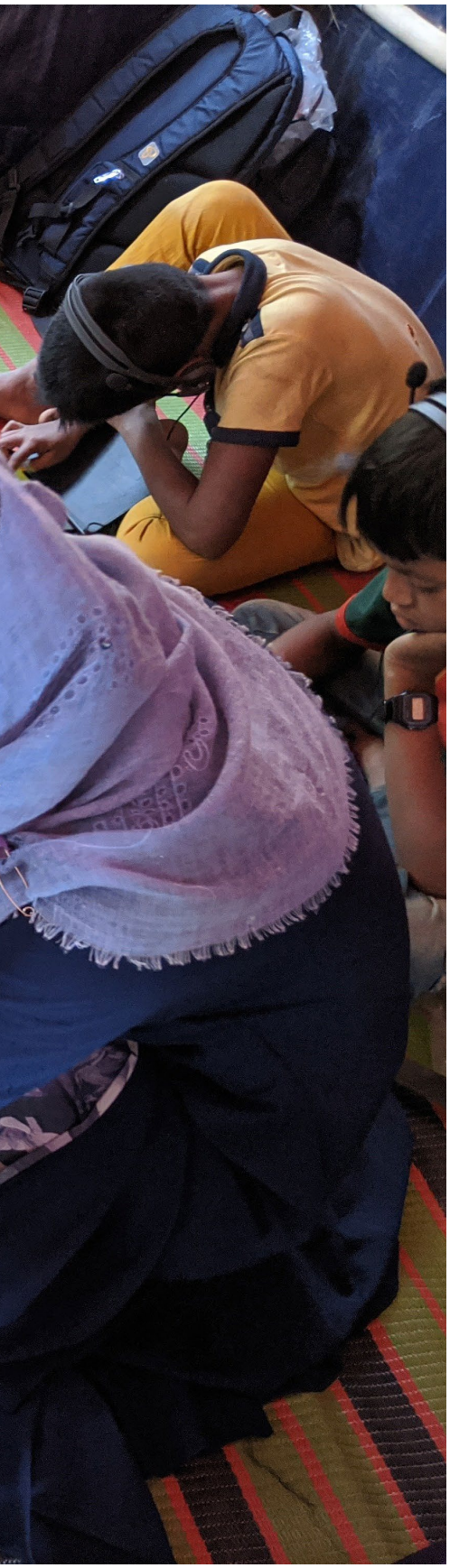
pop up

Pop-Up Learning

Design and implementation research report

YEAR 1

APRIL 2019 – APRIL 2020



This work was made possible with the generous support of our funders and partners:

The original Pop-Up donor cohort consisted of **individuals from Silicon Valley and San Francisco firms and academic institutions** who provided seed capital.

John O'Farrell & Gloria Principe

Aaref & Fariha Hilaly

Nick & Asena McKeown

Philip & Alicia Hammarskjold

F. Bennett Cushman II

Spring Point Partners provided flexible funding for Airbel's education work, which enabled the pilot work of Pop-Up.

Imagine Worldwide was a key partner providing design, implementation and technical support, ongoing thought partnership, and funding for the localization of the Kitkit School software.

The Kovner Foundation has invested flexible funding over five years to Airbel's research and innovation portfolio including the work of Pop-Up.

The Arnhold Foundation has invested flexible funding over five years to Airbel's research and innovation portfolio including the work of Pop-Up.

AUTHORS

Laura de Reynal, Sylvia Sable, Mostafa Kamal, Subrata Khisa, Silvia Diazgranados, Hira Siddiqui

Executive summary

Millions of children in low- and middle-income countries are being failed by the education system and not reaching their development potential. The "learning crisis" faced today is one of access and quality that necessitates reimagining where, how, and with whom learning can take place.

In March 2019, the International Rescue Committee (IRC) set out to respond to this design challenge:

> How might we provide learning at the right level to children, in any learning space, without relying on skilled teachers, within the first 8 weeks of a crisis, at scale?

Autonomous learning **software**, with thoughtful **facilitation** and agile technology **infrastructure**, has the potential to improve education access and quality in emergency settings. We define autonomous learning as a child-led process in which students engage in learning activities and navigate educational content without the need for a skilled teacher.

In order to pilot and rigorously evaluate autonomous learning (AL) programs in emergency contexts, the IRC is implementing a three-part strategy:

1. Adapt the best AL solutions for children who need it most.
2. Design new delivery strategies that respond to different contextual challenges to ensure that last mile learners have access to AL.
3. Evaluate the cost-effectiveness of AL to ensure that AL solutions are set up for scale from the onset.

This report provides an overview of the IRC's first AL project in the refugee settlement of Cox's Bazar in Bangladesh. The project began with an exploratory design research phase and then moved into a full pilot of the program.

Design Phase

The design research phase aimed to better understand user values and context and to iteratively test the operational feasibility of different software, localization methods and delivery models.

- **User values and context**: How might we create a program that is both answering a humanitarian need and providing value to the community?
- **Software and content:** What tablet-based interactive content can keep children engaged while learning foundational academic and SEL skills at the right level and with autonomy?
- **Localization:** How might we use new workflows to enable timely localization of AL software so that programs can be set up quickly to respond to emergencies?
- **Facilitators:** How might we create a simple delivery model where low-skilled facilitators support children's learning experience?
- **Spaces:** How might we utilize informal spaces in temporary homes and centers while ensuring a high-quality and safe learning experience for children?
- **Technology and hardware set-up, storage and distribution:** What operations and infrastructure are needed to support AL programs in crisis-affected contexts?

Design Insights

- **User values and context:** Parents exhibited high agency with regard to where their children went to receive an education in the camp. They made decisions that prioritized education quality, and considered whether learning spaces showed respect for gender norms, such as boys and girls learning separately and female facilitators working close to home, and exhibited good discipline, order and neatness. They also chose education spaces that protected time outside of school so that children could learn about and practice religion. Community leaders from Cox's Bazar were supportive of the program from the start, but Bangladesh government officials required more in-depth engagement to recognize the program's value and adherence to rules and regulations.
- **Software content:** Software and content was well-received by children and children were easily engaged in the software. Children and their communities valued the program because it taught English and numeracy, but they also expressed that they really wanted to learn Burmese. Visual cues, some activity repetition, and child-led learning where children choose their own activities were found to create the best first time use experience, allow children to concentrate on the learning content rather than instructions, and keep children engaged throughout the course of a learning session.
- **Localization:** Children understood the

software content that was localized with Rohingya voice-over. While doing a full cultural adaptation and localization of software is best practice, it may not be always necessary when responding quickly in a crisis.

- **Facilitators:**
    - Facilitators could be children's mothers, sisters or other close community members, and should receive a stipend, regardless of their designation. We heard from IRC staff and community members that if we created a program with no stipend and relied solely on volunteers, we would have high turnover, little implementation control, and the program would have been less valued by the community. Economic empowerment was key to success and buy-in.
    - Facilitators were quick to learn the application through hands-on experience, but needed human support to understand the logistics of the program, including helping children log in to their accounts, how to receive and distribute tablets, and what to do if something on the tablet broke. Facilitators were eager to help children, and against our advice, often jumped in when students faced challenges or roadblocks, sometimes by solving the activities themselves.
- **Spaces:**
    - Caregivers were strong advocates for the program, but it was not possible for parents to host the learning sessions because it required female facilitators to enter another person's home which was not considered appropriate in this context. It was better when the learning session took place at the facilitator's home. Learning sessions could also take place at child-friendly centers in camp.
    - Initial prototypes in the home showed that tablets could be shared by 3-5 children per day. If facilitators were able to run 3-5 sessions per day, each with 8-10 children, we could reach a ration of 1 facilitator per 24 students to 1 facilitator per 50 students. Prototypes in the center showed that tablets could be shared by up to 20 students and facilitators could run up to 5 sessions per day, but only if space in the center was available.
- **Technology and hardware set-up, storage and distribution:** Operations were challenging and storing tablets in the camps was not a valid option for the pilot because security and risk mitigation was a major concern for the IRC operations team in Cox's Bazar. Partnership with other organizations to store tablets and new risk mitigation strategies and asset protection agreements should be explored in the future.

