontinue their engagement if their teacher is not invested in the long term. This highlights the need for future research to evaluate the efficacy of using differentiated facilitation strategies for students with different levels of engagement with OERs. In fact, the Algebra Nation development team informed us that some participating teachers require their students to use the platform for their homework or class activities. This interaction may indicate that medium and light users who engaged with Algebra Nation as required by their teachers were more likely to stop using it when it was no longer required. Our findings suggest that careful attention be paid to lighter users. Wong, Khalil, Baars, de Koning, & Paas (2019) emphasized the importance of acquainting students with self-regulated learning with OER so that they perceive the resources as useful sources for their own learning beyond class time.

Limitations

There are several limitations in this study, two of which are particularly important. First, we were not able to see what teachers or students printed from the online materials such as test items or worksheets. These offline activities were not recorded in the database and may have influenced our analysis. As this may be true in studies using computer logged data from OER-supported online learning environments, future studies should be smaller scale, which would enable exploring student and teacher experiences through various data collection methods such as observations and interviews (Huggins-Manley et al., 2019). Second, the granularity of log data used in this study was not sufficient for detailed inferences regarding teacher behavior because we could not track what features the teachers used in the Teacher Area. However, we did use a coarser variable to examine the relationship between Teacher Area usage and student duration in the platform, leading to meaningful discussion points in the preceding section. Third, the data did not have separate variables for student and teacher usage of the OER during class time and outside of class time, which would help clarify the interaction of student and teacher usage time observed in this study. Fourth, although we controlled potential classroom and school-level factors by including numerous covariates, it was still not possible to capture dynamics in practice. For example, some teachers may have allocated limited time to learning with the platform. Likewise, school principals may have encouraged teachers to use the platform during an exam. Future research should consider using additional methods such as surveys and interviews to better understand how the use of the platform varies in local contexts. Last, future research should test whether the online learning and teaching patterns represented by the identified proxy variables actually result in better learning in OER-supported online learning platforms. Controlled experiments would provide better insights into causal relationships between these learning patterns and student learning outcomes, as well as how to facilitate student learning in similar learning environments.

Conclusion

This study's contribution to research and practice is summarized below. First, we addressed the lack of research on integrating OERs into formal secondary classrooms despite the rapid growth of OERs in K–12 settings. Prior studies were mostly conducted in higher education settings or on the fully online course mode. Therefore, our study contributes to increasing the knowledge base about OER use in K-12 setting. Second, we provided insights into key proxies that may be indicators of student engagement with OER-supported open online learning platforms. Although previous research provided some evidence of the relationship between student learning patterns and their achievement in online learning platforms at scale, little has been found on how to support students with decreasing engagement over time. This longitudinal study identified a few student- and teacher-level proxy variables that can be used to detect at-risk students in the early phases of learning on OER-based platforms, and develop student and teacher support systems such as a monitoring tool and early warning feature. Last, we elucidated student learning patterns in relation to their teachers' usage patterns through a multilevel modeling approach. Given that OERs are increasingly employed in formal classrooms led by a teacher, student learning with OERs is largely influenced by teachers' strategies. Not only did we control possible variables resulting from teacher-level variables, but student background information was also considered for the discovery of useful knowledge. Our findings inform the design and implementation of OER-based online learning in classrooms that are dynamically dependent on contextual and environmental factors.

References

Algina, J., & Swaminathan, H. (2011). Centering in Two-Level Nested Designs. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of Advanced Multilevel Analysis* (pp.285-312). New York: Taylor & Francis.

Allison, P. D. (2009). *Fixed effects regression models*. Los Angeles, CA: Sage.

Asmussen, S., Binswanger, K., & Højgaard, B. (2000). Rare events simulation for heavy-tailed distributions. *Bernoulli*, *6*(2), 303-322.

Austin, P. C. (2017). A tutorial on multilevel survival analysis: methods, models and applications. *International Statistical Review*, *85*(2), 185-203.

Borup, J., & Stevens, M. (2016). Parents' perceptions of teacher support at a cyber charter high school. *Journal of Online Learning Research*, *2*(3), 227-246.

Brahimi, T., & Sarirete, A. (2015). Learning outside the classroom through MOOCs. *Computers in Human Behavior*, *51*, 604-609.

Butcher, N., & Wilson-Strydom, M. (2008). Technology and open learning: The potential of open education resources for K-12 education. In Voogt, J., & Knezek, G. (Eds.), *International handbook of information technology in primary and secondary education* (pp. 725-745). New York, NY: Springer.

