TEACHERS COLLEGE
COLUMBIA UNIVERSITY

Frontier by eSpark Learning: Product Analysis Report

Chuheng Hu, Samantha Pepe, and Jonathan Stelman

Teachers College, Columbia University

Table of Contents

Abstract

To better understand the efficacy of its new product “Frontier,” the educational technology company eSpark Learning partnered with student members of the Learning Analytics program of Teachers College, Columbia University. The collaborative research project links student demographics, assessment performance, and Frontier usage data taken from a Pennsylvania school district in order to identify elements of Frontier’s content, structure, and usage that led to the greatest growth outcomes on the NWEA standardized tests over one school year. The following analysis focuses on Frontier’s impact on reading exam scores, since Frontier is a tool for students that emphasizes reading and writing. This study finds that increasing Frontier usage leads to stronger academic outcomes, particularly among struggling students.

Overview of eSpark Learning

eSpark Learning is an educational technology company founded in Chicago, Illinois. The company was founded in 2010 (“Working at eSpark Learning,” n. d.). eSpark Learning builds tools to help PK–8 grade students develop skills in English language arts and mathematics and reach individual student’s learning goals. The company also offers professional development options that are tailored to each school district’s unique education needs. (“eSpark Learning,” n. d.).

Overview of Frontier

The newest product offering by eSpark Learning is called Frontier. Frontier is a web-based service which delivers "inquiry-based web and chromebook lessons for grades 3–8." (“eSpark Learning,” n. d.). In Frontier, students explore a number of different modules, or “Frontiers.” Each Frontier consists of a curated set of multimedia content, or “activities,” which include videos, articles, infographics, podcasts, or other forms of media. All activities in a Frontier are selected to help students answer a specific, essential question. Essential questions include “Is tackle football too dangerous for kids to play?”; “How does social media impact your life?”; and “Should driverless cars be allowed on the road?”. When a student selects a Frontier, he or she is presented with the curated set of activities. As they access the activity content, students are prompted to write what they learn through an on-screen note-taking feature. All of Frontier’s activities are aligned with state standards, so school districts can be confident that usage of Frontier helps students to attain the required benchmarks.

As a student completes each activity in a Frontier, he or she should gain a broader, more holistic perspective on the topic, which will allow the student to write a thoughtful response to the Frontier’s essential question. For each Frontier, students take notes to document information learned throughout the activities and reflect on their learning.

As a learning tool, Frontier offers scaffolds to support students as they receive, process, and integrate, and synthesize new information. Frontier is designed to increase engagement by offering content that is the appropriate difficulty for students and allowing students to study topics that are interesting and relevant to them.

Overview of school district partner

The data used in this study is taken from Propel Schools, a public charter school system in Pittsburgh dedicated to transforming education in underserved communities through the use of student-centered instruction. The decision to analyze this school system in isolation is due to the relatively homogeneous demographic of students. Homogeneity holds great importance when analyzing student performance, since there are many omitted variables that arise. Some omitted variables this analysis accounts for include geographic location and proportion of economically disadvantaged students. As indicated in the chart below, Propel students have low test scores relative to the state average, and the district has a high percentage of economically disadvantaged students. Propel Schools has complete participation in providing data to eSpark Learning, and the student assessment data available for this product analysis is as complete as the data provided to the Propel School District by NWEA testing service.

Propel School	1	2	3	4	5	6	7	8
Performance - Math	21.7	23.1	32.0	17.0	42.9	22.1	37.0	17.0
Performance - ELA	41.5	47.9	51.4	34.5	67.7	40.3	56.4	38.8
Performance - Science	43.1	38.7	58.9	35.0	65.4	23.7	60.5	43.9
Percent Free or Reduced Lunch	76.0%	58%	22%	87%	81%	80%	57%	61%

[Sources: National School Lunch Program Reports (n. d.), Pa.gov (n. d.)]

[Performance indicators range 0 to 100, 50 being roughly average for PA]

Goals of the research

This research was undertaken to investigate the academic efficacy of Frontier. These analyses are aimed at discovering insights about Frontier that will drive better implementation fidelity, inform more effective content creation, and optimize learning outcomes for students who use Frontier.

To understand implementation fidelity, these studies examine usage trends between classrooms, grade levels, and schools. Comparing usage across these different cross-sections of users may reveal differences in instructional practices at each of these levels. If trends are identified that reveal certain instructional practices are associated with better learning outcomes, eSpark Learning can offer professional development that encourages these best practices.

