Rapid Cycle Product Efficacy Assessment & Feedback Loop Pilots

Case Study: Avonworth School District

Evaluation Findings and Design Insights Report for Digital DreamLabs

By Soniya Gadgil, PhD & Marti Louw, MDes
Learning Media Design Center
Carnegie Mellon University

July 27, 2016

INTRODUCTION

Recent advances in technology offer new opportunities for schools to personalize learning for each student; however, choosing the best products to use in schools can be challenging. School districts are now spending a sizable portion of their budgets on ed tech products. In 2015, U.S. elementary, middle, and high schools spent $6.6 billion on ed tech[1]. The market is flooded with a multitude of instructional and software products, and while many school and district leaders recognize the potential of these products to help personalize learning for each student, they report high levels of uncertainty about how to select the right products for their needs[2]. Existing evidence about ed tech product effectiveness is scarce and school district leaders struggle to access, validate, and apply findings to their unique settings. Faced with limited reliable information, many districts would prefer to conduct local pilots as a way to generate useable product efficacy evidence. However, few school districts have the research capacity to conduct high-quality evaluations that yield the evidence district decision makers need to purchase a new ed tech product.

With support from the Bill and Melinda Gates Foundation and Digital Promise, three Pittsburgh, PA area school districts (Avonworth, Elizabeth Forward, and South Fayette) partnered with researchers from the Learning Media Design Center at Carnegie Mellon University to conduct ed tech product efficacy research during the 2015-2016 school year. In this report, we describe a case study are from the Elizabeth Forward school district, which piloted the Amplify suite of ELA and STEM games in its middle school classrooms.

The three main goals of this project were as follows: 1) to provide guidance and information for participating schools to use when planning pilots, 2) to provide useful measures and practices to support current and future educational technology pilots, internal decision-making, and adoption processes, and 3) to provide developers a valuable, external reflection on their products-in-use. We relied on multiple sources of qualitative and quantitative data to study product efficacy, using a mixed methods approach. While we hope that our findings would provide guidance to other schools make informed decisions about choosing ed tech products, we also recommend that careful attention be paid to each school's unique contextual factors before extrapolating findings from this research to other situations.

[1] *http://www.centerdigitaled.com/higher-ed/US-Education-Institutions-Spend-66-Billion-on-IT-in-2015.html*

[2] *http://digitalpromise.org/wp-content/uploads/2016/02/IDEO-Digital-Promise-Report-Evolving-Ed-Tech-Procurement-in-School-Districts.pdf*

CASE STUDY: AVONWORTH SCHOOL DISTRICT

Avonworth School District, a member of the Digital Promise League of Innovative Schools is a small, suburban school district to the north of Pittsburgh. It was ranked 61st out of 498 Pennsylvania school districts in 2012, by the Pittsburgh Business Times. The total enrollment at the district is approximately 1600 students, of which 15% are from economically disadvantaged households (see Table 1 for district snapshot). A partner of the Remake Learning network, a local collaborative of network of educators and innovators working together to shape the future of teaching and learning in the Greater Pittsburgh Region, Avonworth places a strong emphasis on "...collaboration, technology integration, authentic application and reflection, while inspiring creativity" (cf. remakelearning.org).

Table 1: Avonworth School District Snapshot

Number of Students Served	School Ranking	% Free or reduced lunch	Products Piloted	Target Subject Areas	Grade-Levels in Pilot
1600	61/ 498	15	eSpark, Puzzlets	ELA, Digital Literacy	K,1,2

The district had two distinct goals for choosing products for this pilot.

The first goal was to show sustained student achievement growth, particularly in the ELA domain. Since the past three years, Avonworth Primary Center has used the MAP assessment by NWEA's measures of Academic progress (MAP) three times a year to assess student growth. The eSpark platform that creates a personalized learning plan for each student based on his or her NWEA score provided a promising alternative to curating individual apps, and investing in seat licenses for each student. eSpark also came strongly recommended by another member of the League of Innovative Schools in the region, who had used eSpark with students in lower elementary grades. As a result, Avonworth chose to pilot eSpark in six first grade classrooms.

A second goal was to identify a product that would provide an engaging, early introduction to computational thinking concepts such as sequential thinking, reasoning, and problem solving to students in grades K-2 in their Digital Literacy classrooms. To this end, the district chose the product Puzzlets — a tangible learning game system focused on providing an early introduction to sequential thinking and problem solving in K-2 Digital Literacy classrooms. Puzzlets is developed by a local ed tech company called Digital Dreamlabs that started out of Entertainment Technology Center at Carnegie Mellon University, and is also a member of the Remake Learning network. The lead teacher for Digital Literacy at Avonworth also acts as consultant to Digital Dreamlabs. Although Puzzlets in its current form is intended for the consumer market, and is in the process of being adapted for classroom use, the district felt comfortable choosing to pilot it, given the local connection of Digital Dreamlabs, as well as the strong recommendation from the lead teacher.