Pilot Phase

Immediately after the design phase, we launched a 16-week pilot study with 632 students in two different refugee camps in Cox's Bazar. The primary objective was to gather information about the feasibility and desirability of different implementation models of AL, using two different localized software products (Kitkit School and Can't Wait to Learn), and to learn more about the operations required to conduct this work at scale.

The research conducted during the pilot had four areas of inquiry:

Participant outcomes: What baseline-endline changes do we observe in learning (literacy and numeracy) and social emotional learning (hope and agency) outcomes of children using AL?

Implementation fidelity: What levels of attendance, dosage, progress and engagement and overall levels of facilitator proficiency do we observe in the AL sessions?

Participant experiences: What are the experiences of children, facilitators, caregivers, and community members with the Pop-Up program?

Cost efficiency: What was the average cost per child during this pilot and how can it be optimized?

In order to answer these research questions, we conducted a pilot observational study with mixed methods data.

Participants: For the quantitative sample, we collected data from 521 students, between the ages of 5 to 15, who engaged in AL in 71 delivery sites (482 children in 66 home-based sites and 39 children in 5 center-based sites) who were either using CWTL numeracy (264 children) or Kitkit School literacy and numeracy (257 children). At endline, we tried to collect data from the same children and sites, but were unable to track more than 50% of the children due to COVID-19 disruptions. The qualitative sample consisted of 105 participants, which included 24 caregivers, 63 children, 10 facilitators, 4 community leaders and 4 IRC staff members.

Instruments: We implemented a wide set of tools including: the ASER survey for literacy and numeracy, a hope and agency scale, an adapted Teacher Classroom Observation (TCO) tool, protocols for interviews and focus group discussions, and user interface and user experience testing scripts. All data was analyzed by members of the IRC's Airbel Impact Lab.

Analytic approach: We analyzed the quantitative data looking at descriptive indicators such as means and frequencies, and we documented learning gains by identifying baseline-endline changes in the proportion of children at different levels of performance. In order to analyze qualitative data, we identified recurring themes and aligned them with pre-existing categories of analysis, including program experience, perceived education value and quality, and alignment with social and community values.

Results

- **Learning gains in literacy:** Changes in literacy skills are documented only for children who used Kitkit School because children in Can't Wait to Learn did not have access to a literacy software. At baseline, we observed that 71.60% of children were not yet able to read words correctly (ASER Levels 0 and 1), and 28.40% were only able to read words and not sentences (ASER Level 2). After 4 months of using Kitkit School, there was a positive increase of .21 ASER levels on average. Specifically, we observed that the percentage of children who were unable to read words (Levels 0 and 1) decreased 24.06 percentage points from 71.60% to to 47.54%, as they moved to more advanced literacy levels. Additionally, we observed increases to 40.81% of children reading words correctly (Level 2), 4.04% of children reading a short Grade 1 paragraph correctly (Level 3) and 7.62% reading a Grade 2 passage correctly (Level 4).

- **Learning gains in numeracy:** Children using Can't Wait to Learn spent 4 sessions per week and 45-60 minutes engaged in numeracy. Children using Kitkit School spent approximately half that time in numeracy. We observed positive changes of .46 ASER numeracy levels on average for children who used Can't Wait to Learn. At baseline, 80.3% of children using Can't Wait to Learn were not able to identify double-digit numbers (Levels 0 and 1), and this percentage decreased by 46 percentage points to 34.3% at endline, as children moved to Level 2 and successfully identified the numbers 10-99 correctly. We also observed a positive change of .33 ASER numeracy levels on average for children who used Kitkit School . At baseline, 83.3% of children using Kitkit School were not able to identify double-digit numbers (Levels 0 and 1). At endline, 46% of children moved to Level 2 and successfully identified the numbers 10-99 correctly. Additionally, a small percentage (<1%) progressed to ASER Levels 3 (subtraction) and 4 (division).

- **Changes in hope and agency:** We observed positive improvements in children's hope and agency after participating in 4 months of AL. At endline, the percentage of children with low levels of hope and agency decreased by 5 percentage points and the percentage of children reporting high levels increased by 9 percentage points.

- **Implementation fidelity:**

  - **Dosage:** Students learning on Can't Wait to Learn in the home experienced an average dosage of 54 minutes learning on the tablet per session or 90% of the intended dosage. Students learning on Can't Wait to Learn in the center experienced an average dosage of 51 minutes learning on the tablet per session or 113% of the intended dosage. We are unable to accurately report on the dosage for Kitkit School due to usage tracking issues. However, initial analyses of Kitkit School data showed that participants spent (roughly) equivalent time on each subject in Kitkit School. This implies that the dosage for math for Kikit School would be about a half of the total dosage and less than half of the numeracy dosage received by children learning on Can't Wait to Learn (because it was math-only).

- **Facilitators' proficiency** was assessed with regard to behaviors such as providing positive encouragement, being attentive to the needs of children, playing games and practicing calming activities, using positive discipline strategies to respond to misbehavior, and helping children find their own solutions. Data from two classroom observation sessions indicate that 24% of the facilitators were rated as exhibiting emerging evidence of proficiency, and approximately 76% were rated as exhibiting either good or exemplary levels of proficiency.

- **Participant experiences with AL:** In qualitative interviews, children, caregivers, facilitators and community members described Pop-Up as a valuable program that brought quality education to the community. They reported appreciating the home-based model for its ability to i) enable small-group learning that is visible by caregivers, ii) enable women and girls to stay close to home and iii) enable quiet and clean learning environments, despite the small spaces of each home.

- **Cost efficiency:** The cost per child including both Kitkit School and Can't Wait to Learn was $575 per child for the initial pilot study and early-stage investments, including software localization, equipment, infrastructure, and administrative costs. The cost per child by software was $723 per child for Kitkit School (numeracy and literacy) and $412 per child for Can't Wait to Learn (numeracy). Localization of the software comprised 51% of the total budget (40% Kitkit School; 11% Can't Wait to Learn). The cost per child excluding localization was **$284 per child**. In this model, equipment, including the hardware (tablets and other tech set-up), made up the majority of the budget at 36%, followed by national staff (32%) and country operations and support costs (16%). Both localization and equipment are a one-time cost incurred when launching the program, and tablets can be used for multiple years across multiple cohorts of children thus improving the cost per child over time. Costs of a new pilot program differ from those of an established IRC program and we expect to see cost efficiencies once the program is tested and established. Initial scale projections estimate a **cost per child of $151** when reaching 32,000 children excluding localization costs and HQ costs which will not be incurred during the scale phase of a project.

Limitations

The pilot study aimed to collect evidence about the feasibility and desirability of implementing AL using two different localized software products (Kitkit School and Can't Wait to Learn), and to learn more about the operations required to conduct this work at scale. This study did not attempt to identify the impact of different AL software or delivery modes on children's learning and SEL outcomes. We strongly discourage the reader from attempting to draw conclusions on impact because we did not use a control group. Additionally, we discourage the reader from attempting to compare software packages or delivery models based on the baseline-endline findings that we included in the pilot because the study samples were very small and not powered to conduct comparisons, and because COVID-19 disrupted data collection activities at endline, further reducing our sample sizes.

