

Data-informed Curriculum Sequences for a Curriculum-Integrated Game

Ruth Okoilo Akintunde

rookoilo@ncsu.edu

North Carolina State University

Raleigh, NC, USA

Preya Shabrina

pshabri@ncsu.edu

North Carolina State University

Raleigh, NC, USA

Veronica Catete

vmcatete@ncsu.edu

North Carolina State University

Raleigh, NC, USA

Tiffany Barnes

tmbarnes@ncsu.edu

North Carolina State University

Raleigh, NC, USA

Collin Lynch

clynch@ncsu.edu

North Carolina State University

Raleigh, NC, USA

Teomara Rutherford

teomara@udel.edu

University of Delaware

Newark, DE, USA

ABSTRACT

In this paper, we perform a predictive analysis of a curriculum-integrated math game, ST Math, to suggest a partial ordering for the game's curriculum sequence. We analyzed the sequence of ST Math objectives played by elementary school students in 5 U.S. districts and grouped each objective into difficult and easy categories according to how many retries were needed for students to master an objective. We observed that retries on some objectives were high in one district and low in another district where the objectives are played in a different order. Motivated by this observation, we investigated what makes an effective curriculum sequence. To infer a new partially-ordered sequence, we performed an expanded replication study of a novel predictive analysis by a prior study to find predictive relationships between 15 objectives played in different sequences by 3,328 students from 5 districts. Based on the predictive abilities of objectives in these districts, we found 17 suggested objective orderings. After deriving these orderings, we confirmed the validity of the order by evaluating the impact of the suggested sequence on changes in rates of retries and corresponding performance. We observed that when the objectives were played in the suggested sequence, we record a drastic reduction in retries, implying that these objectives are easier for students. This indicates that objectives that come earlier can provide prerequisite knowledge for later objectives. We believe that data-informed sequences, such as the ones we suggest, may improve efficiency of instruction and increase content learning and performance.

CCS CONCEPTS

- **Applied computing → Interactive learning environments;**
- **Computing methodologies → Classification and regression trees.**

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

LAK '20, March 23–27, 2020, Frankfurt, Germany

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-7712-6/20/03...\$15.00

<https://doi.org/10.1145/3375462.3375530>

KEYWORDS

Serious Game Analytics, Curricular Sequencing, Retries, Educational Games

ACM Reference Format:

Ruth Okoilo Akintunde, Preya Shabrina, Veronica Catete, Tiffany Barnes, Collin Lynch, and Teomara Rutherford. 2020. Data-informed Curriculum Sequences for a Curriculum-Integrated Game. In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20), March 23–27, 2020, Frankfurt, Germany*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3375462.3375530>

1 INTRODUCTION

In recent times, there has been an increased exploration of serious games analytics for skill improvement [19]. Whether game data is collected *in situ* (data about game and players collected from within the game environment) or *ex situ* (data about game and players collected from outside the game environment), the analytics of data collected from serious games has helped to improve performance [19] and inform stakeholders on better ways of teaching with serious games [7]. Beneficial patterns gleaned from descriptive [18] and predictive modeling of educational games are being used to advise academic decisions [12, 21].

In this paper, we perform detailed analytics on a mathematics curriculum-based serious game, ST Math. The importance of effective sequencing of objectives in mathematics curriculum cannot be over-emphasized, because of the impact of basic mathematics skills on performance of future more advanced concepts. A study by Dowker et al. shows that many adults have been traumatized by mathematics [9] and one way to alleviate this is by introducing math games and puzzles to children as early as kindergarten [14]. However, a math game with a curriculum sequence that does not adequately teach math skills is counterproductive [2].

1.1 Preliminary Analysis

We conducted an exploratory analysis to discover the different sequences in which five different districts and schools in the United States played ST Math. At first glance, we noticed an important pattern: that the rise and fall in number of retries needed before passing some objectives co-occurred with the position of such objectives in the curriculum sequence. Figure 1 shows how 5 districts played 5 ST Math objectives.

PERCENTAGE RETRIES OF 5 OBJECTIVES PLAYED BY 5 DISTRICTS

DISTRICTS A, B, C, D, E

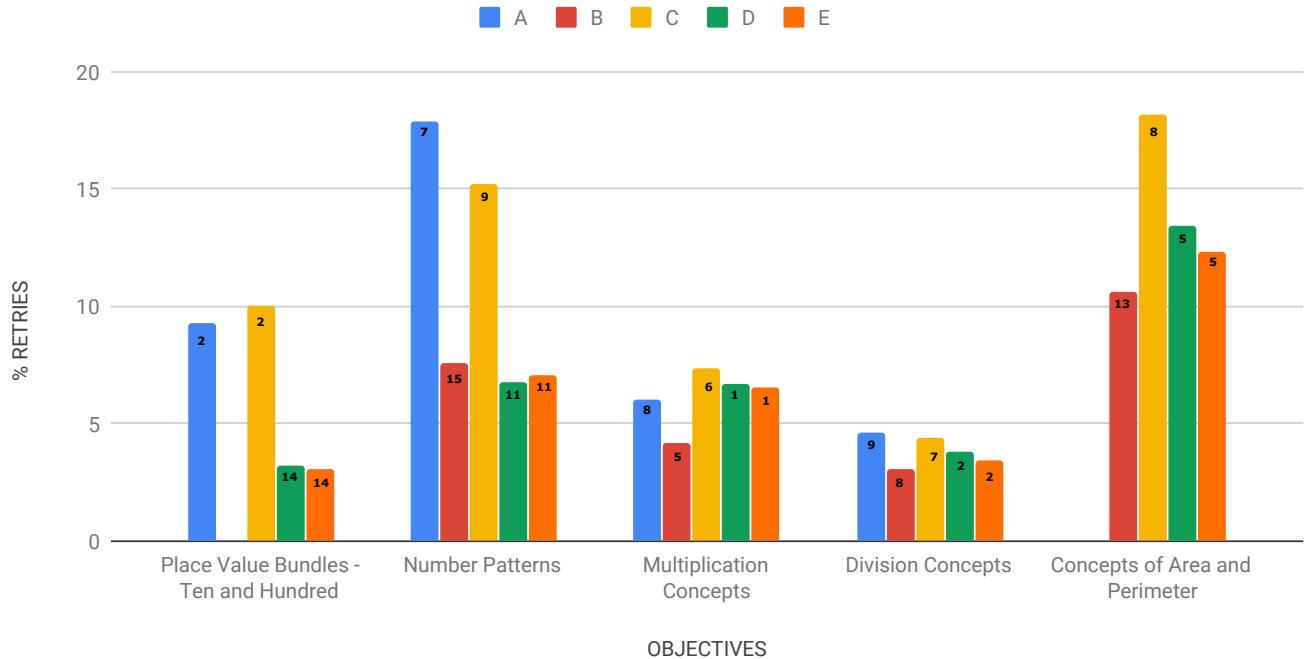


Figure 1: Percentage Retries of 5 Objectives Played by 5 Districts. The number in each bar represents the position of the objective in the curriculum of each district.