The research team worked closely with the school district to identify the goals and objectives of the pilot, develop assessment instruments to measure key product efficacy dimensions, collect data, and provide feedback to the product developers. The research team did not have a say in the choice of product, because the products were chosen prior to their engagement in the pilots. The primary point of contact at Avonworth for both pilots was the building principal of Avonworth Primary Center. For the eSpark pilot, the research team interfaced with seven teachers, one from each of the six classrooms, and one special education teacher. For the Puzzlets pilot, they worked with the lead teacher for Digital Literacy who taught all three classrooms in the pilot.

Given the disparate curricular goals and learning objectives of each product, the research team developed custom assessments strategies to measure key dimensions of product efficacy — student learning, student engagement, teacher support, teacher satisfaction, and administrator satisfaction. Methods used included surveys, focus groups, in-depth interviews with students, teachers, and administrators, classroom observations, and analysis of log data provided by product developers. Both products were piloted with younger students from grades K-2, so written surveys were not used to measure student learning and engagement, but classroom observations and short interviews with students conducted during classroom observations were relied on.

Table 2 provides an overview of the different measures used to study product efficacy for each of the products piloted by Avonworth, noting when a measure was deployed more than once.

Table 2: Overview of research methods and measures used

		Products	
Dimension of Product Efficacy	**Measure**	**eSpark**	**Puzzlets**
Student Learning	Log Data	✓	✓
	Pre-Post Survey		
	Post-test		✓
	NWEA assessments	✓	
	Interview / Focus Group	✓	✓ (x 2)
Student Engagement	Log Data	✓	✓
	Survey		
	Interview / Focus Group	✓	✓ (x 2)
	Classroom Observations	✓	✓ (x 2)
Teacher Support	Survey	✓ (x 2)	
	Interviews/ Focus Groups	✓ (x 2)	✓
	Observation of PD sessions	✓ (x 3)	
Teacher Satisfaction	Survey	✓ (x 2)	

	Interviews/ Focus Groups	✓ (x 2)	✓
Administrator Satisfaction	Interview	✓	✓

See Figure 1 for a timeline for research activities at Avonworth. In addition, the research team conducted monthly check-in calls or visits for each of the products with the teachers and administrators.