Conclusions and recommendations

The results of the Pop-Up Learning pilot study in Bangladesh confirm that AL software is a promising avenue for alternative education in crisis settings. However, it requires further investment in nimble infrastructure, software features and rigorous research to ensure the model can be deployed quickly at the onset of a crisis, in a cost-effective manner. The pilot study helped us to obtain preliminary evidence to support hypotheses with regard to different dimensions of the program:

- Displaced children who are out of school can acquire foundational academic skills and SEL skills through tablet-based AL.
- Displaced out-of-school children are able to attend home- and center-based learning sites, in a safe and stable way.
- Tablet-based interactive content can keep children engaged while learning foundational academic and SELskills.
- Low-skilled caregivers or community members can be recruited to work as facilitators with a stipend through community touchpoints, lightly supported by the IRC, and are able to supervise AL sessions successfully.
- Localization into niche dialects and languages can be faster and cheaper than current practice; new workflows can enable timely localization.

To inform the next phase of work, we lay out 15 recommendations for the IRC and its partners. We elaborate on these recommendations in the full report below.

For program implementation:

1. Store tablets near learning spaces.
2. Invest in charging systems that are adapted to local constraints.
3. Increase facilitator model reach.
4. Clarify facilitator roles and positioning in the community.
5. Test the value of workbooks for blended learning models.
6. Structure sessions to include more SEL activities and stretching breaks.

For software and product design:

7. Ensure children are guided through the curriculum.
8. Improve First Time Use (FTU) experience and optimize for the absence of training.
9. Include ways for children to ask for help in the software.
10. Propose different ways of learning a concept in the software.
11. Integrate more adaptive features to tailor content to children's learning levels.

For research:

12. Improve tablet analytics and work towards standardized metrics.
13. Use valid and reliable tools of learning that are fit for purpose to measure learning gains.
14. Conduct further operational research to test feasible ways to store and distribute tablets and to deliver the program safely during the COVID-19 crisis.
15. Build the evidence about the impact of AL on children's learning and SEL outcomes and about cost-effective ways to implement AL at scale.


INTERNATIONAL
RESCUE
COMMITTEE
pop up
President

1.0 Autonomous learning in crisis settings: Our vision and global strategy

1.1 The Learning Crisis

Millions of children in low- and middle-income countries are failed by the education system and are not reaching their development potential. The "learning crisis"[1] faced today is one of access and quality.

ACCESS:

Despite the progress on increasing access to schools made over the last decade, more than 262 million children are not in school.[2] That is 1 out of 5 in the world. And in crisis and conflict settings, 10 million primary school-aged children are displaced, and 40% of refugee children are out of school.[3]

QUALITY:

When in schools, children are also not learning at their developmental potential. Six out of 10 children are not able to achieve proficiency levels in reading and mathematics by the time they complete primary school. In 2017, that is more than half the children in the world. After several years of schooling, many still cannot read, write, or do basic mathematics, which has dramatic consequences on children's life outcomes—risking unemployment, poor physical and mental health, and early child marriage later in life.[4]

The need for innovative educational products and implementation methods is crucial. We must find alternative ways to enable children in crisis settings to have access to quality education, to reach their full developmental potential.

At the International Rescue Committee (IRC), we are developing and testing innovative solutions that improve access to quality education for refugees, displaced people, and underserved populations around the world. In March 2019, we set out to respond to this design challenge:

> How might we provide learning at the right level to children, in any learning space, without relying on skilled teachers, within the first 8 weeks of a crisis, at scale?

1.2 The solution: Autonomous learning adapted to crisis settings

We believe that evidence-based autonomous learning software with thoughtful facilitation and agile technology infrastructure can improve education access and quality in emergency settings.

What is Autonomous Learning?

We define autonomous learning (AL) as a child-led process in which students are directing their own learning and navigating educational content without the need for a skilled teacher. Many companies are currently creating and offering AL software, using different features to create an engaging and interactive user experience for children, such as gamification techniques, adaptive algorithms, use of videos and quizzes, and so on.

How does AL work?

By leveraging a high-quality, interactive, and guided curriculum on a tablet, students can take charge of their own learning without relying on the physical infrastructure of classrooms or internet connection. AL software provides a practical solution for children when there are few skilled teachers available. The software usually works offline and synchronizes data on a mobile server to enable parents or teachers to monitor children's progress. These digital learning experiences enable children to rediscover the joy of learning. Children also gain agency as they learn at their own pace, with their peers and with high-quality, learning resources.

Evidence-based autonomous learning programs have the potential to be transformational in helping to solve the global learning crisis. Existing evidence around autonomous learning is overall promising and indicates that such programs might be a cost-effective solution in the short to medium term for providing education in challenging environments.[5]

1. World Bank. (2018). Word development report 2018: Learning to realize education's promise. Washington, DC: World Bank.
2. UNESCO. (2017). More than one-half of children and adolescents are not learning worldwide. Fact Sheet No. 46. UNESCO.
3. Galen, H. (2018). Estimating the number of forcibly displaced school-age children not accessing education. Background paper prepared for the 2019 Global Education Monitoring Report, Migration, displacement and education: building bridges, not walls.

   UNICEF. (2020). Child displacement. UNICEF.
4. UNESCO. (2017). More than one-half of children and adolescents are not learning worldwide. Fact Sheet No. 46. UNESCO.
5. Muralidharan, Karthik, Abhijeet Singh, and Alejandro J. Ganimian. 2019. "Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India." American Economic Review, 109 (4): 1426-60.

   Research by Nicola Pitchford exploring the use of innovative mobile technology to support the acquisition of basic skills (numeracy, literacy, English) by primary school children in Malawi, the UK, and several other countries.

   Levesque, Karen, Sarah Bardack, Antonie Chigeda. 2020. "Tablet-based Learning for Foundational Literacy and Math: An 8-month RCT in Malawi." imagineworldwide.org

The IRC has been exploring the use of various AL software in the past few years, to enable vulnerable populations to learn without skilled trainers or teachers in very remote contexts. Some of our existing partners to cite a few are: Kitkit School, Can't Wait to Learn, Onebillion and Mindspark.

However, few evaluations of AL have been conducted in displacement and crisis settings. Additionally, the current ecosystem of partners and projects does not yet allow us to understand precisely whether AL works to improve the learning and transition outcomes of out-of-school (OOS) children during an emergency setting. As such, the IRC and partners aim to cultivate a portfolio of design and research projects to answer a global research agenda:

1. What **operational systems and human resources** are required to quickly and effectively deploy Autonomous Learning programs in last-mile, crisis-affected settings?
2. What **delivery models** are **desirable, feasible and scalable** when reaching the most vulnerable, last-mile learners in crisis-affected settings?
3. What **behavioral supports** are useful in improving uptake and adherence for last-mile learners in crisis-affected settings?
4. What is the **cost** of localizing Autonomous Learning materials and implementing different delivery models in crisis-affected settings?
5. What is the **impact** of Autonomous Learning on children's **literacy, numeracy, social-emotional learning (SEL) and transition outcomes?**
6. Is Autonomous Learning a **cost-effective** replacement solution to provide access to education opportunities and improve the learning and SEL outcomes of out-of-school children in crisis affected settings? Is AL a **cost-effective** complementary solution to improve learning, SEL, and transition outcomes for children in crisis contexts?

What is the IRC's strategy on AL?

With this global agenda in mind, we have developed a three-part strategy:

1. **Adapt the best AL solutions for children who need it most.** This requires working hand-in hand with software partners to localize and adapt their products for vulnerable populations according to various contexts and cultures.
2. **Design new delivery strategies that respond to different contextual challenges to ensure that last mile learners have access to AL.** This requires providing a flexible infrastructure, for example solar power, charging stations, mobile servers, offline and intermittent connectivity; and the right level of human support through volunteers, caregivers, and teachers to support children in their learning experience based on the needs. This will enable deployment in the most challenging contexts, while ensuring that children receive the intended dosage for the lowest cost possible.
3. **Evaluate the cost-effectiveness of these programs to ensure that AL solutions are set up for scale from the onset.**

For our first autonomous learning pilot study in Bangladesh, the IRC has partnered with Enuma (Kitkit School) and War Child Holland (Can't Wait to Learn).