To inform more effective content creation, these studies investigate relationships between features of Frontier activities and the learning outcomes of students who use those Frontiers. Individual Frontiers can be described according to their subject areas, the media types of their composition activities, the percentage of students who are initially drawn to start the Frontier, completion rates of students who use the Frontier, and the quality of notes written about the Frontier. Exploring all of these features serves to identify those qualities of Frontiers that lead to the greatest student growth outcomes. In identifying these qualities, eSpark Learning can design more content that incorporates those features that are associated with the most growth.

Research Questions

A number of hypotheses were investigated throughout the research period. The following questions are among those with the most salient conclusions. We address the investigation of each of these hypotheses in detail below.

I. Does increasing weekly time on Frontier lead to academic growth?

Since Frontier is primarily used to improve students' reading and writing skills, the following analysis investigates if increased usage of Frontier is related to higher growth on standardized reading assessments. The NWEA MAP assessment is a computerized adaptive test that is administered three times annually. The exam tests childrens' knowledge in mathematics, reading, and science. For this analysis, students performing below the 50th percentile on their Fall reading exam were included. Based on students' fall MAP performance, their spring performance is then predicted and documented (PA.Gov, n. d.). Table 1 shows the results of a logistic regression predicting whether projected score growth was met based on the student's weekly usage. The significance at the .05 alpha level suggests that underperforming students who consistently use Frontier are significantly more likely to meet their expected growth target than underperforming students who do not consistently use Frontier. For example, based on the logistic regression, underperforming students who utilize Frontier for 45 minutes per week are 21% more likely to meet expected growth than underperforming students who utilize Frontier for 5 minutes a week.

Logistic Regression- Average Weekly Time Spent on Frontier vs Meeting Expected Growth (yes/no)				
	Estimate	**Standard Error**	**Z Value**	**P Value**
Constant	-0.572	0.237	-2.42	0.0156
Average Time Weekly	0.021	0.009	2.2	0.0204**

[** Significant at .05 alpha level]

Table 1

Using a decision tree model, we look to determine which Frontier usage variables, if any, are significant in predicting if a student meets projected growth on the spring NWEA MAP assessment. The sources of data input are student usage data and fall to spring assessment data. Both data were subset by students performing lower than the 50th percentile on their fall assessment. Student usage data includes such variables as number of Frontiers completed, number of activities completed, average length of note, and other variables indicating user usage. The spring assessment data indicates whether each student has met his or her projected growth. We select to measure whether a student meets their projected growth target as the outcome to be predicted. The decision tree is considered to be have significant predictive power if the accuracy of prediction is greater than 50%. Given these usage variables, we produce a decision tree with a rate of accuracy of 65%, a significant result. We can conclude that given a student's level of usage, we can more accurately predict compared to random chance whether a student meets projected growth.

A training set is created by subsetting 66% of the students, leaving the remaining 34% as a test set. To avoid overfitting and reduce the computational load, we assign a complexity parameter equal to 0.025. This prevents the algorithm from selecting any splits that do not increase the R^2 value of the model by at least this amount (Therneau & Atkinson, 2017).

Figure 1 shows the decision tree. The predictor variables used in tree construction are average note length, average time per session, start note length, and total minutes of usage. Because the algorithm selects those variables that best minimize the loss function, it can be inferred that these variables that are represented in the tree are the most salient features for predicting whether students meet projected growth on their spring MAP assessment. The generated decision tree first splits observations on average time per session; students with average minutes per week lower than 3.5 are passed to a terminal node, and classified as not meeting projected growth, while the rest are passed to the next node for further classification. Observations passed to this node are next split according to their total minutes; students with fewer than 312.5 total minutes are passed to the left node, while the rest are passed to the right. We see that observations in both of these nodes are split according to the starting note length, and a final decision is made at fourth level, which classifies observations according to average note length.