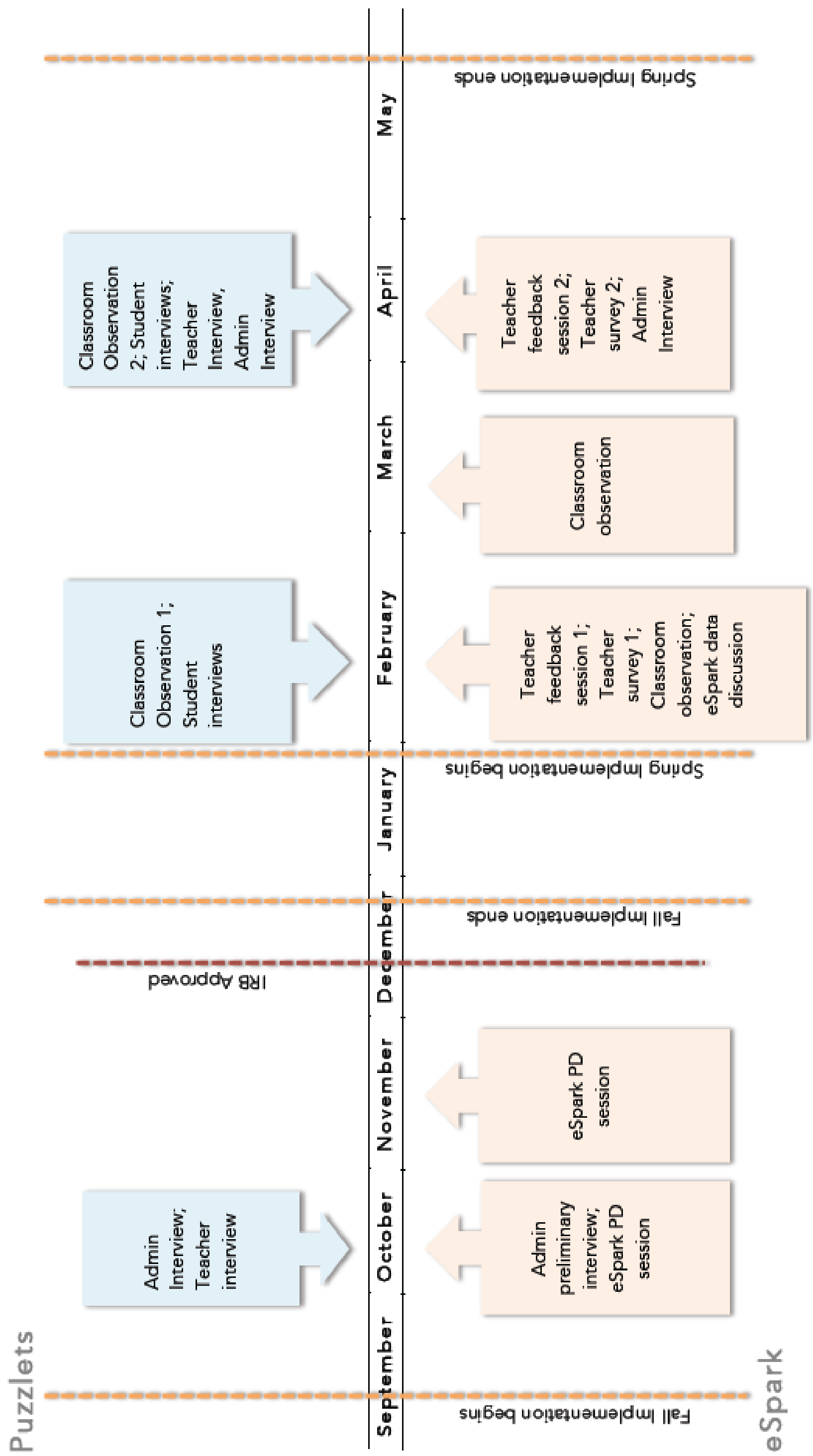

Figure 1: Avonworth Pilots Timeline

PRODUCT OVERVIEW & PILOT GOALS

Puzzlets is a hands-on learning game system that includes 22 "*Puzzlet*" tiles that fit into a play tray, which connects wirelessly via a Bluetooth connection to an iPad game app "Cork the Volcano". Using operator (directional) and variable tiles, players are required to map out a sequence of moves that enable the character to navigate through a series of screen-based scenario challenges that become increasingly more complex as players complete levels. Based on interviews with the cognizant school administrator and teacher, the established goals for the *Puzzlets* pilot at Avonworth school district were as follows:

- Test out *Puzzlets* learning affordances for communication, collaboration, creativity, and computational thinking in K-2 classrooms (teacher interview, principal interview, grant proposal).
- Evaluate whether *Puzzlets* provided an engaging alternative to other educational technology courseware

CURRICULAR NEED

Avonworth Primary Center is engaged in developing a new curriculum for Digital Literacy, focused on bringing 21st century skills such as communication, collaboration, creativity, and to provide early exposure to learning activities that support the growth of computational thinking skills. The school wanted to pilot a product that would support the above objectives, while also being engaging to students. The former technology curriculum focused primarily on basic computer skills such as keyboarding, mouse and software tool use, which was neither very engaging for students, nor was preparing them sufficiently for future-oriented fluency with technology and computing.

PRODUCT IMPLEMENTATION PROCESS

One classroom each from grades K-2 was chosen to participate in this pilot. Each class rotated through a 14-week curriculum, meeting once every six days for 45 minutes. The teacher used a station rotation model where children rotated through 3 station setups *Puzzlets*, *Type Rocket* and *ScratchJr*, and were able to spend approximately 15 minutes at each. After the first two weeks, the Scratch Jr. station was dropped as the teacher felt students mostly spent their time coloring or doing tasks that did not contribute to their learning computational skills. As a result, after the first two weeks of the implementation, students worked on two stations instead of three, spending between 15-20 minutes at each. They worked on *Puzzlets* in pairs, working with the same partner throughout the course of the pilot.

Figure 2: Student pairs engaged in Puzzlets gameplay

Only those students whose parents or guardians had provided consent participated in the study. Unconsented students engaged in classroom activities, but were not interviewed and their data was not included in the final data analysis. Sixteen (16) kindergarteners, twenty-two (22) 1st graders, and eighteen (18) 2nd graders participated in the study, for a total of fifty-six (56) participants.

RESEARCH QUESTIONS & STUDY DESIGN

For the *Puzzlets* product pilot, the research examined the four pre-defined product efficacy dimensions—student learning, engagement, teacher support, and satisfaction, along with an additional dimension of administrator satisfaction—using a mixed methods approach. Two rounds of classroom observations were conducted in addition to student interviews, and semi-structured teacher and administrator interviews at the beginning and end of the pilot period. Log data made available by the developers were analyzed to examine students' progress over the course of the pilot. In addition, as this was a product initially developed for the home market and in the early stages of being adapted for classroom use, the research team conducted a standard heuristic analysis of product usability and learning affordances to generate a set of design recommendations.