The number in each bar represents the position of the objective in the curriculum of each district. There are three groups of objectives in figure 1: *Group 1*: Objectives that are difficult (i.e. have a high number of retries) or easy (i.e. have a low number of retries) depending on their position in the curriculum sequence, *Group 2*: Objectives that are always easy no matter their position in the sequence, and *Group 3*: Objectives that are always difficult, no matter their position in the sequence. Note that district B did not play "Place Value Bundles, Tens and Hundreds" and district A did not play "Concepts of Area and Perimeter." From the plot, we can see that two objectives, namely "Number Patterns" and "Place Value Bundles, Tens and Hundreds" fall into the first group, because their percentage retries dropped when played later in the sequence. When Number Patterns was played as the 7th objective in district A, it accounted for over 17% of all retries for district A, but when it was played as the 15th objective in district B, it was less than 8% of the total retries for district B. "Multiplication Concepts" and "Division Concepts" fall into group 2, whereas "Concepts of Area and Perimeter" falls into group 3. We suggest that group 2 objectives should be played earlier in the sequence, whereas group 3 objectives should be played later, towards the end of the sequence of the first 15 objectives. Our aim is to design a recommended partial ordering for objectives in group 1 so as to reduce the number of retries. To achieve this, we replicated Liu et al.'s method for suggesting a curricular sequence that models the potential predictive relationships

between objectives [21]. The results from our replication study are suggested orderings of objectives in ST Math. We noted these results and designed a partial ordering of objectives as shown in Figure 4. We then evaluated the resulting partial sequence and its association with retries. We performed these procedures on ST Math's *in situ* gameplay data of 3,328 3rd graders during the 2017-2018 school year. These procedures can be applied to other serious games to derive data-informed curricular sequences, potentially improving students' learning and performance.

2 RELATED WORK

A number of studies have performed detailed serious game analytics and applied insights to improving performance of players and the quality of the game environment.

2.1 Insights from Serious Game Analytics

Liu et al. [17] showed that replaying games was not necessarily beneficial, and some patterns may be associated with work-avoidance behaviors. In another study, Liu et al. [22] applied learning curve analyses under different cognitive assumptions to discover game levels where students failed to learn or transfer, identifying potential game design changes that could result in better learning. Horn et al. [13] grouped players into clusters by studying player strategies in order to provide proper feedback at failure points.

From these studies we can see how analytics have been used for improving educational games.

2.2 Methods for Curricular Sequencing

A number of studies have implemented diverse methods to improve curricular sequencing driven by real student data. Chen and colleagues (2016) [3] introduced the Combined student Modeling and prerequisite Discovery (COMMAND) algorithm which uses Bayesian networks to learn prerequisite structure of skills. Their work generates prerequisite graphs of mathematics skills. Yang Chen et al. (2015) [4] discovered pre-requisite structure of skills by using association rule mining on a large student performance dataset. Doroudi et al. (2016) [8] introduced Sequencing Constraint Violation Analysis (SCOVA), which provides a method for evaluating sequencing constraints and how they impact post-test performance. Our approach to curricular sequencing, however, produces a partially ordered sequence which is driven by student's retry behavior. Inferring prerequisite structure of a mathematics curriculum by evaluating its impact on the number of retries is yet to be explored.

2.3 Importance of Replication Study

In our study, we replicated the 2017 study by Liu et al. [21] to discover the objective sequencing that might help students achieve more in terms of their learning outcomes. However, we applied their method on a larger dataset to determine a more generalized sequencing. Replication studies have been used to evaluate research and validate the findings of prior pieces of research[5]. These studies have come up with remarkable findings in varied fields of research, including social sciences, software engineering, computational biology, business and nursing (e.g., [6, 10, 15, 16]).

3 ST MATH DATA

3.1 ST Math

ST Math is a curriculum-integrated educational math game that acts as a year-long supplemental program to the existing math curriculum in elementary classrooms[24]. The game is designed to enhance students' K-12 math skills through spatial-visual representations and informative feedback. ST Math content is divided by grade-level and aligned to relevant Common Core or other state standards. ST Math games, levels, and puzzles are nested within objectives, i.e., broader mathematical concepts. Students apply mathematical knowledge to solve puzzles relevant to each objective's content. As students pass each level within an objective, they are exposed to increasingly difficult content under the umbrella of that same mathematical concept.

ST Math students start the game with a series of introductory games. Students are then routed to the first objective of their assigned content. Before engaging in the puzzles within each objective, students must complete a pre-test. The students need to pass all the levels under an objective to pass that objective. On completion of an objective, the students participate in a post-test as a prerequisite to proceed to the next objective. ST Math objectives assigned to each grade-level can be divided into two categories: default and optional.

At the beginning of a level attempt, the students are given two lives. When the students fail to complete a puzzle accurately, they lose a life. To pass a level, the students need to complete all the puzzles under that level without losing both of their lives.