As both software programs present unique approaches and strengths, it is interesting to pilot both of them to better understand how different solutions can be delivered and adapted to various implementation models.

We underwent an extensive design phase where multiple software programs were field-tested with Rohingya children and with adults over the course of 3 months in the refugee settlements in Cox's Bazar, Bangladesh to understand which would work best in the given context. In the final selection of the software, a number of criteria were considered including relevant curriculum for the target population, interactive methods, adaptive methods, ability for children to use it without adults, works offline in the current context and interest to partner from the software companies.

Can't Wait to Learn	Kitkit School

▪ A structured numeracy curriculum[6] that covers Pre-K to Grade 3 with 73 different mini-games and 10 characters in the game world.
▪ Game design based on simulation of building a village through an animated map.
▪ Contains mini-games, videos, and highly localized story-telling with various characters.
▪ Winner of the UNESCO King Hamad Bin Isa Al-Khalifa ICT in Education Prize (February 2019).
▪ Winner of the Dutch Coalition for Humanitarian Innovation (DCHI) public prize for the best humanitarian innovation 2018.
	▪ Kitkit School is a comprehensive digital learning solution with 3 main components:
	  ▪ A tightly-scaffolded curriculum that covers Pre-K to Grade 3 in literacy (11 courses, 280 sessions, and 1,200 activities) and numeracy (11 courses, 260 sessions, 1,200 learning activities) with embedded quizzes and mini-assessments.
	  ▪ Built-in library with hundreds of books and learning videos.
	  ▪ A suite of tools to support creative self-expression through music and art.
	▪ Game design based on a nature theme, featuring a literacy and math coop with eggs and creatures that children hatch and grow as they learn.
	▪ Grounded in Universal Design for Learning, with an intuitive child-friendly interface.
	▪ Co-winner of the Global Learning XPRIZE competition (May 2019).

6. While Can't Wait to Learn has a literacy curriculum in English, Arabic, and French (in progress), they did not have English as an Additional Language (EAL), which is required for this group. Children enrolled in the pilot study only had access to the numeracy curriculum.

2.0 Pop-Up in Bangladesh:

Program design and geographical context

2.1

Working in Cox's Bazar in Bangladesh:

Planting the seed for global programming

Why Bangladesh?

Bangladesh is host to over 855,000 Rohingya refugees, nearly 400,000 of whom are school-aged children. The vast majority of Rohingya refugees live within 34 refugee camps crowded together in the country's Cox's Bazar district. From a global perspective, Bangladesh is representative of the typical challenges faced in many humanitarian contexts, making it a strategic place to begin piloting and testing our autonomous learning work. These challenges include:

- **Limited space in the camps.** The majority of refugee camps in Cox's Bazar, Bangladesh do not meet the minimum density requirements set out by UNHCR.[7] Lack of space severely hampers the construction of learning centers; where there is space, priority is given to lifesaving service facilities.
- **Mixed learning levels.** Due to the large numbers of children who have arrived, the focus has generally been on providing immediate access rather than making distinctions between different educational levels among children.
- **Difficulties in recruiting qualified teachers.** This is particularly challenging from the Rohingya population because of the lack of enrollment in formal school and teaching experience for many Rohingya adults prior to displacement.
- **Complex language issues and lack of schooling experience.** Children have been displaced from Myanmar where formal education is done in Burmese, but the great majority have not been enrolled in school before their displacement. They also speak Rohingya, a dialect that does not have a written script. Moreover, the Bangladesh government prohibits Rohingya children from learning its national language, Bangla. All of these parameters make design decisions regarding language of instruction extremely difficult.

While progress has been made on access to informal education for children since the IRC began responding to the Rohingya crisis in Bangladesh two years ago, the quality of educational services available in the camps remains of great concern. An unpublished assessment conducted by 25 education sector partners and shared at the December 2019 education sector meeting showed that among children enrolled in temporary learning centers (TLC), learning progression was poor: after one year of educational programming, only 37.6% of children were promoted from Level I to Level II, and only 49.2% from Level II to Level III.[8] The 2019 Education Needs Assessment by REACH showed that only 2% of TLCs fully meet minimum sector-specific teaching and learning materials requirements, which means that creating a joyful and effective learning environment in the camps continues to be a big challenge. The Needs Assessment also found competency and skills of teachers was a concern for all stakeholders: 52% of caregivers reported that they would like to see improvements to teaching at the learning centers and the most commonly-reported request for improvement was more training for teachers (reported by 40% of caregivers). The lack of adequate learning environments and skilled teachers impacted children's motivation to join the learning sessions: 22% of out-of-school boys and 16% of out-of-school girls were reportedly not attending sessions because what is taught in TLCs is not useful or age-appropriate (6-14 years).[9]

7. Inter Sector Cooperation Group (2019). Cox's Bazar: Rohingya Population Density by Camp in Ukhia. Inter Sector Cooperation Group.

8. As per Learning Competency Framework and Approach (LCFA): Level I: Equivalency: Pre-Primary, Level II: Equivalent to Grade I and II Competencies, Level III: Equivalent to Grade III, IV and V Competencies Level IV: Equivalency to Grade 6, 7, 8

9. UNICEF & REACH. (2019). Education needs assessment: Cox's Bazar, Bangladesh. UNICEF.

2.2 Design research phase

How we created Pop-Up

In the design of Pop-Up, we aimed to develop a product that could address these complex challenges and provide quality educational programming to Rohingya children across Cox's Bazar. In April 2019 we launched into the design research phase of Pop-Up, spending three months developing a better understanding of user values and context and iteratively testing different software, localization methods, and delivery models to determine which combination would enable high-quality learning within the first few weeks of a displacement crisis. Our design research focused on six main areas of inquiry:

USER VALUES AND CONTEXT

How might we create a program that is both answering a humanitarian need and providing value to the community?

SOFTWARE AND CONTENT

What tablet-based interactive content can keep children engaged while learning foundational academic and SEL skills at the right level and with autonomy?

LOCALIZATION

How might we use new workflows to enable timely localization of AL software so that programs can be set up quickly to respond to emergencies?

FACILITATORS

How might we create a simple delivery model where low-skilled facilitators support children's learning experience?

SPACES

How might we utilize informal spaces in temporary homes and centers while ensuring a high-quality and safe learning experience for children?

TECHNOLOGY AND HARDWARE SET-UP, STORAGE, AND DISTRIBUTION

What operations and infrastructure are needed to support AL programs in crisis-affected contexts?

The design research findings summarized below are based on the analysis of field notes, notes from interviews, and observational data documented during program design iterations.

User values and context

How might we create a program that is both answering a humanitarian need and providing value to the community?

The majority of participants we spoke to indicated that education was valued in the community. However, most caregivers expressed frustration with the quality of TLCs in the camps and so did not always require children to attend available educational programs. Stakeholders working in the humanitarian sector reflected on the shortcomings of educational opportunities in the camps and noted they were looking to identify new solutions.

Participants indicated that parents had the ultimate decision regarding where their children went to receive an education in the camp and whether they attended the sessions. We found that the factors most important to parents when assessing education opportunities included the perceived quality of education programming, respect for gender norms, a learning environment that was disciplined, orderly and neat, and the protection of time to learn about and practice religion.

Community leaders we spoke with were eager to see more education opportunities offered to their communities and were supportive when approached to share more about the program.

One of the major obstacles to implementing Pop-Up was the regulation mandated by the Bangladesh government of no mobile phones or devices in the camp. Not surprisingly, Bangladesh government officials in charge of the camp expressed hesitation in having children learn on tablets in the camp. It took over three months of in-depth engagement with government officials to gain permission to run the program. The IRC was required to show that tablets did not have connectivity and that children, facilitators and parents could not access any other learning applications on the tablet.