Two classroom observations were conducted— one at the beginning of the implementation (first week of February), and one at a midpoint (first week of April). These observations were designed to be a window into student engagement and student learning. For the first classroom observation, two researchers visited the classrooms, and made structured observations using coding sheets to capture measures of student learning and engagement which were developed in consultation with the teacher (see Appendix 1 for Classroom Observation Protocol and Appendix 2 for Observation Recording Sheet). Six pairs of students from each class were randomly selection. The researchers audio-recorded their talking interactions, and made notes about behaviors, gestures and teacher interactions on observation sheets. Additionally, as the end of a play session students were asked a series of questions about whether they liked using *Puzzlets*, whether it was easy or hard, and then gave them a choice task to indicated their preference for playing *Puzzlets*, *Scratch Jr* and *Type Rocket* (see Figure 3).


Figure 3: Screenshots of interfaces of Puzzlets, Type Rocket, and Scratch Jr.

After each rotation was over, researchers noted whether students chose to continue playing or moved on to the next rotation immediately. We also looked for technical challenges and points of intervention by the teacher.

During the second classroom observation, students were asked to complete an additional debugging task in *Puzzlets*. The *Puzzlets* board was presented to the students with an incorrect sequence of tiles (See Figure 4), and students were asked to rearrange the tiles such that the character will move in the correct fashion.

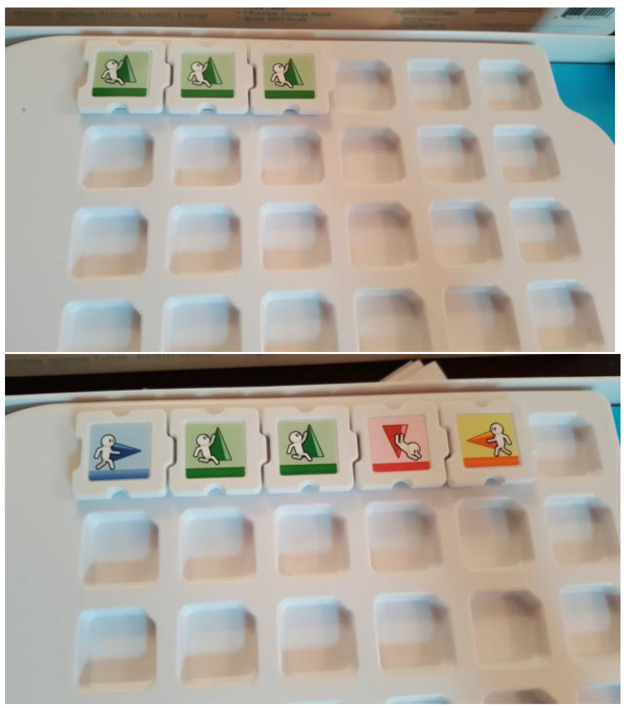
Figure 4: Debugging tasks presented during Classroom Observation 2

In addition, during both classroom observations, researchers looked for evidence of collaboration, and noted technical challenges experienced.

EVALUATION FINDINGS

In the following sections, findings across each of the product efficacy dimensions — student engagement, student learning, teacher support, teacher satisfaction, and administrator satisfaction are summarized.

To understand student learning and engagement, the research team relied on various qualitative and quantitative measures. Qualitative measures included student interviews and classroom observations, and quantitative measures included a debugging task designed to capture students' understanding of correct and incorrect sequences, and analyses of log data provided by Digital Dream Labs. Log data provided information on event properties, such as time taken to complete a level, number of pieces used to construct a sequence, number of tries needed to complete a level, number of raindrops (additional incentives embedded in game) collected, in addition to user properties such as grade level, log in and log out times etc. The log data is at the level of student pair.

Student Engagement

Evidence from interviews and classroom observations indicated that students across the three grade levels were highly engaged and enjoyed playing *Puzzlets*. The level of engagement was sustained from the first observation (conducted on the first session of game-play), to the second classroom observation (conducted on the fifth session of game-play). Based on log data, students in each grade played *Puzzlets* for a total of approximately 60 minutes across five sessions from February to early April.

Some students commented that they liked the game-playing aspect of *Puzzlets.*

> *"It's fun…it's like a little video game."*

--1st grade student on Puzzlets

Others liked the tangible interaction with the play tray

> *"… the tray makes it really fun… you can make the character on the screen move."*

-- Kindergarten student on *Puzzlets*.

For the station-choice task in our first classroom observation, 27 out of 36 students selected *Puzzlets* as their first choice, 6 chose *Type Rocket*, and 3 chose *ScratchJr*. About a third of the student pairs continued to play *Puzzlets* even after the rotation was signaled to be over, and a few others moved on to the next rotation, but expressed disappointment about doing so.

Student Learning

Analysis of log data revealed that while all students progressed in the game since they started playing, their progress differed by grade level. All students started playing at level 1, and each level contains 7-8 sub-levels. The highest level reached by kindergartners was 2-7, by first-graders was level 3-1, and by second graders was 3-5.