3.2 Objectives in ST Math

The data used for this study was collected from ST Math gameplay by 3rd graders in 5 US districts during the 2017-2018 school year. The 3rd grade curriculum contains 32 objectives in total; however, for this study we focus on the first 15 objectives, described in Table 2. The order in which ST Math objectives are played is different across districts, classrooms, and students. Teachers have the freedom to assign different objectives to different groups of students in the same classroom at a particular time, although prior research [20] suggests teachers rarely do this, instead most often reordering at a classroom level. We extracted the largest number of students that played the first N objectives in an identical sequence. This statistics is shown in Table 1. From Table 1, we can see that as N increases, the number of students reduces. For this study, we selected the column where N = 15. This represents a trade off between the number of objectives and the size of the dataset i.e., number of observations or students involved in the study. Our dataset therefore contains 826 students from District A, 1,826 students from District B, 376 students from District C, 137 students from District D and 163 students from District E.¹ Thus, we observed the way a total of 3,328 students from the 5 districts played the first 15 objectives. The sequence of objectives was extracted from the gameplay data using the following logic: An objective (e.g., A) is considered ordered prior to another objective (e.g., B) if the majority of its gameplay happened prior to the completion of the other objective. For example, we consider B as ordered before A if a student attempted the first level of objective A, was switched by the teacher to objective B, completed all levels in objective B, and only then completed the rest of the levels in objective A. In this situation, we say that most of the gameplay in objective B happened before objective A. The first 15 objectives played by each district are not exactly the same across districts. For example district A and B might have the same first 13 objectives (played in different order) but the 14th and 15th objectives might be different, hence table 2 has more than 15 objectives. It could be the case that district A played "Multiplication" and "Division" as its 14th and 15th objectives whereas B played "Comparing Fractions" and "Number Patterns" as its 14th and 15th objectives, but B later played "Multiplication" and "Division" as the 16th and 17th objectives. In this case, "Multiplication" and "Division" as the 16th and 17th objectives, would not be a part of this study for district B.

3.2.1 Default and Optional Objectives. ST Math objectives are labeled as "default" or "optional." Those labeled as default are ones in alignment with the Common Core or state standards and have been deemed by the state or district as key content to be covered during the school year. Optional objectives cover material that is supplemental to this curriculum or that covers default objectives in a different or expanded way. Teachers can switch between similar

¹This is a serious limitation to this study; in future work, we plan to perform undersampling or matched samples to account for the differences in sample sizes from different districts

Table 1: Statistics of Students that Played ST Math in the Same Order

Districts	Total # of students	Highest # of students that completed the first N objective in the same order		
		N = 10	N = 15	N = 20
A	7629	1374	826	436
B	3916	2438	1826	1387
C	5158	527	376	214
D	2468	315	137	74
E	2727	344	163	87

default and optional objectives at their discretion. For example: a default objective “Place Value Bundles, Tens and Hundreds” could be replaced by “Place Value, Tens, Hundreds, and Thousands.” We discovered one impact of using optional objectives in place of default objectives. It should be noted that SO1P: “Place Value Bundles, Tens and Hundreds” is difficult when played earlier and easy when played later. However, district B did not attempt this objective; they attempted an *optional objective* called SO1P1000: “Place Value Bundles, Tens, Hundreds, and Thousands” in its place and recorded a low number of retries. It could be that the teachers in district B noticed that SO1P was difficult for students and decided to assign SO1P1000: instead, because it provides a similar concept. We played both SO1P and SO1P1000 to explore this and we observed that knowledge of some prerequisite objectives were required for SO1P, but not required for SO1P1000, because the latter is relatively intuitive.

3.2.2 Retries in ST Math. Retries are defined as all attempts on a level other than the first attempt that occur before the student passes the level. Retries are different from replays in ST Math. Replays on a level are all attempts that occur *after* the level is passed. This study solely focuses on reducing the number of retries for objectives in ST Math. We consider a reduction in retries as indicative of a reduction in the difficulty of objectives. Retries may indicate a lack of content understanding or difficulty with the mechanics of the particular game. Given the design of ST Math to focus on small steps of incremental learning, a large number of retries indicate difficulty beyond what is optimal. A large number of retries prevents students from progressing and covering content. In addition, the need for frequent retries may encourage students to game the system by passing a level using patterns in the game without learning[11, 22]. Our aim is to derive a partially ordered sequence of objectives that places prerequisite objectives earlier in the sequence order. This would provide the students with adequate prerequisite knowledge that in turn reduces difficulty and potentially improves performance and content coverage. We gleaned insights on what objectives should precede the next by replicating a prediction study[21] that suggested objective orderings in ST Math.

4 METHODOLOGY

Our analysis involved three major steps. In the first step, we extracted the total number of retries for each of the first 15 objectives.

This number was aggregated at the district level. We also extracted performance for each student on each objective and the number of levels completed by each student. For performance, we used the first attempt performance for each level, as this provides information on student skills regarding the level content before they experience any feedback. Thus, first attempt performance on a level shows how much prerequisite knowledge the students have acquired from other objectives before playing the current objective and subsequent attempts may not truly reflect each student’s knowledge of the mathematics content [22]. We define objective performance for each student as the average of first attempt performance on all levels in an objective. To get the district performance on an objective, we took the average of all district students’ performance in that objective. We noticed that objectives with higher total number of retries recorded lower average objective performance, whereas those with lower number of retries had higher objective performance. This supports the fact that a high number of retries signifies difficulty of that objective.

In the second step, we replicated the 2017 study by Liu et al. [21] that introduced a novel data-informed method to determine pairwise orderings for objectives in the curricular sequence. This method uses mixed effects regression to identify what objective statistically significantly predicts performance in another objective. If an objective A is played before another objective B in the curriculum sequence but does not statistically significantly predict it, and there exists a situation where objective B is played before objective A and B predicts A, then B should come before A. We will label this scenario with a dark red tile in figure 2. Dark red tiles generally suggest a revision of the order of objectives involved (e.g., the suggestion that objective B should come before objective A is a dark red tile). Our aim is to see which objectives predict performance on the goal objectives (where goal objective is each of the first 15 objectives in the curriculum). We first performed pairwise Spearman correlations on all 15 objectives, resulting in 210 correlations for each district. Next, we found the average pairwise correlations for each district, which we refer to as *mean_corr* and filtered out objective pairs with correlation values less than the *mean_corr*. To ensure there were enough observations, we filtered out objective pairs with fewer than 30 observations. All objective pairs passed this constraint. Next we used mixed effects regression to determine which objectives were statistically significant predictors of the goal objective.