Software and content

What tablet-based interactive content can keep children engaged while learning foundation academic and SEL skills at the right level and with autonomy?

There is a wealth of tablet-based education content, but not all of it is high quality and even fewer educational software applications have been rigorously evaluated to determine their impact on learning.

As a first step towards developing an AL program for Rohingya refugee children in Bangladesh, we conducted background desk research to learn more about the tablet-based solutions in the market that provide adaptive and autonomous learning for primary school children. We identified four software applications with some evidence indicating effectiveness and tested these in Bangladesh with Rohingya children to determine which applications seemed more adaptable to the context and more interesting to the users.

Through this process we learned that children and their communities would value a program that teaches English and numeracy, but that they would also like to learn Burmese. We also found applications with more visual cues improved the first time use experience for children. Some activity repetition was also valuable because it allowed children to master the instructions for a game and concentrate on learning the content. For solutions where the user was asked to frequently switch between activities, children spent more time learning the instructions for an activity than learning the content. Child-led learning where children choose the activities and games they wanted to engage with was also perceived by children as more fun and tended to keep children engaged learning on the application for longer.

We also found that even when using fully designed software applications, there were many decisions to be made about how to program the application and game for a specific population. For example, as there is little available data regarding literacy and numeracy levels, learning pace, and technology literacy for Rohingya children, we had to rely on benchmarks from other similar countries and contexts as a starting point. During the design phase, we tested these assumptions, using our initial findings to inform application programming for the pilot. We found that the availability of a dashboard as part of a software application was particularly important as the component would allow IRC staff members to track children's activity and progress in real-time and when needed, work with a given facilitator to provide targeted support to a child falling behind.

Overall, the design research reinforced our awareness of the need for more evidence-based software adapted for use in humanitarian contexts. Although there are few existing models, we did find that the software in the market is malleable and open to feedback from implementing organizations like the IRC regarding how to best create solutions for these complex settings. Going forward, we believe the IRC has a key role to play in giving feedback to software partners so that solutions can be responsive to the needs of our clients across the many contexts where we work.

Localization

How might we use new workflows to enable timely localization of AL software so that programs can be set up quickly to respond to emergencies?

Localization of tablet-based education content into niche dialects and languages is a large bottleneck in reaching more children with appropriate content in their native language during a crisis. High price tags and long timelines prevent more content from being deployed quickly in an emergency.

During the design phase, we tested the understandability of educational content using a Rohingya voice-over created by experienced translators who had worked with the Rohingya population in Cox's Bazar for several years. Through many tests in the camps with children, parents, and teachers on both literacy and math content, we learned that children would be able to understand the applications and games if they were accompanied by a Rohingya voice-over for instructions and explanations.

The localization process included writing translations (using Bangla script), recording audio for game strings and video dubbing, editing audio files to match videos, testing audio and video in the field to assess comprehension, making edits to the written script, recording the final audio, and making final edits to the audio files to match the videos. Our software partners were responsible for packaging all the final content into a localized application.

In order to localize more quickly, we asked software companies whose applications we were testing to create packages of content to be localized all at once. Packages might include, for example, a string of game instructions and problems that needed to be translated and recorded, as well as instructional videos that needed to be dubbed. We built an at-home recording studio and rented a quiet hotel room to do the recording in. The set-up we prototyped and eventually used was simple and cheap and could most likely be quickly replicated anywhere at the onset of a crisis.

Throughout the process we tested our content with the users and stakeholders to ensure we were developing quality educational content that could be understood by our clients. Pop-Up translators met with teachers from TLCs, parents, children, and community leaders to test their translated audio recording and get feedback. This improved content quality and understandability. Our team also tested outsourcing the content localization to Rohingya diaspora. We learned that working with Rohingya diaspora or other out-of-country language speakers does not always result in high quality localization due mostly to disparities in dialects. We recommend finding staff that have worked closely with the target population in the past and leaning heavily on field testing to ensure understandability.

However, in-house localization did not always make the most efficient use of our translators' skills. We determined that in the future it may be possible to outsource some parts of the process, particularly technical work like cleaning and editing audio files, without compromising quality.

Finally, doing a full cultural adaptation and localization of software is always best practice, but may not always be essential. We found that videos produced in Uganda with Uganda actors were suited to this population when localized with a Rohingya voice-over. Being more flexible with the process of cultural adaptation allowed us to focus our main efforts on the localization of the software using the Rohingya language. This process helped ensure children quickly received content they could understand and learn from.

Facilitators

How might we create a simple model where low-skilled facilitators support children's learning experience?

During the design phase we explored multiple means of recruiting facilitators, including through the formal IRC hiring processes where jobs were posted on camp bulletins and by engaging with community religious leaders to help identify and reach individuals who fit a specific set of criteria.

Initially we opened the position to any individual with or without an education to test our working hypothesis that anyone, even those without a formal education, can guide children in their learning in an autonomous learning setting. However, we received push back from the community who found it more appropriate for facilitators to have at least a minimal level of education.

We also worked through different possibilities of who we might recruit as facilitators. At first we explored recruiting men and women, but gender norms in the community prohibited having men and women in the same training sessions and facilitator meetings. We also found that men were not usually home all day, while women spent most of their time close to home and were familiar and trusted by neighborhood children. Community members we spoke with believed it was most appropriate to have married women as facilitators, a recommendation we followed in all but in a few instances when we had younger, unmarried women (aged 18 - 20) as facilitators. These younger women most often lived in smaller blocks in camp with a lot of family members participating in the program.

We learned through interviews with community members and stakeholders that facilitators could be children's mothers, sisters or other close community members, but they would need to receive a stipend, regardless of their designation. Various individuals noted that if we created a program with no stipend that relied solely on volunteers, we would have high turnover, little implementation control, and it would be less valued by the community. Economic empowerment was key.

We also learned that facilitator familiarity was important and facilitators were intrinsically motivated when their child, their niece, nephew, or cousin or their neighbors' children were enrolled in their group. A sense of pride and leadership in the community helped keep facilitators accountable to their students and the program. By building on existing social networks and community structures, our program was perceived as high quality from the start.

During the design phase we prototyped and tested several different facilitator support models including an introduction and training session and a Community of Practice (COP) meeting meant to support facilitators throughout the program. We observed that facilitators quickly gained confidence with the tablet based on a short in-person demo and explanation of the application. The most valuable way to gain confidence in the application was for facilitators to play the games themselves and work together if they got stuck. In-person discussion and quizzing of the facilitators about the app was appropriate in the culture and helped build confidence. However, in future projects, it would be valuable to include a video orientation to the tablet to ensure that the introduction to the application was standardized and nothing was missed.

In-person, human support from IRC staff was required to explain the logistics of the program including how to help children log in to their accounts, how to receive and distribute tablets, and what to do if something on the tablet breaks. We found follow-on training through COP meetings was best done every other week. Project logistics and technical were top of the agenda at these meetings, but facilitators also benefited from sharing about their challenges and experiences with one another.

During the design phase we also aimed to better understand what type of support was needed from the facilitator to ensure children engaged with the content. We learned that support is needed to help children log in to their account, but engagement with the software came naturally and quickly for children. Once children were given permission, they would begin to play the game without prompting.

Our research also demonstrated that while the software games were designed to be autonomous, modifications still need to be made to improve the child user experience. During sessions children at times would appear confused, stuck or bored with the content. Facilitators were expected to track children's engagement in order to help move them past these more difficult moments. In order to properly track children the facilitators needed to have a basic understanding of how to identify when a child was off track, how to diagnose what was making them go off track, and have mastered several methods of redirecting the child. This type of support did not require mastery of the content itself, but it did require facilitators to have the ability to recognize how children should be learning on the software and if their experience was deviating from that.

We also observed that facilitators had a hard time not jumping in to teach a concept themselves if students faced challenges or roadblocks, and were often eager to help children, sometimes by solving the activities themselves.