The average number of tiles used to construct a sequence, and how this number can be used a proxy for complexity of sequences constructed. Figure 5 shows the average number of tiles on the final sequence for each day of game play broken up by grade. Error bars indicate standard errors. From the graph, it is seen that on day 1, students used close to three tiles on average to construct the final sequence, while on day 5, they used over five tiles on average to construct their final sequence. There was a trend of grade 2 students using slightly more tiles on average, than grade 1 and kindergarten students; however, this difference was not statistically significant.

Note that this finding should be interpreted with caution, because as students gain more expertise on the game, the number of tiles used may not be associated with complexity of sequence constructed, as they learn to accomplish the same result with a fewer number of tiles.

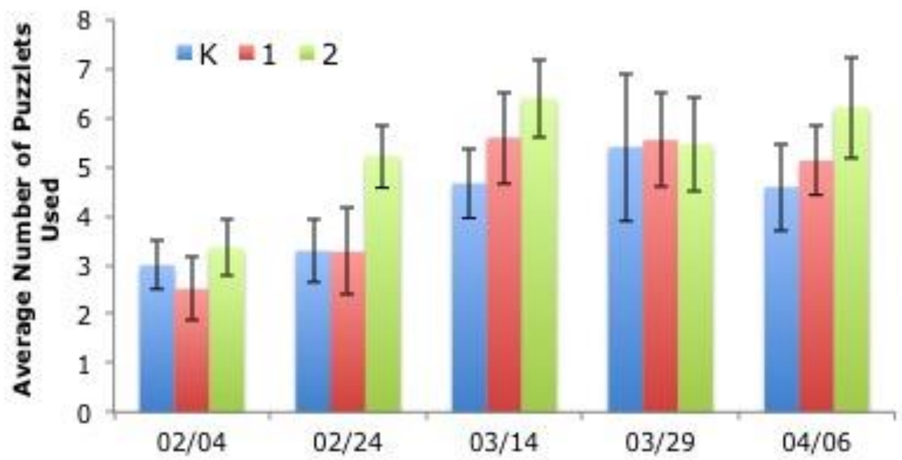

Figure 5: Average number of Puzzlets used by grade

An alternative way of assessing complexity of final sequences constructed would be to look at how many times the operators — x2 and x3 were used. These are relatively complex operators, compared to the directional tiles. Out of 784 final sequences constructed, students used either the x2 or the x3 operator 58 times. Figure 6 shows the breakdown of frequency of use of complex operators by grade-level. Second-graders were over twice as likely as first-graders and nearly thrice as likely as kindergartners to use the complex operators in their final sequences. Thus, older children are more likely that younger ones to adopt and use the more complex operator.

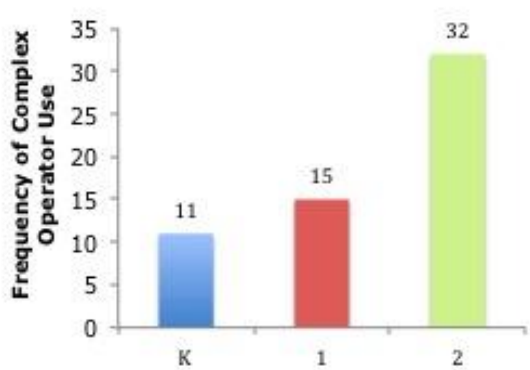

Figure 6: Average number of complex operators used by grade

Performance on the debugging task also showed some differences by grade level. All pairs (3 kindergarteners, 6 first-graders, and 5 second-graders) correctly solved the first debugging task. One out of two pairs of kindergartners correctly solved the more complex second debugging task correctly. Among first graders, all six pairs correctly solved the second debugging task, but one of the pairs required several tries. Among second graders, five out of six pairs solved it correctly, and one pair ran out of time.

Although not a controlled experimental task with random assignment and carefully monitored times on task, the debugging task provides some insight into student progress and their problem-solving skills within the game. Students' progress through the game levels and performance on the debugging task suggests that they learned problem-solving strategies within the game context after four in-class sessions using the game.

Teacher Support

Teacher support was assessed based on interviews with the teacher and principal (See Appendix 3 for Teacher Interview Protocol and Appendix 4 for Administrator Interview Protocol), and communications with Digital Dream Labs. Overall, teacher support provided by the company was timely and helpful. Individual accounts for students are not currently available, but Digital Dreamlabs set them up for student pairs in the three classrooms being observed, so that their data could be tracked systematically over the course of the pilot. The lead developer from the company was also present during the first classroom observation, to help with getting started and troubleshooting issues.