4.1 Linear Mixed Effects Regression

Linear mixed effects models are mixed models used for regression analysis involving both fixed and random effects [1]. These models are an extension of linear regression models such that the linear predictor contains random effects in addition to the usual fixed effects. Linear mixed effects regression allows us to evaluate the significance of fixed effects in created models. The fixed effects in our model represent the relations between the independent variables and our outcome. To account for nesting of students within classrooms, we modeled classroom as a random effect, with a random intercept for each classroom. The *MixedML* function in Python provided us with the p-value of each fixed effect variable’s association with the outcome (goal objective). We set our statistical significance

Table 2: A Description of Objectives in the Curriculum

Objectives	Description and sample games
SO1O: Place Value Concepts	Students write ones, tens, and hundreds in the correct order using digits 0-9. For example, 'How Many Petals' game requires students to count and provide the correct numeral to represent the pile of petals.
SO1P: Place Value Bundles, Tens and Hundreds	Students build on their knowledge of place value concepts developed in earlier objectives to learn the relationship between the ones and tens place and the tens and hundreds place. Students learn that ten or hundred are a bundle of 10 ones or tens. For example: Greenies Regrouping game involves regrouping the ones or tens or both in order to represent the total number in standard expanded form.
ROUND3DGT: Rounding Three-Digit Numbers	Students learn to round two-digit numbers to the nearest ten and three-digit numbers to the nearest hundred. Students round whole numbers using place value skills and the number line. For example: Number Funnels Tens Place game involves rounding decimals to the nearest given value. The game also teaches place value concepts up to the tenths place.
SO1Y: Addition and Subtraction with Regrouping	Students apply strategies for regrouping on numbers up through three digits. For example: Intro to Regrouping game involves using the petals model; students add two three-digit whole numbers with regrouping in the ones or tens place.
SOE: Scale and Measurement in Graphing	Students are required to draw bar graphs from scaled keys and graph quantities with incrementally marked axes. For example: Bar Graph Bridge game involves constructing vertical and horizontal bar graphs for a data set given as single observations or in a set of observations.
A&S_1000: Addition and Subtraction within 1,000	Students apply their knowledge of place value concepts to add two whole numbers. Numbers range from one to three digits. Students also perform addition and subtraction on a number line. For example: Candy Factory game involves identifying the number of tens and ones for a given two-digit whole number.
SO3M: Number Patterns	Students use knowledge of addition and multiplication to count sequences of numbers. For example: Make it Linear game involves identifying the common difference in an increasing or decreasing arithmetic sequence represented in numerical form and with virtual manipulatives in order to extend a sequence of numbers or identify missing numbers in a sequence.
SO1F: Multiplication Concepts	Students gain conceptual understanding of multiplication by repeated addition. Students build an understanding of the differences between addition and multiplication. For example: Bouncing Shoes game involves using repeated addition within the model to determine how many of one animal are needed to fill the given number of shoes.
SO28: Division Concepts	Students gain conceptual understanding of division by manipulating visuals to represent division problems. Students associate division with dividing items into equal groups. For example: Equal Division game involves dividing blocks into equal parts.
MULT&DIV_S: Multiplication and Division Situations	Students are presented with situations including equal groups where they use multiplication and division to solve these problems. For example: Fruit Monster game involves determining how many pieces of fruit are needed to feed the monsters. Students explore the relationship between inputs and outputs using ratios within a visual model.
MULT&DIV_R: Multiplication and Division Relationships	Students disambiguate between multiplication and division using the number line. They are also introduced to the inverse relationship between the two operations. For example: Building Expressions game involves dividing whole numbers by forming equal groups of dots.
MULT: Multiplication	Students practise multiplying and dividing within 100. For example: How Many Legs game involves multiplying whole numbers using repeated addition.
SO29: Division	Students practice solving division problems. Students use visual models to understand and interpret division statements. For example: Area Divide game involves dividing the tiles into equal groups, with and without remainders. The correct answer is demonstrated using an area model.
UNKNOWNs: Unknowns in Two-Step Problems	Students represent two-step word problems using equations and solve them. For example: Mice Island Two-Step Problems game involves filling in the missing number to make the equation true. This game teaches addition and subtraction of one- and two-digit whole numbers.
SO2S: Fraction Concepts	Students represent a whole number as a fraction. For example: Pie Monster game involves representing the given fraction or whole number with circles divided into equal parts.
SOJ: Fractions on the Number Line	Students represent a whole number as a fraction on the number line. For example: Fraction Trap game involves estimating on a number line the location of fractions.
LP&Mode: Line Plots and Mode	Students generate fractional measurement data and show the variability by making a line plot. They discern between multiple population attributes and identify one variable that is expressed in a line plot and identify the mode, minimum, and maximum given in a line plot. For example: Mode is Most game involves identifying the mode of a given collection of shapes or letters.
SO1P1000: Place Value Bundles, Tens, Hundreds, and Thousands	Students apply knowledge of prior objectives such as place value concepts, tens, hundreds, or thousands to learn addition and subtraction techniques based on place value. They also develop the skills to compose and decompose tens, hundreds, and thousands. For example: Building Blocks to 100 games involves filling in the missing addend to make a sum of 100 or 1000.

threshold at $p < 0.05$. To construct a linear mixed-effects model, we extracted all the objectives played before each goal objective. These served as potential predictors (fixed effects) of the goal objective, as shown in equation 1.

Next we applied mixed effects regression to predict P_l (where P_l is a student's performance on the goal objective multiplied by the number of levels of that objective completed by the student) using P_{l_1} to P_{l_k} (which represents the student's objective performance multiplied by number of levels the student completed for objectives 1 to k played before the goal objective), and *position* O of the goal objective in the current curriculum sequence as fixed effects while *Classroom C* factor was used as random effect. Coefficients of predictor objectives that met our p-value threshold were marked as "good predictors," whereas the rest were not considered as predictors. Only predictor-goal objective pairs that had correlations greater than the *mean_corr* were included in the analysis. A total of 58 predictor objectives were found to be statistically significant predictors for the 15 goal objectives.

$$P_l = \beta_0 + \beta_1 O + C \sim N(0, \sigma^2) + \sum_{i=1}^k \beta_i P_{l_i} \quad (1)$$