Spaces

How might we utilize informal spaces in temporary homes and centers while ensuring a high-quality and safe experience for children?

We aimed to build on existing behaviors in the community when seeking space to host Pop-Up Learning sessions. In this particular context, the program was operated within a conservative Muslim culture with tight knit communities, even after displacement. Families often resided in the same home or a single block of homes nearby to one another. Women and girls were encouraged to stay close to home.

Initially, we prototyped facilitators moving house to house for three sessions over the course of the day, serving children from their block at other families' homes. This prototype failed. Female facilitators and their families felt uncomfortable entering someone else's home and it was hard to find enough homes that could accommodate a group of 8 students without disrupting family activities. In response, we tested delivering three sessions per day in the facilitators home. Children were asked to walk a short distance, within the same block of camp. We found this did not cause a disruption and because facilitators were paid a stipend and this work was considered a high-quality job, facilitators' husbands and family members were happy to leave the room where the learning sessions were taking place.

Initial prototypes in the home showed that tablets could be shared by 3-5 children per day. If facilitators were able to run 3-5 sessions per day, each with 8-10 children, we could reach a ration of 1 facilitator per 24 students to 1 facilitator per 50 students. Prototypes in the center showed that tablets could be shared by up to 20 students and facilitators could run up to 5 sessions per day, but only if space in the center was available. Ultimately, the ratio would improve depending on the number of sessions facilitators could teach in one day and the number of students per group. We found we could add one more session per day by shortening the length of the sessions to one hour maximum and targeting girls for afternoon sessions who were about 10 years and older and not required to attend afternoon religious studies at the madrassa.

We found community relationships that enabled parents to visit their children's learning sessions throughout the day mitigated protection risks for children visiting facilitator's homes. Additionally, the close relationship between parents in the community and the facilitator meant the facilitator was often a relative or close friend of the parents who had children in the group. Social accountability played a large part in the success of the program and mitigating protection risks.

Outside of the home we also prototyped hosting sessions in child-friendly centers operated by the IRC. While this was an adequate test of a center-based model, if we were to scale, we would likely have to build our own centers designated for Pop-Up because competition for time in the shared child-friendly centers was high and we were not guaranteed adequate time to host our groups. Additionally, due to their location outside of residential blocks in camp, we saw that learning sessions in the centers were not able to capitalize on community relationships as much as learning sessions in the home.

Technology and hardware set-up, storage, and distribution

What operations and infrastructure are needed to support AL programs in crisis-affected contexts?

During the design phase, our team investigated multiple potential storage solutions for the tablets. Options included storing tablets in a locked cabinet in the Bangladesh government office inside of camp, in permanent or temporary IRC buildings in camp, or in permanent or temporary buildings belonging to other humanitarian organizations. Ultimately we had to look outside of the camp for storage solutions due to a lack of available buildings with permanent walls and secure locks and hesitation from the Bangladesh government office to provide storage assistance to the IRC and not to other organizations. Finding an adequate storage site was made even more difficult by the fact that camps were so sprawling in Cox's Bazar and traffic was so dense that it can take up to 3 hours to travel in between camps, not including travel back to the main city of Cox's Bazar and the IRC's main office.

We invested less time identifying electricity solutions to charge the tablets. Although solar power is a viable option in this context, we did not pursue it for this pilot because regular electricity was available at approved and secure tablet storage locations in camp.

One major challenge we faced in developing the Pop-Up program was finding ways to connect tablets to the internet. The Bangladesh government had banned the use of mobile phones and other devices in camp and had requested telecommunications companies to block connectivity within the camp. Due to these regulations we were only able to back up user data from the tablets using an internet connection outside of the camp.

The distribution models we prototyped varied depending on where the tablets were stored. One option included IRC staff transporting tablets from the storage site by car and handing them off to facilitators at the camp entrance road. In another camp where the pathways to the residential areas were winding and steep, IRC staff walked the tablets from the camp entrance road into the residential areas of camp and met 2-4 facilitators at designated touch points. Each facilitator and IRC staff member had a waterproof backpack that allowed for easy and safe carrying of tablets.

Risk mitigation requirements by the IRC team in Bangladesh meant some potential solutions were out of reach. Tablets had to be taken to a locked cabinet, in a secure building which meant storing tablets in the camp was not an option. An IRC health center and field office, both outside of camp, were identified as the best potential hardware storage solution given the safety and infrastructure constraints. Both locations were equipped with electricity for charging the devices and internet connection to back up user data from the tablets. Although the IRC team was able to store the tablets in a secure site with electricity and internet connection, the tablets remained far from camp and required distribution by car and foot each day.

A longer inception period and more concentrated efforts to partner with other organizations with permanent structures inside of camp could still result in finding a storage location in camp and should be explored in further iterations of the project. Additionally, new risk mitigation strategies and asset protection agreements should be explored in future iterations of the project in order to allow for more agile technology infrastructure and distribution models.

2.3 Program components

After three months of prototyping, testing, and learning about the Bangladesh context, we created Pop-Up Learning with the following components:

Software	▪ We piloted two software programs together, Kitkit School and Can't Wait to Learn, to understand if certain features lead to different usage and experiences for children and facilitators. ▪ Kitkit School provided children with numeracy and literacy (English as a Second Language), curriculum from Pre-K to Grade 3 with Rohingya audio/video instructions . ▪ Can't Wait to Learn provided children with a numeracy curriculum from Pre-K to Grade 3 with Rohingya audio/video instructions.
Facilitators	▪ The facilitators were low-skilled female caregivers from the community, including mothers and older sisters, who received a stipend to work as facilitators and who were recruited for their motivation and interest in running learning sessions. ▪ All facilitators had received some education themselves ranging from some primary school through to matriculation into university. The average level reached was Level 6 and only 2 out of 31 achieved matriculation into university. ▪ Their role was one of implementation and they were not asked to provide any teaching to children. ▪ They received a kickoff training of 2 days, followed by a biweekly peer-to-peer learning circle or community of practice (COP) with the Pop-Up team. COP meetings were designed to provide ongoing support to facilitators and to help IRC staff understand what was working well and what aspects of the program needed to be improved. ▪ Facilitators carried 8-10 tablets in their Pop-Up backpacks and ran several learning sessions per day with groups of 8 children.
Infrastructure and ops	▪ Tablets were stored in the IRC office in Ukhia and a dedicated logistics team brought them in and out of camps every day. Depending on traffic between the IRC field office and the camp, transporting tablets could take anywhere between 1 and 3 hours to transport one-way. ▪ Tablets were charged through regular electricity. ▪ Tablets synced with the server every time they were returned to their charging location.
Spaces	▪ The program took place in facilitators' homes and in various informal learning spaces available in the camps. ▪ Two camps were selected for the pilot, Camp 8E and Camp 22. Camp 8E is a large and dense camp centrally located in the refugee settlement. Camp 22 is farther away from Cox's Bazar and the main IRC field office and is smaller and less dense than Camp 8E. We chose these two locations to test out program feasibility in a variety of camp contexts in Cox's Bazar.
Dosage	▪ Children enrolled in sessions with a minimum of 45-60 minutes spent learning on the tablet per session. All home groups and one center group aimed for children to receive 60 minutes of learning on the tablet. Two center groups aimed for children to receive 45 minutes of learning on the tablet. Learning was limited to 45 minutes in two center groups because other programming separate from Pop-Up was already taking place in the center. Remaining session time was dedicated to SEL,[10] session introduction, and wrap-up activities. ▪ Dosage for Can't Wait to Learn was spent exclusively on mathematics while dosage for Kitkit School was split between mathematics, literacy, and tools to support creative self-expression through arts and music. Given the inherent differences in the two software products, it is not possible to directly compare associated learning outcomes.