Given that the product is still being adapted for use in classrooms, the company does not have a fully implemented professional development plan and supporting materials and resources for teachers. While this was not a problem in the present pilot as the teacher has been the educational advisor for the product and is very familiar with its use in the classroom. Teacher support and tested training materials will need to be produced for larger-scale adoption in schools.

Teacher Satisfaction

Teacher satisfaction was examined through semi-structured teacher interviews. Only one teacher was involved in this pilot. As previously noted, this teacher acts as an advisor to the product developer, Digital Dream Labs, and has conducted her Master's thesis research based on *Puzzlets*. Despite these confounding factors, the key findings relating to teacher satisfaction were as follows:

- The teacher noted that she got a lot of positive feedback from students and parents, and that she found students to be engaged when working on *Puzzlets* in the classroom
- Students are increasingly getting more competitive when using *Puzzlets*
- Avonworth is not a 1:1 iPad school, so sharing of devices is necessary. *Puzzlets* was not originally designed for classroom use, so currently a maximum of only three logins is available on each device. So, students' work can't be saved, and often students have to play resume a level played by a student who used device before them.
- Being able to see the data on student progress would be a big plus.

Self-reported satisfaction is not necessarily indicative of teachers less involved with the product development. As more teachers get involved in using this product in the classroom, there will be an opportunity to study this dimension in greater detail.

Administrator Satisfaction

Administrator satisfaction was assessed through a semi-structured interview with the principal of the Primary Center (see Appendix 4 for complete instrument). Overall, the principal was satisfied with the product. He noted that the developers were responsive to suggestions, and were appreciative of feedback. They talked directly to the kids using *Puzzlets*, and gathered feedback and suggestions. *Puzzlets* was an exceptional case in that, the district would not normally choose to pilot a product that is still in development and does not contain basic functionality for an ed tech product, such as a teacher dashboard, and student log-in screen. However, the teacher's deep knowledge of the product, as well as the developers' responsiveness were key factors for the principal to decide to pilot *Puzzlets.*

In terms of suggestions for improvement, the principal noted that the curriculum/ scope and sequence needs to be more developed. There also needs to be a plan for more comprehensive professional development, which will guide the teachers on getting started, diagnosing and troubleshooting issues, and other relevant topics.

Additional Observations / Findings

Collaborative Behaviors

The evidence for collaboration was mixed. During the first round of observations, players did not seem to communicate very much with each other about strategies. Most often, one player took control of the *Puzzlets* board, while the other took control of the screen. The division did not seem to be the most beneficial for collaboration, because one player generally decided the strategy, while the other executed.

In the time between the first and second rounds of observations, the teacher introduced some strategies for collaboration. Of of the strategies involved one player taking on the role of "driver" and the other the "passenger", and switching roles midway between their play. Researchers noted more evidence of turn taking and discussion about strategy was observed in the second round of classroom observations.

Technical Challenges

Technical challenges were frequent during the first classroom observations. The Bluetooth connection kept dropping off for some of the students. This product has not yet been optimized for classroom use, where there will be multiple devices in operation at the same time, so interference of signals across devices is possible. Out of 18 pairs observed, 4 pairs had issues with the Bluetooth connection, whereas two others had issues with logging in. Technical challenges were less frequent in the second round of observations. Only one of the pairs had to stop the game because of an issue with the Bluetooth connection.

DESIGN INSIGHTS

To align the product better with the curricular goals of Avonworth, the following design recommendations were provided to Digital Dreamlabs:

- To encourage greater collaboration and communication, it may be helpful for the game interface to prompt switching, or to otherwise create more multiplayer aspects that necessitate the input of both players.
- Sequential thinking:

    - Feedback on accuracy of the sequence constructed by the students would promote the learning of sequential thinking skills. For example, currently, players can place multiple tiles and, upon execution, only use the first few. However, they still pass the level, and receive no feedback that they placed more tiles than necessary.
    - The [left] and [right] tiles indicate both direction and movement, which can be unintuitive.
    - The [up] key indicates jumping, while the [down] key is stop. Several students in our observations did not know that down denotes stop, but thought that it makes the character jump down. Making this consistent and intuitive would support development of reasoning skills.
- Creativity: Currently, the game offers bonuses in each level for fewest tiles, shortest time, etc. However, these bonuses are not obvious, and they are not necessary for progression in the game. It may be helpful to bring them to the forefront and make them requirements on some levels to encourage players to engage in planning and deeper thinking to come up with creative solutions to the problems.
- Teacher dashboard: *Puzzlets* being a product in an early stage of development currently does not come equipped with a teacher dashboard. Consequently, there is no way to save a students' work, which could be an issue for schools that do not have 1:1 tablets. There is no way to create class-wise rosters, so teachers can keep track of students' work. Teachers also need to be able to access key metrics of student usage and performance, such as
    - Log in and log out times
    - Turn taking
    - Current level
    - Identifying points of intervention if a student is stuck on a level
- Each student should have their own login so that they can save their work from one session to the next, and their work can be assessed better.