4.2 Dark Red Tiles

The result of the mixed effects regression analysis includes all objectives, the number of goal objectives they predicted, and the number of objectives that predicted them. Dark red tiles at (X,Y) in figure 2 indicate that Objective X comes before Objective Y in (any) one district's curriculum but when Y is played before X by students in another district, our model shows that Y predicts X. In other words, the ordering YX results in better student performance on X, and suggests that Y should come before X. Dark red tiles were found by checking 5 different districts that played ST Math with different curriculum sequences. If any one district's curriculum (referred to as base curriculum) orders XY, but X does not predict Y, and another district uses YX, and Y predicts X, then that relationship is shown as a Dark Red Tile and Objective Y should be played before Objective X. After forming dark red tiles from 5 districts, we recorded a total of 17 dark red tiles. Using these suggestions, we created a partial order of objectives following the rules of these 17 dark red tiles. In addition to replicating the the Liu study [21] across 5 districts, we also associated the suggested orderings with the number of retries. The third and final step therefore involved deriving a partial ordering of objectives and evaluating the suggested orderings to observe if the number of objective retries was reduced in cases where students already followed our suggested data-informed curriculum sequence

5 RESULTS AND INTERPRETATION

From figure 2, we can see a dark red tile, which suggests that SO1F: "Multiplication Concepts" should be played before SO1O: "Place Value Concepts." In addition, from our analysis in step one, we observed that when SO1F is played before SO1O, the latter is less difficult. Also, from the dark red tile for SO1F: "Multiplication Concepts" and SO3M: "Number Patterns," whenever SO1F is played before SO3M and SO3M is played later in the sequence, the latter is less difficult. If MULT&DIV_S: "Multiplication and Division Situations" is played before SO1P: "Place Value Bundles, Tens and Hundreds," the latter is less difficult. Whenever ROUND3DGT: "Rounding Three-Digit Numbers" is played before SO1P, the latter is less difficult. Therefore, in districts that followed the suggested ordering, we see reduced number of retries compared with districts that did not follow the suggested order.

and Hundreds," the latter is less difficult. Whenever ROUND3DGT: "Rounding Three-Digit Numbers" is played before SO1P, the latter is less difficult. Therefore, in districts that followed the suggested ordering, we see reduced number of retries compared with districts that did not follow the suggested order.

5.1 Partial Ordering

A partial order is a binary relation R over a set S satisfying the following axioms[23]: For all i,j,l in S

- The order S is **transitive**, if $iRj \in S$ and $jRl \in S$, then $iRl \in S$ for all $i \neq j \neq l$
- if $iRj \in S$, then $jRi \notin S$
- $iRi \notin S$

For this study, the relation R means "should be played before." From the dark red tiles suggestions in figure 2, we come up with partial ordering for objectives that fall into group 1. Figure 4 shows the partial orderings.

5.2 Suggested Partial Ordering of Group 1 Objectives

We use the inference from our replication study to suggest a partial ordering for objectives in group 1. After separating the group 2 and 3 objectives, we used the dark red tiles to create a binary partial order between the dark red tiles that involve group 1 objectives and their prerequisite objectives. From figure 4, it is suggested that ROUND3DGT: "Rounding Three-Digit Numbers," MULT&DIV_S: "Multiplication and Division Situations," and SO1P: "Place Value Bundles, Tens and Hundreds" should come before SO3M: "Number Patterns." Possible reasons for this include: ROUND3DGT teaches students the concept of nearest whole numbers, the puzzle environment for MULT&DIV_S is quite similar to that of SO3M, and higher levels of SO1P teaches basic concepts like adding tens and hundreds required for solving higher levels of SO3M. These most likely prepare the students for success in SO3M. Also, prerequisites for MULT: "Multiplication" include SO1Y: "Addition and Subtraction with Regrouping" and MULT&DIV_R: "Multiplication and Division Relationships." MULT moves a step higher from MULT&DIV_R. MULT&DIV_R involves multiplication within tens. MULT involves multiplying 2 digit numerals, i.e., within hundreds. MULT also involves repeated addition; such skills are learnt from SO1Y. The suggested partial orderings were found to reduce the rate of retries of objectives in the partial order shown in figure 4.

5.3 Group 2: Easy Objectives

From our analysis in step one, we noted there were objectives that fell into group 2, i.e., these objectives had a low number of retries no matter their order in the sequence. Objectives SO1O, SO1F, A&S_1000, SO28, SO29, SOE, and SO2S fell into this category. This likely means that the games for these objectives were quite intuitive, their game design was user friendly, or students were able to easily demonstrate their knowledge of the concept. With a mind toward avoiding conflicts with other ordering decisions (e.g., aligning ST Math with classroom curriculum), these objectives should be played at the beginning of the sequence of 15 objectives. Because they are easier objectives, playing them first could help students become