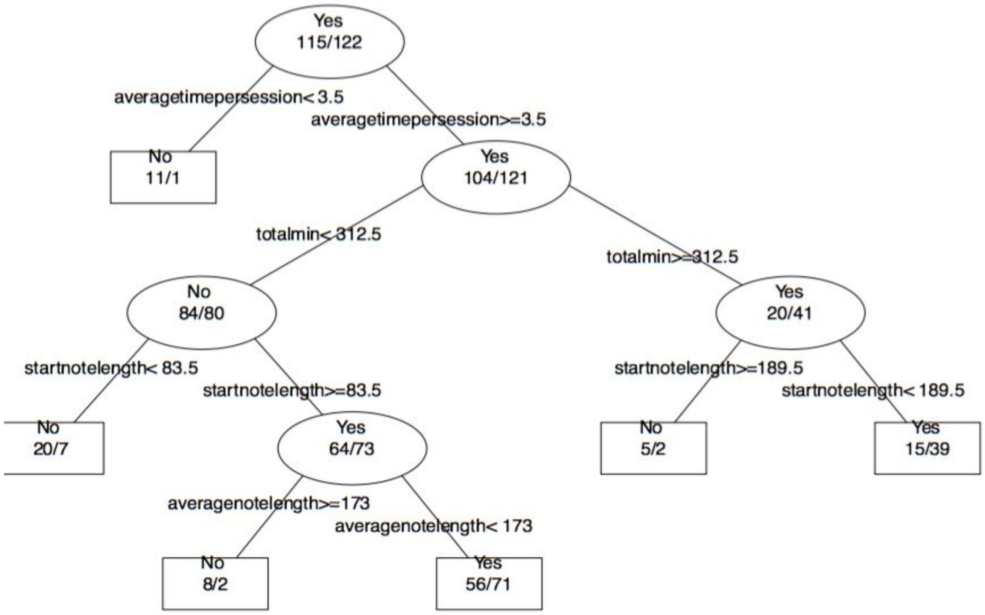

Figure 1

When applying this classification tree on the test set, the accuracy of prediction = 0.647, meaning 64.7% of the time, while applying this decision tree to the test set, the prediction given by the tree matches with the actual result.

II. Do students who meet their projected growth targets have similarities or identifiable trends in usage?

We hypothesize that students who meet projected growth targets and those who do not meet their targets have different topical preferences when selecting Frontiers to complete. We want to test if there is some particular type of Frontiers that are utilized more often by those who meet projected growth targets. The key takeaway from the analysis is that for each Frontier topic category, students who met their projected growth target used more Frontiers than those students who did not meet their projected growth targets. This also suggests that increased usage of Frontier, regardless of Frontier topic, is correlated with meeting growth targets.

The sources of data input are student usage, fall to spring assessment data and list of Frontier's activities categorized by topic. The student usage data file contains variables such as number of Frontiers, number of activities, average length of note and other indicating variables of user usage. The fall and spring assessment data are filtered for only reading assessment scores.

		Arts	**History**	**Health**	**Other**	**Political science**	**Pop culture**	**Science**	**Sports**	**Total Number**
Met Projected Growth	**Total number**	2326	72	153	9201	321	3083	5375	2690	377
	Average Number of Frontiers / Student	6.170	0.191	0.406	24.406	0.851	8.178	14.257	7.135	
Did Not Meet Projected Growth	**Total number**	2321	57	126	8556	270	2527	4300	2745	421
	Average Number of Frontiers / Student	5.513	0.135	0.299	20.323	0.6413	6.002	10.214	6.520	
Significance Level (between group means)		0.371	0.585	0.521	**0.014**	0.39	**0.047**	**0.014**	0.542	

Table 2

Table 2 displays the total number and average number of Frontiers students completed under each topic for both groups: student who met projected growth and student who did not meet projected growth. We use one-way ANOVA to analyze the difference between means of the two groups. The p value for the difference between group means is also shown in Table 2.