A comprehensive professional development plan, which will guide the teachers on key issues such getting started, diagnosing and troubleshooting issues, and keeping track of students' progress would be helpful.

FEEDBACK LOOP

The research team shared their feedback with Digital Dream Labs in a meeting with their lead developer and CEO. In this meeting they discussed the usability issues identified in heuristic analysis, and went over our design insights and recommendations. They also provided Digital Dreamlabs with references to academic research on tangible interfaces to teach programming (e.g., TIDAL group at Northwestern University led by Michael Horn and Kinderlab at Tufts University led by Marina Umaschi-Bers) who are conducting rigorous work on how such interfaces can encourage computational thinking and reasoning. The feedback will be shared with the rest of the development team at Digital Dream Labs, and would feed into the next product development cycle.

CONCLUSION

As an early-stage product, *Puzzlets* shows promise in promoting 21st century skills of communication, collaboration, creativity, and computational thinking in students. In order to improve its educational value, we made several recommendations on usability, intuitiveness of experience, increasing complexity, and promoting communication and collaboration. To improve teacher support, the company is in the process of developing a teacher dashboard, through which teachers can access data on students' progress. The company is also developing a comprehensive professional development plan, which will help teachers integrate it into their curriculum.

The teacher and administrator at Avonworth were satisfied with the student engagement aspect of *Puzzlets*. They were also satisfied with the level of support provided by the developers, and plan to continue the partnership with Digital Dreamlabs for the next school year.

Computational skill development is an expressed goal and desired outcome for this product. To better assess and identify student learning with *Puzzlets* it would be important to determine which knowledge components of computational thinking are being targeted. A pretest and a posttest could assess students' knowledge of those components gained through using *Puzzlets*. Additionally, providing teachers and students with vocabulary to tie *Puzzlets* play experiences to computational expressions and terms such as operators, sequencing, and parsimony would help increase the promise and value of this product as an educational tool. With these learning goals and constructs defined, future studies could examine more in-depth the learning potential of *Puzzlets* as an ed tech product that supports the development of computational literacy skills in learning play.

APPENDICES

Appendix 1: Classroom Observation Protocol

Digital Promise League of Innovative Schools
Product Efficacy and Feedback Loops Project

Avonworth School District: Puzzlets Student Observation Study

Purpose:
The goal of this study is to observe students using the *Puzzlets* (DDL) platform to look for markers of interest and engagement, problem solving strategies, debugging skills, collaborative behavior.

Participants:
Students in K-2 Digital Literacy classes at Avonworth Primary Center whose parents/ caregivers have consented to participation in study.

Protocol Description:
The lead teacher, uses station rotation model in her digital literacy classes to ensure all students have access to the ed tech learning experiences. The three stations available in the computer lab are Puzzlets, Scratch Jr., and Type Rocket. Each station has four game setups available, and student work in teams of two. Each rotation runs 10-12 minutes long. The maximum number of students in the lab is 3 X 4 X 2 =24. We will observe students using Puzzlets only. There will be 6-8 students (3-4 pairs) using the Puzzlets during every rotation who will be divided teams. There will be two CMU researchers, and each researcher will focus two groups of student pairs.

The researchers will observe *Puzzlets* game play for first 5 minutes. If the children fall silent, the researchers may prompt them to talk by asking "can you tell me what you are doing?" Mid-task, the researcher will ask the first team questions 1-5 (below) which should not take longer than 3-4 minutes. The researchers will then move to the second team and ask the same set of questions. The first group can resume play after responding to the researcher questions. At the end of the rotation (12 minutes), the researchers will hand out participation stickers. Non-participating children will receive a different kind of stickers. We plan to take over the shoulder and non-identified photographs of the students, and audio record their answers to questions.

Observation Categories (see Observation sheet):

- Play Strategy: Planned vs. trial and error
- Problem solving (debugging strategies used)
- Collaborative Talk & Behaviors (turn-taking, disagreement resolution)
- Off task behaviors
- Tech challenges (e.g. the Bluetooth connection, log-ins)

Tally Sections (Footer; see Observation sheet):

- Questions & requests to teacher
- Turn taking (Are the children taking turns, i.e., does one child handle the iPad and the other the Puzzlets board and switch during the rotation?)
- Session Close: End/Continue Play

Questions for Students:

1. Can you explain to me how this game works?
2. Can you tell me what these tiles do?
3. What do you think of this game?
4. How do you like playing this game?
5. Which one of these games do you most like to play (show three start pages)?