Burd, E. L., Smith, S. P., & Reisman, S. (2015). Exploring business models for MOOCs in higher education. *Innovative Higher Education*, *40*(1), 37-49.

Cargile, L. A., & Harkness, S. S. (2015). Flip or Flop: Are Math Teachers Using Khan Academy as Envisionedby Sal Khan?. *TechTrends*, *59*(6), 21-28.

Crowther, M. J. (2017). Multilevel mixed effects parametric survival analysis. *International Journal of Oral & Maxilofacial Implants, 32*(5), 1097-1102.

Dennen, V. P. (2008). Pedagogical lurking: Student engagement in non-posting discussion behavior. *Computers in Human Behavior*, *24*(4), 1624-1633.

Engle, D., Mankoff, C., & Carbrey, J. (2015). Coursera's introductory human physiology course: Factors that characterize successful completion of a MOOC. *The International Review of Research in Open and Distributed Learning*, *16*(2). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/2010

Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers & education*, *106*, 83-96.

Friesen, N., & Wihak, C. (2013). From OER to PLAR: Credentialing for open education. *Open Praxis*, *5*(1), 49-58.

González-Gómez, D., Jeong, J. S., & Rodríguez, D. A. (2016). Performance and perception in the flipped learning model: an initial approach to evaluate the effectiveness of a new teaching methodology in a general science classroom. *Journal of Science Education and Technology*, *25*(3), 450-459.

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. (1995). *Multivariate Data Analysis* (3rd ed). New York: Macmillan.

Heo, H. J., & Choi, M. R. (2014). Flipped learning in the middle school math class. *Advanced Science and Technology Letters*, *71*, 94-97.

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational research review*, *12*, 45-58.

Hox (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

Hu, E., Li, Y., Li, J., & Huang, W.-H. (2015). Open educational resources (OER) usage and barriers: A study from Zhejiang University. China. *Educational Technology Research and Development, 63*(6), 957–974.

Huggins-Manley, A. C., Beal, C. R., D'Mello, S. K., Leite, W. L., Cetin-Berber, D. D., Kim, D., & McNamara, D. S. (2019). A commentary on construct validity when using operational virtual learning environment data in effectiveness studies. *Journal of Research on Educational Effectiveness*, *12*(4), 750-759.

Light, D., & Pierson, E. (2014). Increasing student engagement in math: The use of Khan Academy in Chilean classrooms. *International Journal of Education and Development using ICT*, *10*(2), 103-119.
Mishra, S. (2017). Open educational resources: removing barriers from within. *Distance Education*, *38*(3), 369-380.
Moreno-Marcos, P. M., Muñoz-Merino, P. J., Maldonado-Mahauad, J., Pérez-Sanagustín, M., Alario-Hoyos, C., & Kloos, C. D. (2019). Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs. *Computers & Education*, 103728. https://doi.org/10.1016/j.compedu.2019.103728.
Khalil, H., & Ebner, M. (2014, June). MOOCs completion rates and possible methods to improve retention-A literature review. In *EdMedia+ Innovate Learning* (pp. 1305-1313). Association for the Advancement of Computing in Education (AACE).
Kim, D. (2018). A Framework for Implementing OER-Based Lesson Design Activities for Pre-Service Teachers. *The International Review of Research in Open and Distributed Learning, 19*(4). https://doi.org/10.19173/irrodl.v19i4.3394.
Kim, D., Lee, I. H., & Park, J. H. (2019). Latent class analysis of non-formal learners' self-directed learning patterns in open educational resource repositories. *British Journal of Educational Technology, 50*(6), 3420–3436.
Kim, D., Yoon, M., Jo, I. H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. *Computers & Education, 127*, 233–251.
Kovanovic, V., GaĹˇevic, D., Dawson, S., Joksimovic, S., Baker, R. S., & Hatala, M. (2015). Does time-on-task estimation matter? Implications for the validity of learning analytics findings. *Journal of Learning Analytics*, *2*(3), 81-116.
Leite, W. L., Cetin-Berber, D. D., Huggins-Manley, A. C., Collier, Z. K., & Beal, C. R. (2019). The relationship between Algebra Nation usage and high-stakes test performance for struggling students. *Journal of Computer Assisted Learning*. *35*, 569-581.
Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, *80*, 179-196.
Marple, S., Jaquet, K., Laudone, A., Sewell, J., & Liepmann, K. (2019). *Khan Academy in 7th Grade Math Classes: A Case Study*. San Francisco, CA: WestEd.
Martinez-Maldonado, R., Clayphan, A., Yacef, K., & Kay, J. (2014). MTFeedback: Providing notifications to enhance teacher awareness of small group work in the classroom. *IEEE Transactions on Learning Technologies*, *8*(2), 187-200.
McShane, M. Q. (2017). Open educational resources: is the federal government overstepping its role?. *Education Next*, *17*(1), 18-25.
Misra, P. (2014). Online training of teachers using OER: Promises and potential strategies. *Open Praxis*, *6*(4), 375-385.
Muir, T. (2014). Google, Mathletics and Khan Academy: students' self-initiated use of online mathematical resources. *Mathematics Education Research Journal*, *26*(4), 833-852.
Ngimwa, P., & Wilson, T. (2012). An empirical investigation of the emergent issues around OER adoption in Sub-Saharan Africa. *Learning, Media and Technology, 37*(4), 398–413.