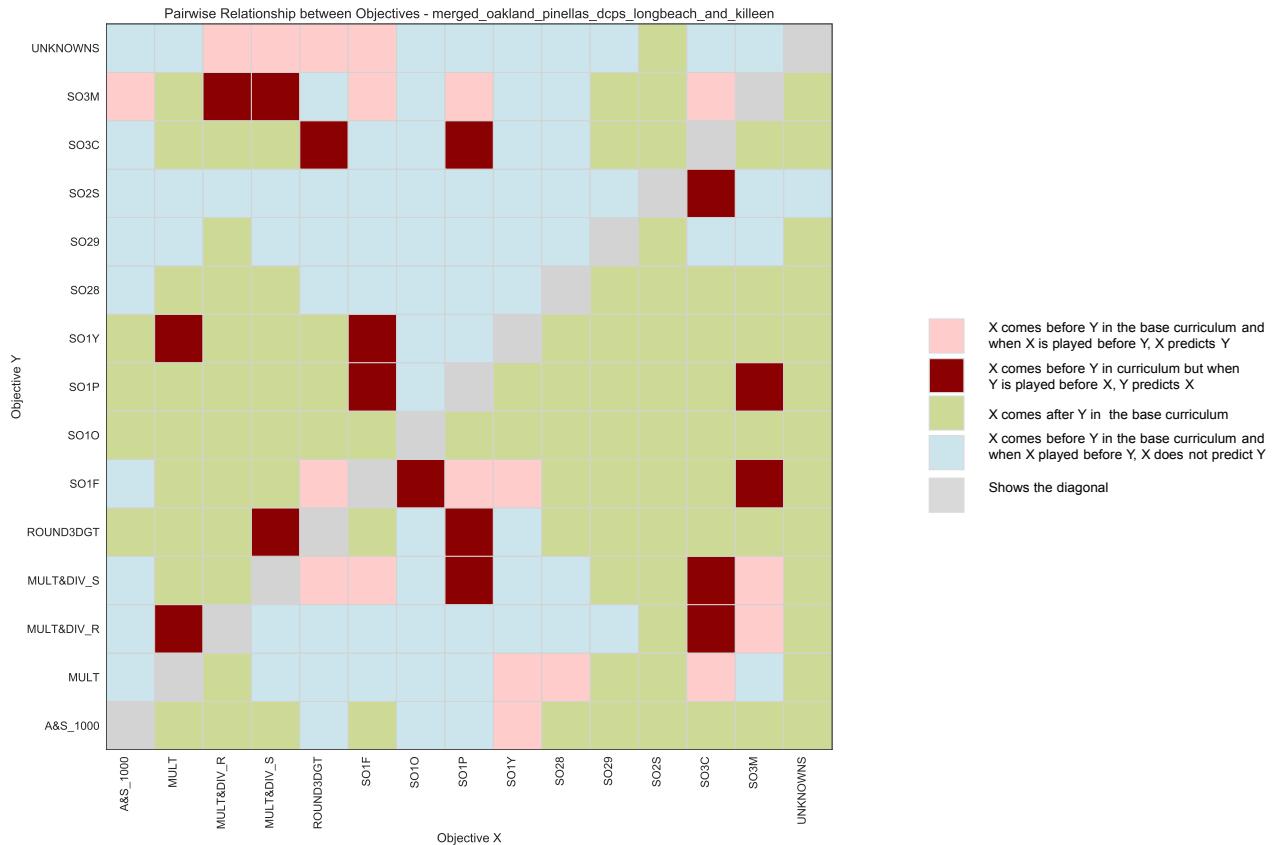


Figure 2: Pairwise Prediction of Objectives

area = 4 square in.
 perimeter = 10 in.

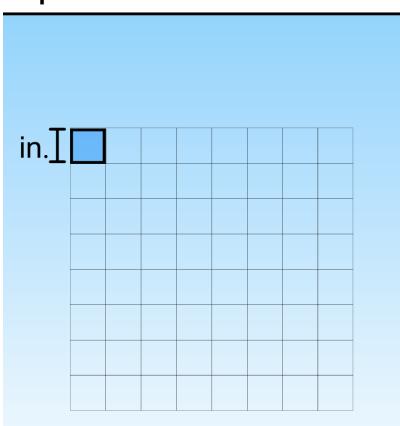


Figure 3: A Puzzle in Concepts of Area and Perimeter Objective (Copyright © 2019 MIND Research Institute)

more familiar with the game environment. Furthermore, some of these objectives, such as SO2S, provide prerequisite knowledge for SO3C, as indicated by a dark red tile in figure 2.

5.4 Group 3: Difficult Objectives

SO3C: "Concepts of Area and Perimeter" is a difficult objective no matter the order in the sequence. Although the tasks in SO3C require students to apply their knowledge of addition and multiplication, even districts with students that had performed well in SO1F: "Multiplication Concepts" and SO1Y: "Addition and Subtraction with Regrouping" had a hard time with SO3C. There may be game mechanic reasons for this, such as a potential lack of intuitiveness of the game. For example, figure 3 shows a puzzle that requires students to draw a rectangle with an area of 4 square inches and perimeter of 10 inches. Firstly, students might not easily figure out the question being asked, which exposes a game design issue. Secondly, the question requires a single answer that satisfies two different constraints, area and perimeter. Students might draw a 2x2 shape which gives a correct area but a wrong perimeter, because the correct answer is a 4x1 shape. We suggest that this objective should be played at the end of the sequence of 15 objectives. Students may need some form of guidance in playing higher levels of this objective. Although we have no definitive evidence, such as