Materials:

- *Cameras + batteries*
- *Observations sheets*
- *2 DVR recorders + batteries*
- *Consent forms*
- *Clipboards*
- *Participant and nonparticipant stickers*
- *Pens*
- *Watch/Timer*
- *Screenshots for each of the three stations*

Appendix 2: Classroom Observation Recording Sheet

Digital Promise League of Innovative Schools
Product Efficacy and Feedback Loops Project

Avonworth School District: Puzzlets Student Observation Recording Sheet

Student ID:	Observer:	Instructor:	Date:
Start Time:	End Time:	DVR #:	Photo Consent?
Type of Play Strategy (ex. trial and error)			
Problem-Solving (debugging strategy)			
Collaboration			
Off-Task Behavior			
Tech Challenges			
Ask for Assistance		Turn-Taking	

Session End: Ended Play Continued Play

Post-Task Questions:

1. Can you explain to me how this game works?

2. *Can you tell me what these tiles do? (Note: Point to a few of the directional tiles, as well as the X 2 and X 3 tiles)*

3. *What do you think of this game?*

4. *What do you like (or not like, depending on answer to question 3) about this game?*

Notes:

Appendix 3: Teacher Interview Protocol

Digital Promise League of Innovative Schools
Product Efficacy and Feedback Loops Project

Avonworth School District:
Teacher Interview Questions

The goal of this interview is to talk about the Puzzlets pilot from a teacher's standpoint. I'll begin by asking you some questions about your pilot process in general, then some questions that are specific to Puzzlets, and then finally some reflection questions about what went well, and what can be improved. We have covered some of this information in our previous conversations, but I'd like to be comprehensive, so don't worry if you feel like you are repeating some things I may already know.
Do you have any questions before we begin?
Do I have your permission to record this interview?

School EdTech Decision-making

1. How does your school make a decision to adopt a new ed tech product? What stakeholders are involved, and in what ways?
2. What is the process for teachers and students to provide feedback to decision-makers, and how influential is that feedback?

Product Implementation

3. How did your school decide to use Puzzlets?
    a. *How did you get involved? Who else is involved? What are their roles?*
    b. How did it differ from other pilots?
4. What were the specific goals for piloting Puzzlets? In other words, what challenges were you hoping it would address and what would success look like?
5. What learning objectives does this product address?
6. How did you teacher(s) integrate Puzzlets into your curriculum? What adaptations or modifications have be made over the course of the pilot? (e.g., Can you talk a bit about how you decided upon the number of stations; weekly rotations vs rotations within a single class)
7. Can you comment on how students interacted with Puzzlets and how that changed over the course of the pilot?
8. To what degree has Puzzlets been successful in meeting your original goals and objectives?
9. What challenges did you face during the implementation of Puzzlets? Would you be doing anything differently for the next round?
10. What changes to Puzzlets would make it a better product for schools hoping to use it in the future?
11. Your experience may have been a little different given that you were so intimately familiar with the product. But do you see teachers who have not have experience with it?
12. Is there any other feedback you'd like to give Digital Dreamlabs?
13. How likely are you going to want to use this product again, if the changes you suggested were made?
14. What advice would you share with teachers, administrators, or edtech companies about conducting educational technology pilots?

Appendix 4: Administrator Interview Protocol

Digital Promise League of Innovative Schools
Product Efficacy and Feedback Loops Project

Avonworth School District:
<u>Administrator Interview Questions</u>

The goal of this interview is to understand the product efficacy pilots at Avonworth from an administrator perspective. I will begin by first asking questions about your pilot process in general, and then focus on each of the products that were piloted this year.
Do you have any questions before we begin?
Do I have your permission to record this interview?

School EdTech Decision-making

1. How does Avonworth go about choosing an edtech product?
2. Please describe is your pilot process.
3. What challenges do you typically face when piloting products?
4. How does your school make a decision to adopt a new edtech product? What stakeholders are involved, and in what ways?
5. How would you characterize the criteria, evidence or insights you are using to make your decisions? What factors do you weigh?
6. What is the process for teachers and students to provide feedback to decision-makers, and how influential is that feedback?

Puzzlets

1. How did your school decide to use Puzzlets?
2. How did you find out about Puzzlets? Who else was involved? What are their roles?
3. How does it differ from other pilots?
4. What were the specific goals for piloting Puzzlets? In other words, what challenge were you hoping it would address and what would success look like
5. What learning objectives does this product address?
6. To what degree has Puzzlets been successful in meeting those goals?
7. Please tell us how your teacher(s) has integrated Puzzlets into your curriculum? What adaptations or modifications have be made over the course of the pilot?
8. What challenges did you face during the implementation of Puzzlets? Would you be doing anything differently if you had to do it again?
9. How satisfied are you with Puzzlets at this point?
10. What changes to Puzzlets would make it a better product for schools hoping to use it in the future? Any other feedback you'd like to give Puzzlets?
11. How likely are you to use this product again, if the changes you suggested were made?