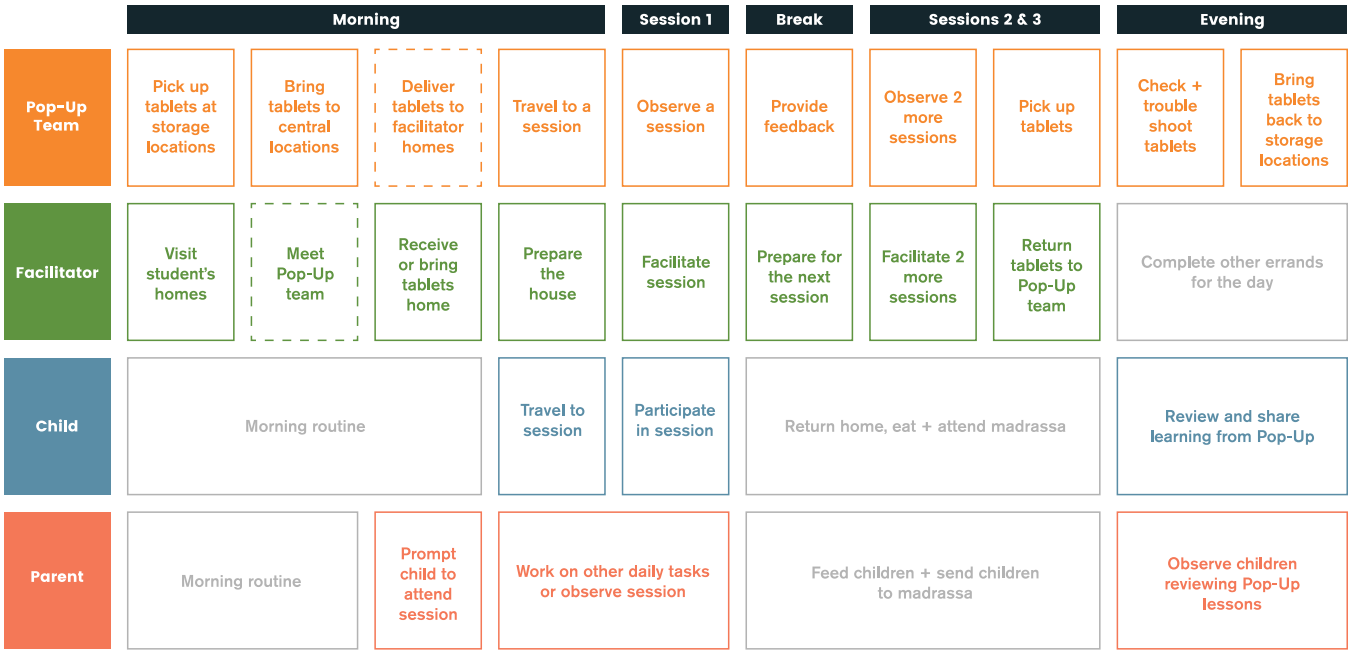
We rolled out our pilot with these components in mind and adjusted the programming as needed as we learned about feasibility. The following illustrated journeys show the weekly and daily programming of Pop-Up Learning, as well as what a typical learning session looked like.

What's happening: The weekly Pop-Up routine

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Pop-Up Team	Deliver + collect tablets; support facillitators	Deliver + collect tablets; support facillitators	Deliver + collect tablets; support facillitators	Deliver + collect tablets; support facillitators	Run CoP meetings (bi-weekly)		
Facilitator	Run Pop-Up sessions	Run Pop-Up sessions	Run Pop-Up sessions	Run Pop-Up sessions	Participate in CoP meetings (bi-weekly)		Visit homes + check on Pop-Up experience
Child	Participate in Pop-Up session	Participate in Pop-Up session	Participate in Pop-Up session	Participate in Pop-Up session			Talk to facilitator about Pop-Up
Parent	Send child to Pop-Up session	Send child to Pop-Up session	Send child to Pop-Up session	Send child to Pop-Up session			Talk to facilitator about Pop-Up

10. Initially the program was designed to include SEL activities from IRC's Safe Healing and Learning Classrooms curriculum. However, given implementation complexities, we were unable to train facilitators to deliver these activities. In practice, SEL activities were limited to one activity that facilitators were taught during their training. The inclusion of additional activities and games was decided by each facilitator individually or the IRC field staff, but these activities were not SEL-specific or taken from IRC's evidence-based SEL curriculum.

What's happening: The daily Pop-Up routine



What's happening: The Pop-Up session

		5 min		5 min	45–60 min	5 min	5 min	5 min
Child	Wait for session to start, sitting quietly	Play train game	Sit down	Receive tablet + headset	Play on the tablet	Return tablet + headset	Share today's learnings	Play train game until exiting the room
Parent	Wait for students to arrive	Instruct students to play train game, sing, recite poem	Seat students in designate	Hand out tablets + headsets	Observe + support students	Collect tablets + headsets	Ask students about learning	Instruct students to play train game, sing, recite poem

3.0

Pilot research phase

How we tested Pop-Up

After a phase of exploratory design research and product localization, we launched a 16-week pilot study with 632 students in Camp 8E and Camp 22 refugee camps in Cox's Bazar. Our primary objective was to gather information about the feasibility and desirability of different implementation models of AL, using two different localized software products, and to learn more about the operations required to conduct this work at scale.

Specifically, the pilot study aimed to answer the following research questions:

1. PARTICIPANT OUTCOMES

What baseline-endline changes do we observe in learning (literacy and numeracy) and SEL (hope and agency) outcomes of children using AL?

- What are the baseline-endlines changes in participant outcomes for each delivery model (home-based and center-based)?
- What are the baseline-endline changes in participant outcomes for each software?

2. IMPLEMENTATION FIDELITY

What levels of attendance, dosage, progress and engagement and overall levels of facilitator proficiency do we observe in the AL sessions?

- What levels of attendance do we observe?
- What amount of time do children spend engaged with different software programs? [11]
- How far do children progress through the game? What learning pace, time on task, and success rates do we observe among participant children? [12]
- What overall levels of implementation quality do we observe in the AL sessions? Do we observe any indicative trends in the quality of the learning environment provided by facilitators in home-based and center-based delivery models? What levels of implementation quality did we observe for each software program?

3. PARTICIPANT EXPERIENCES

What are the experiences of children, facilitators, caregivers, and community members with the Pop-Up program?

- What are the perceptions of children, facilitators, caregivers, and community members about i) the relevance and usefulness of the program and ii) the quality of facilitation and the learning environments?
- What are facilitators' perceptions about the benefits and challenges of participating in the COP?
- What aspects of the learning experience do children like and dislike?
- What aspects of the learning experience are easy and difficult for children to navigate?
- What are children's perceptions and experiences with different software programs? Which features of each software program do children prefer? [13]

4. COST EFFICIENCY

What is the average cost per child experienced during this pilot and how can it be optimized?

- What are the key cost drivers per child?

11, 12, 13. This study is not powered to compare softwares and inherent differences in the two software products do not allow a direct comparison.


RESCUE
pop up
RESCUE
pop up

4.0 Method

We conducted a pilot study with the aim of gathering information about changes observed in the learning and SEL outcomes of children, and the experiences of stakeholders with different components of the program. The study included pre and post quantitative data collected from students through surveys and learning assessments, facilitator observations implemented at two points during the program, and qualitative data in the form of interviews with children, facilitators, community leaders and IRC staff, focus groups with caregivers, and session observations in the last few weeks of the program.

4.1 Participants

The quantitative baseline sample of the study includes data from 521 students (482 in home-based sites and 39 in center-based sites), who were either using Can't Wait to Learn numeracy (264 children) or Kitkit School literacy and numeracy software (257 children), and who received AL within 71 delivery sites (66 home-based and 5 center-based). Our sample consisted of children between the ages of 5 to 16. Children recalled eating 2.9 meals the day before and lived with 1.4 adults and 2.5 children on average. Ninety-seven percent of children lived with their mother and 91% lived with their father. Forty-seven percent reported that their mother could read. Forty-two percent of students in our sample reported that they attended some form of schooling before being displaced to Bangladesh, and 56% had never attended school. Of those who were registered in school, only 69% reported always attending school, 5% often, 15% sometimes, 9% rarely and 2% never attended school.