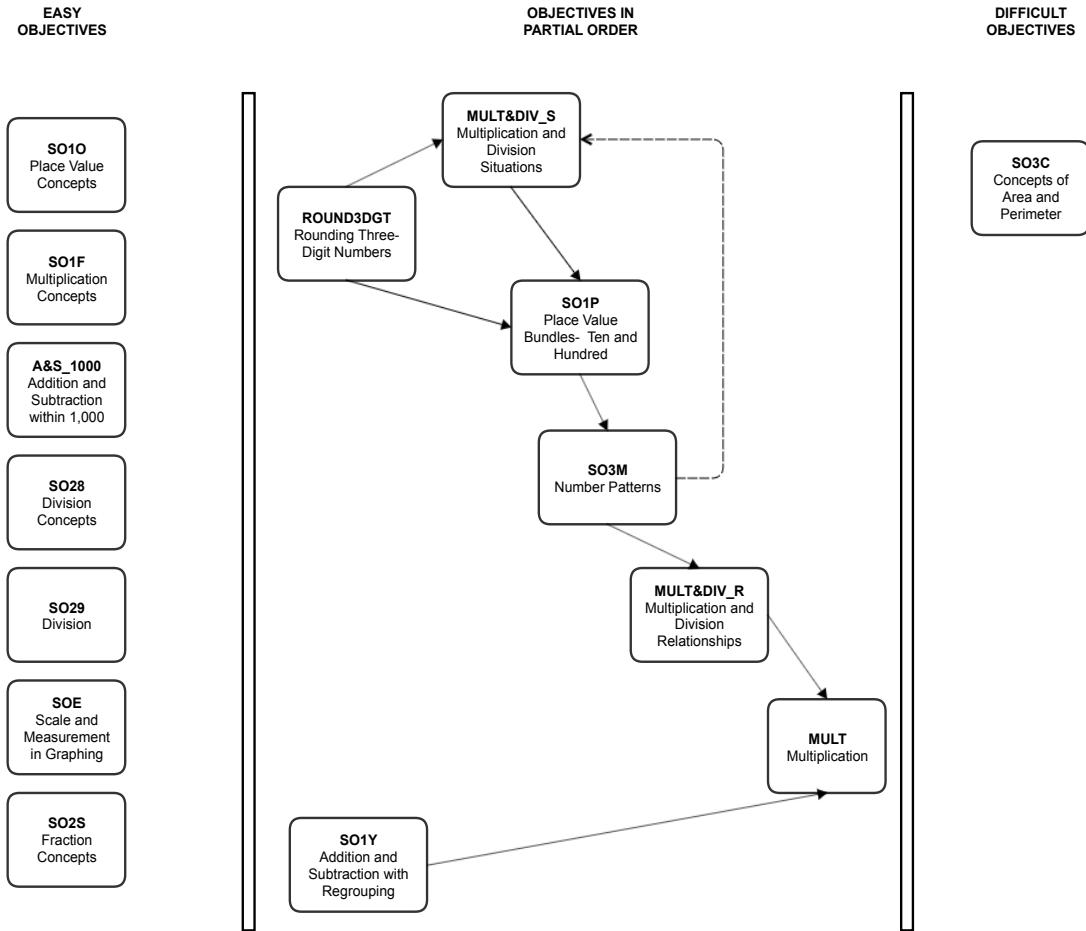


Figure 4: Partial Ordering of Objectives. Dotted Arrow Shows Anomaly to Partial Ordering

that from clinical interviews, as to why students struggled with this game, our results suggest that it may be beneficial to redesign this game to make it more achievable, or build prerequisite levels that assist students in relating perimeter and area.

5.5 Validating Previous Research

Previous research[21] labeled SO3M: "Number Patterns" to be a very difficult objective, because it presents students with difficult tasks, such as multi-step addition, subtraction, and multiplication. From this study, we not only confirm this, we also provide potential reasons behind this. "Number Patterns" requires students to count even and odd numbers forward e.g., 2, 4, 6, 8, 10 as well as backwards e.g., 9,7,5,3,1. The current study confirms that "Number Patterns" is difficult; however, it was difficult for only 2 districts as shown in figure 1. Because more districts are involved in this analysis, we discovered scenarios in 3 other districts where SO3M was easier. From our analysis, we discovered that SO3M was played

later in the curriculum sequence by these 3 districts. In addition, from our replication study, we also inferred that without the right prerequisite knowledge, as suggested by our partial order in figure 4, SO3M would be a difficult objective. Our study also validates the predictive relationship method used by Liu et al. [21] to suggest objective sequence. By using this method, we discovered partial orders that reduced objective difficulty, indicating that this method can be effective for suggesting objective ordering.

Other findings in the prior study [21] show that SO1Y: "Addition and Subtraction with Regrouping" predicts SO1F: "Multiplication concept" but does not predict SO29: "Division", SO3C: "Concepts of Area and Perimeter," or SO2S: "Fraction Concepts." This study confirms all these as well.

5.5.1 Anomaly in Partial Ordering. We recorded an anomalous relationship from our results. This abnormal relationship is represented by the dotted line in figure 4. It indicates that SO3M: "Number Patterns" should be played before MULT&DIV_S: "Multiplication and

Division Situations." However, other dark red tiles suggest that MULT&DIV_S should be played before SO1P and SO1P before SO3M. Another red tile suggests that SO1F should be played before SO3M. From this we can deduce that SO1F and MULT&DIV_S should be played before SO3M. We confirm from our analysis in step one that when SO1F and MULT&DIV_S are played before SO3M, SO3M becomes easier. But when SO3M is played before MULT&DIV_S, SO3M is difficult, so although one red tile suggests this order, we do not accept it, because it doesn't contribute to our goal of reducing the number of retries. One explanation for this anomaly could be that MULT&DIV_S alone does not directly predict SO3M, but a combination of SO1F: "Multiplication concept" and MULT&DIV_S predicts SO3M.

6 CONCLUSION

Our study uses a pairwise predictive technique to suggest objective orderings, validate prior work, and make recommendations that could reduce the number of retries within the ST Math objectives. This kind of student modelling is beneficial to both the serious game and analytics communities. In particular, modelling student retry patterns during gameplay across a varied population and different curricular implementations can suggest whether game difficulty is a result of curricular sequences that do not leverage natural prerequisite structures in the game or whether the objective game design is posing particular difficulty for students. In the future, a similar analysis could be performed for the bottom 17 objectives to suggest an ordering of the whole objective sequence for the ST Math Curriculum.

One limitation of this work is that we cannot be sure about the source of game difficulty for students. Since ST Math is a curriculum integrated math game intended to be used concurrently with the content of the school math curriculum, other factors outside of the ST Math system may affect the number of retries within each district. These factors include variation in district implementation of ST Math and different instructional methods for administering math curricula. For example, some teachers might teach pre-requisite skills in class before bringing students onto ST Math. Although we made classroom factor a random effect in our prediction analysis to account for classroom effects, future work could gather *ex situ* data to understand the impact of both class and district-wide outside-of-ST Math curricula on gameplay behavior.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1544273.