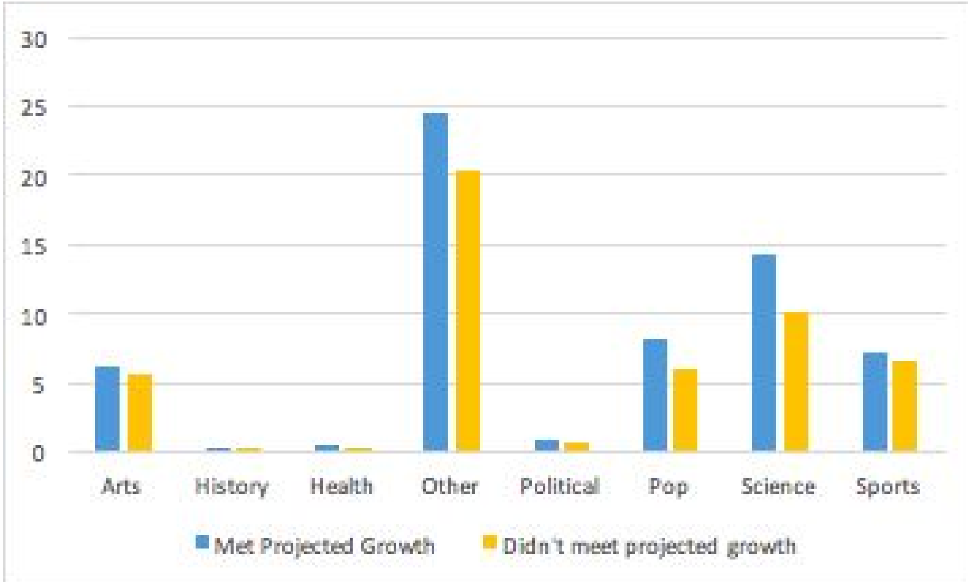

Figure 2

Figure 2 shows the differences between groups for each topic. According to Table 1, the difference between Other, Pop Culture and Science are significant (p-value<0.05) As a result, we conclude that completion of Frontiers focused on Science and Pop Culture significantly associate with meeting projected growth targets.

III. What Frontier topic categories are most interesting to lower performing students?

The interests of individual students may be related to his or her ability. With this in mind, Frontier gives students the ability to select activities from a range of topics that may be relevant to their interests. Understanding which topic categories are most interesting to different groups of students can help curriculum designers ensure that there is ample content that is interesting to different student groups. We hypothesize that high achieving students may be more interested in more "academic" topics, such as science and literature. To test this hypothesis, students are grouped by performance level, and then we count the number of students in each group who have engaged with Frontiers in each topic. This research reveals that Frontiers about science topics are attractive to both high- and low-achieving students.

This analysis suggests that students who were initially below the 50th percentile in the fall and who met projected growth by spring completed significantly more science Frontiers than their counterparts who did not meet projected growth.

The sources of data input are student usage data, fall to spring assessment data, and a list of Frontier's activities categorized by different topics. The first two data sources were subset to include only students performing below the 50th percentile on the fall MAP assessment, and filtering for only reading assessment scores.

	Arts	History	Health	Other	Political Science	Pop Culture	Science	Sports
Group 1: All students performing below 50th percentile on fall NWEA MAP assessment								
# students who touched topic	70	2	12	238	5	74	79	72
avg number of activities	17.7	4	12	20.7	12.6	16.1	34.1	20.6
Group 2: Students who MET projected growth on spring NWEA MAP assessment								
# of students who touched topic	43	2	7	122	4	43	46	39
avg number of activities	18.4	8	12.9	22.7	10.3	18.5	38.2	20.8
Group 3: Students who DID NOT MEET projected growth on spring NWEA MAP assessment								
# of students who touched topic	27	0	5	115	1	31	33	33
avg number of activities	17.3	0	13.5	8.748	22	13.5	29.3	20.9
Significance (p value)	0.073	0.333	0.482	0.249	0.605	0.062	0.064	0.644

Table 3

Table 3 displays the average number of activities students viewed under each topic. The first row examines usage of all students who performed below the 50th percentile on the fall NWEA MAP assessment (Group 1). Of those, we then divide the subset of students who met projected growth on the spring assessment (Group 2), and the subset of students who did not meet projected growth on the spring assessment (Group 3). This table allows us to understand the most frequently viewed topic among the under 50 percentile performers as well as whether these topics later contributes to whether they met projected growth.

The average number of activities for each topic i is given by

$$\overline{X}_a = \frac{n_a}{n_s}$$

where

- n_a is the total number of activities in topic *i* completed by students below the 50th percentile, and
- n_s is the total number of students who touched at least one activity from topic *i* (Therneau & Atkinson).