Pérez-Paredes, P., Ordoñana Guillamón, C., & Aguado Jiménez, P. (2018). Language teachers' perceptions on the use of OER language processing technologies in MALL. *Computer Assisted Language Learning*, *31*(5-6), 522-545.

Pursel, B. K., Zhang, L., Jablokow, K. W., Choi, G. W., & Velegol, D. (2016). Understanding MOOC students: motivations and behaviours indicative of MOOC completion. *Journal of Computer Assisted Learning*, *32*(3), 202-217.

Robinson, T. J., Fischer, L., Wiley, D., & Hilton III, J. (2014). The impact of open textbooks on secondary science learning outcomes. *Educational Researcher*, *43*(7), 341-351.

Schaeffer, C. E., & Konetes, G. D. (2010). Impact of learner engagement on attrition rates and student success in online learning. *International Journal of Instructional Technology & Distance Learning*, *7*(5), 3-9.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: Sage Publications.

So, H. J. (2009). When groups decide to use asynchronous online discussions: collaborative learning and social presence under a voluntary participation structure. *Journal of Computer Assisted Learning*, *25*(2), 143-160.

Sun, J. C. Y. (2014). Influence of polling technologies on student engagement: An analysis of student motivation, academic performance, and brainwave data. *Computers & Education*, *72*, 80-89.

Tseng, F. C., & Kuo, F. Y. (2014). A study of social participation and knowledge sharing in the teachers' online professional community of practice. *Computers & Education*, *72*, 37-47.

Waldrop, M. M. (2013). The virtual lab. *Nature, 499*, 268–270. doi:10.1038/499268a

Weller, M., De Los Arcos, B., Farrow, R., Pitt, B., & McAndrew, P. (2015). The impact of OER on teaching and learning practice. *Open Praxis*, *7*(4), 351-361.

Wiley, D., Bliss, T. J., & McEwen, M. (2014). Open Educational Resources: a review of the literature. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 781–789). Springer.

Wiley, D., Hilton III, J. L., Ellington, S., & Hall, T. (2012). A preliminary examination of the cost savings and learning impacts of using open textbooks in middle and high school science classes. *The International Review of Research in Open and Distributed Learning*, *13*(3), 262-276.

Wiley, D., Webb, A., Weston, S., & Tonks, D. (2017). A preliminary exploration of the relationships between student-created OER, sustainability, and students' success. *The International Review of Research in Open and Distributed Learning*, *18*(4). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/3022

Wong, J., Khalil, M., Baars, M., de Koning, B. B., & Paas, F. (2019). Exploring Sequences of Learner Activities in Relation to Self-regulated Learning in a Massive Open Online Course. Computers and Education, 140, 103595. https://doi.org/10.1016/j.compedu.2019.103595

Wong, T. L., Xie, H., Zou, D., Wang, F. L., Tang, J. K. T., Kong, A., & Kwan, R. (2019). How to facilitate self-regulated learning? A case study on open educational resources. *Journal of Computers in Education*, 1-27.

Zheng, S., Rosson, M. B., Shih, P. C., & Carroll, J. M. (2015). Understanding student motivation, behaviors and perceptions in MOOCs. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing* (pp. 1882-1895). ACM.

Ziegelmeier, L. B., & Topaz, C. M. (2015). Flipped calculus: A study of student performance and perceptions. *Primus*, *25*(9-10), 847-860.