REFERENCES

- [1] 2000. *Linear Mixed-Effects Models: Basic Concepts and Examples*. Springer New York, New York, NY, 3–56. https://doi.org/10.1007/0-387-22747-4_1
- [2] 2007. *In Order to Learn How the Sequence of Topics Influences Learning*. Oxford University Press. <https://global.oup.com/academic/product/in-order-to-learn-9780195178845?cc=us&lang=en#>
- [3] Yetian Chen, José P. González-Brenes, and Jin Tian. 2016. Joint Discovery of Skill Prerequisite Graphs and Student Models. In *Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016, Raleigh, North Carolina, USA, June 29 - July 2, 2016*, 46–53. http://www.educationaldatamining.org/EDM2016/proceedings/paper_89.pdf
- [4] Yang Chen, Pierre-Henri Wuillemin, and Jean-Marc Labat. 2015. Discovering Prerequisite Structure of Skills through Probabilistic Association Rules Mining. In *Proceedings of the 8th International Conference on Educational Data Mining, EDM 2015, Madrid, Spain, June 26–29, 2015*. 117–124. <http://www.educationaldatamining.org/EDM2015/proceedings/full117-124.pdf>
- [5] Enrico W. Coiera, Elske Ammenwerth, Andrew Georgiou, and Farah Magrabi. 2018. Does health informatics have a replication crisis? *JAMIA* 25, 8 (2018), 963–968. <https://doi.org/10.1093/jamia/ocy028>
- [6] Catherine Ecock Connelly. 1986. Replication research in nursing. *International Journal of Nursing Studies* 23, 1 (1986), 71–77.
- [7] Mayer-Schoenberger V. Cukier, K. N. 2013. The rise of Big Data: How it's changing the way we think about the world. (2013), 92(3). <https://doi.org/10.1145/3116595.3116620>
- [8] Shayan Doroudi, Kenneth Holstein, Vincent Alevin, and Emma Brunsell. 2016. Sequence Matters, But How Do I Discover How? Towards a Workflow for Evaluating Activity Sequences from Data. In *Workshop and Tutorial Proceedings of the 9th International Conference on Educational Data Mining, Raleigh, NC, USA, June 29, 2016*. <http://ceur-ws.org/Vol-1633/ws3-paper5.pdf>
- [9] Looi CY Dowker A, Sarkar A. 2016. Mathematics Anxiety: What Have We Learned in 60 Years?. In *Front Psychol*. <https://doi.org/10.3389/fpsyg.2016.00508>
- [10] Michaela Greiler, Kim Herzig, and Jacek Czerwonka. 2015. Code ownership and software quality: a replication study. In *2015 IEEE/ACM 12th Working Conference on Mining Software Repositories*. IEEE, 2–12.
- [11] Erik Harpstead and Vincent Alevin. 2015. Using Empirical Learning Curve Analysis to Inform Design in an Educational Game. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY 2015, London, United Kingdom, October 5–7, 2015*. 197–207. <https://doi.org/10.1145/2793107.2793128>
- [12] Drew Hicks, Michael Eagle, Elizabeth Rowe, Jodi Asbell-Clarke, Teon Edwards, and Tiffany Barnes. 2016. Using game analytics to evaluate puzzle design and level progression in a serious game. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK 2016, Edinburgh, United Kingdom, April 25–29, 2016*. 440–448. <https://doi.org/10.1145/2883851.2883953>
- [13] Britton Horn, Amy K. Hoover, Jackie Barnes, Yetunde Folajimi, Gillian Smith, and Casper Harteveld. 2016. Opening the Black Box of Play: Strategy Analysis of an Educational Game. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY 2016, Austin, TX, USA, October 16 - 19, 2016*. 142–153. <http://dl.acm.org/citation.cfm?id=2968109>
- [14] K. Highfield R. Hunting S. Lee B. Perry L. Thomas J. Bobis, S. Dockett. 2010. Playing with Mathematics Play in Early Childhood as a Context for Mathematical Learning. https://www.researchgate.net/profile/Janette_Bobis/publication/267414818
- [15] Douglas H Johnson. 2002. The importance of replication in wildlife research. *The Journal of Wildlife Management* (2002), 919–932.
- [16] Andrea S Larsen and David H Olson. 1989. Predicting marital satisfaction using PREPARE: A replication study. *Journal of Marital and Family Therapy* 15, 3 (1989), 311–322.
- [17] Zhongxiu Liu, Christa Cody, Tiffany Barnes, Collin Lynch, and Teomara Rutherford. 2017. The Antecedents of and Associations with Elective Replay in An Educational Game: Is Replay Worth It?. In *Proceedings of the 10th International Conference on Educational Data Mining, EDM 2017, Wuhan, Hubei, China, June 25–28, 2017*. http://educationaldatamining.org/EDM2017/proc_files/papers/paper_32.pdf
- [18] Zhongxiu Liu, Rui Zhi, Andrew Hicks, and Tiffany Barnes. 2017. Understanding problem solving behavior of 6-8 graders in a debugging game. *Computer Science Education* 27, 1 (2017), 1–29. <https://doi.org/10.1080/08993408.2017.1308651>
- [19] Christian Loh, Sheng Yanyan, and Dirk Ifenthaler. 2015. *Serious Games Analytics: Theoretical Framework*. 3–30. https://doi.org/10.1007/978-3-319-05834-4_1
- [20] Zhongxiu Peddycord-Liu, Veronica Cateté, Jessica Vandenberg, Tiffany Barnes, Collin F. Lynch, and Teomara Rutherford. 2019. A Field Study of Teachers Using a Curriculum-integrated Digital Game. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI 2019, Glasgow, Scotland, UK, May 04–09, 2019*. 428. <https://doi.org/10.1145/3290605.3300658>
- [21] Zhongxiu Peddycord-Liu, Christa Cody, Sarah Kessler, Tiffany Barnes, Collin F. Lynch, and Teomara Rutherford. 2017. Using Serious Game Analytics to Inform Digital Curricular Sequencing: What Math Objective Should Students Play Next?. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play, CHI PLAY 2017, Amsterdam, The Netherlands, October 15–18, 2017*. 195–204. <https://doi.org/10.1145/3116595.3116620>
- [22] Zhongxiu Peddycord-Liu, Rachel Harred, Sarah Karamarkovich, Tiffany Barnes, Collin Lynch, and Teomara Rutherford. 2018. Learning Curve Analysis in a Large-Scale, Drill-and-Practice Serious Math Game: Where Is Learning Support Needed?. In *Artificial Intelligence in Education - 19th International Conference, AIED 2018, London, UK, June 27–30, 2018, Proceedings, Part I*. 436–449. https://doi.org/10.1007/978-3-319-93843-1_32
- [23] Deepak Rajan and Philip S. Yu. 2006. Discovering Partial Orders in Binary Data. In *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), 18–22 December 2006, Hong Kong, China*. 510–521. <https://doi.org/10.1109/ICDM.2006.57>
- [24] Teomara Rutherford, George Farkas, Greg Duncan, Margaret Burchinal, Melissa Kibrick, Jeneen Graham, Lindsey Richland, Natalie Tran, Stephanie Schneider,

Lauren Duran, and Michael E. Martinez. 2014. A Randomized Trial of an Elementary School Mathematics Software Intervention: Spatial-Temporal Math. *Journal*

of Research on Educational Effectiveness 7, 4 (2014), 358–383. <https://doi.org/10.1080/19345747.2013.856978> arXiv:<https://doi.org/10.1080/19345747.2013.856978>