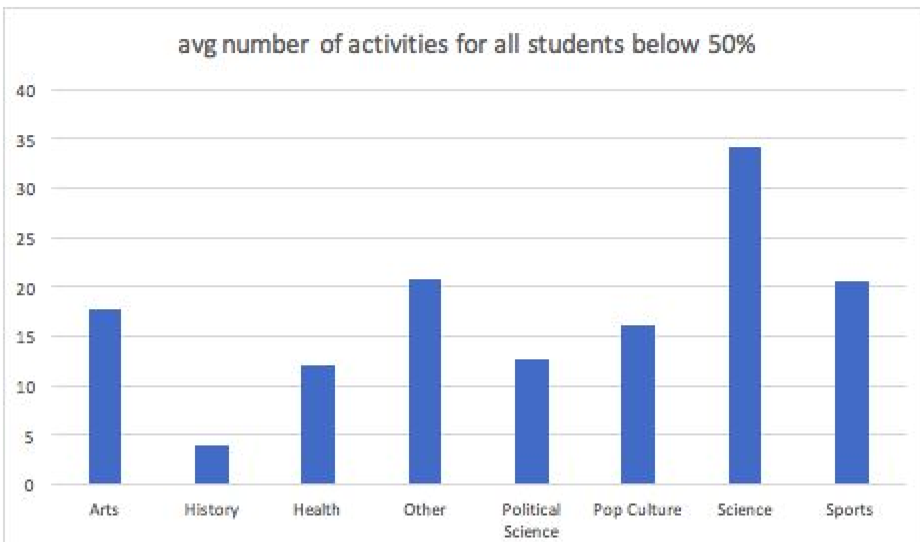

Figure 3

From Figure 3, it is clear that science is the most popular topic among the students below the 50th percentile. Referring back to Table 2, the average number of activities during winter to spring per student is 34.1.

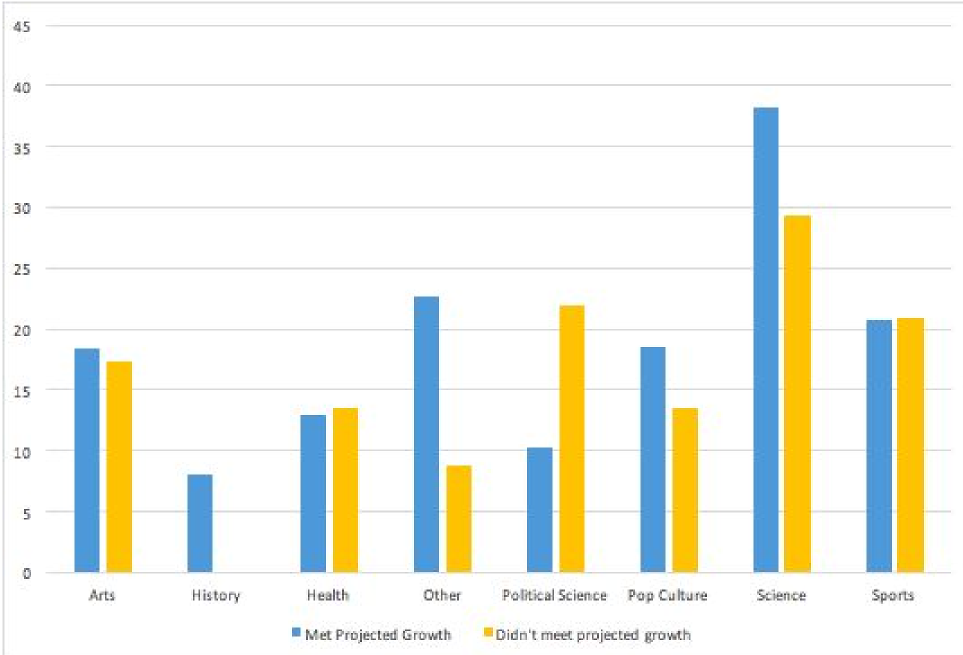

Figure 4

Figure 4 shows the difference between those who later met projected growth and those who didn't meet projected growth. According to Table 3, the difference in performance between these groups is significant with 90% confidence. In other words, students who were initially below the 50th percentile who met projected growth completed significantly more science Frontiers than their counterparts who did not meet projected growth. This suggests a positive relationship between completing science Frontiers and meeting projected growth among underperforming students.

References

eSpark Learning - Differentiated Instruction at Scale. (n. d.). Retrieved from http://www.esparklearning.com/

National School Lunch Program Reports. (n. d.) Retrieved from http://www.education.pa.gov/Teachers%20-%20Administrators/Food-Nutrition/Pages/National-School-Lunch-Program-Reports.aspx#tab-1

PA.Gov. (n. d.). Retrieved from

http://www.education.pa.gov/Teachers%20-%20Administrators/Food-Nutrition/Pages/National-School-Lunch-Program-Reports.aspx#tab-1

Therneau, T. M. and Atkinson, E. J. (2017). An introduction to recursive partitioning using the rpart routines. Mayo Foundation. Retrieved from https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf

Working at eSpark Learning. (n. d.). Retrieved from https://www.glassdoor.com/Overview/Working-at-eSpark-Learning-EI_IE695431.11,26.htm