At endline, we tried to collect data from the same children and sites, but we were unable to successfully track more than 50% of children because the team had to interrupt research and implementation activities due to COVID-19. The endline sample includes data from 258 children (239 in home-based and 19 in center-based sites), of whom only 35 were using Can't Wait to Learn and 223 were using Kitkit School software.

	Baseline			Endline		
	Can't Wait to Learn	Kitkit	**Total**	Can't Wait to Learn	Kitkit	**Total**
Home	245	237	**482**	16	223	**239**
Center	19	20	**39**	19	0	**19**
Total	**264**	**257**	**521**	**35**	**223**	**258**

The total sample for the qualitative research consisted of 105 participants, which included 24 caregivers, 63 children, 10 facilitators, 4 community leaders, and 4 IRC staff members. The sample size comprised 53 people from Camp 8E and 48 people from Camp 22 and 4 staff members that worked across both camps.

4.2 Instruments

For the pilot study, we collected quantitative and qualitative data using the following instruments

Background questionnaire: A survey that gathers demographic information such as age, sex, fathers' level of education, mothers' ability to read, access to reading materials at home, and schooling experience before arriving in Bangladesh and after arriving in Bangladesh.

Hope and agency scale: A 6-item survey that uses a 5-point Likert scale (0=none of the time, 1=few times, 2=sometimes, 3=most of the time, 4=all of the time) to capture the frequency of children's experience with hope and agency.

ASER[14] literacy: A performance-based tool that captures levels of reading according to what a child can or cannot do:

- **Beginner:** A child cannot identify 4 out of 5 letters they attempt to read.
- **Letter level:** The child can correctly identify 4 out of 5 letters they attempt to read.
- **Word level:** The child can correctly identify 4 out of 5 words they attempt to read.
- **Paragraph level (Grade 1 level text):** A short 4 sentence passage of approximately 19 words at Grade 1 level that the child reads "like they are reading a sentence, rather than a string of words." The child can make 2 to 3 mistakes in reading words in the paragraph.
- **Story level (Grade 2 level text):** A 7 to 10 sentence story of approximately 60 words at Grade 2 level that the child reads "like he is reading a sentence, rather than a string of words" and "fluently with ease."

ASER numeracy:

- **Beginner:** The child cannot correctly identify 4 out of 5 randomly selected numbers from 1–9.
- **Level 1 - Number Recognition (1–9):** The child can correctly identify 4 out of 5 randomly selected numbers from 1–9.
- **Level 2 - Number Recognition (11–99):** The child can correctly identify 4 out of 5 randomly selected numbers from 11–99.
- **Level 3 - Subtraction - 2 digits with borrowing:** The child can correctly solve 2 out of 3 subtraction problems with borrowing.
- **Level 4 - Division - 3 digits by 1 digit:** The child can solve 1 division problem.

14. The Annual Status of Education Report is an assessment tool used in India and Pakistan to quickly and broadly assess reading and numeracy outcomes.

Adapted Teacher Classroom Observation (TCO) tool: A classroom observation tool with 11 items that enumerators use to rate the quality of a learning session according to facilitators' classroom management (e.g. facilitators are providing positive encouragement to children as they learn, facilitators are attentive to different needs of children throughout the learning session, facilitators observe each child and are available for questions) and children's engagement (e.g. learners appear focused on their activities during the session, learners are comfortable asking for help). The tool uses a 4-point scale (0=no evidence or negative, 1=emerging evidence, 2=proficient, 3=exemplary).

For the qualitative data collection, we used the following tools:

In-depth interview protocol:

- **For students,** our in-depth interview guide focused on understanding how children experienced the Pop-Up program in regard to the learning process, the facilitator, and the physical environment. We conducted semi-structured conversations with the following modules:
    - **Module 1 - Enrollment and attendance:** Designed to understand how children experienced their participation in the program.
    - **Module 2 - Experience in the program:** Designed to understand how children experienced a typical Pop-Up session, such as asking children what usually happened, what they enjoyed, what could be improved, what they found interesting and fun on the tablet, and what they found challenging. The module also covered some questions about school or learning centers to understand how the child perceived this program in relation to more traditional learning experiences.
    - **Module 3 - Experience with facilitators:** Designed to understand how children typically interacted with their facilitators, such as if they got help when stuck, what worked well for them, and what could be improved.
- **For facilitators,** our in-depth interview guide focused on understanding how facilitators experienced and executed their roles within the Pop-Up program, such as how they experienced the support and training they received and their perceptions of the value of the Pop-Up program as a learning intervention. We conducted semi-structured conversations with the following modules:
    - **Module 1 - Motivation for participation and engagement:** Designed to understand how facilitators learned about this program, why they applied for this position, and how sustained their engagement was throughout the 4 months.
    - **Module 2 - Experience with the program:** Designed to understand how facilitators perceived their role in the program, such as what parts of their role they found rewarding, what part they found challenging and why, how they felt about their position, and how their community and families perceived their job. We also asked about the flow of a typical session and the activities they usually performed with children.
    - **Module 3 - Perception on children's experience with the curriculum:** Designed to understand what facilitators observed during children's tablet time, such as if children seemed frustrated, engaged, experienced difficulties, asked for help, and so on.
    - **Module 4 - Perception of training, community of practice, and IRC support:** Designed to understand how facilitators experienced their light touch Pop-Up training and their on-going support from IRC staff.

- **For IRC staff** in charge of running the project, our in-depth interview guide focused on understanding the major lessons learned regarding the design, implementation, and monitoring of the Pop-Up program as well as sourcing recommendations for strengthening Pop-Up beyond the pilot phase. We conducted semi-structured conversations with the following modules:
    - **Module 1 - Software localization:** Designed to reflect on the process used to localize both software programs.
    - **Module 2 - Delivery model:** Designed to reflect on the different components of the program's delivery model, its structure, and the way it reaches children.
    - **Module 3 - technology infrastructure:** Designed to understand how the technology infrastructure supported the program and how moving the tablets around could be improved.
    - **Module 4 - facilitator support:** Designed to understand interactions with Pop-Up facilitators and how IRC staff perceived their support.
- **User interface and user experience testing scripts:** Conducted with the tablets, these testing scripts were designed to understand what aspects of the technology facilitated or inhibited learning and had a positive user experience, how children navigated the game interface, and what could be improved in terms of software features.
- **Focus group protocol:** We conducted focus group discussions with parents of Pop-Up students, with the objective to understand how parents understood the purpose of Pop-Up, what motivated them to enroll and maintain their children in the program, and what benefits and areas for improvement they observed regarding their children's participation in the program.
- **Session observation protocol:** During session observations, we aimed to observe how a session started, progressed, and ended; how children interacted with each other and with their facilitators; if children appeared to progress on the games as intended, or if they skipped parts of the games or got stuck; and if any outside disruption was observed.

4.3 Process of data collection

For quantitative data collection, we hired a team of 18 enumerators and provided them a 3-day training which focused on how to obtain consent from parents and assent from children to participate in the feasibility study and how to administer the ASER assessment followed by the baseline survey. Enumerators administered the ASER assessment individually with each child, using pen and paper for up to 30 minutes of their time. Any child who scored between 0-2 on the ASER test for literacy and mathematics was invited to enroll in the program. If the child agreed to enroll, the enumerator administered the baseline survey directly after the ASER assessment. The survey was administered individually though the use of a tablet and took on average 45 minutes to complete. Throughout the program, project assistants received training to conduct classroom observations using the TCO and visited learning sites to conduct two rounds of 60 to 90-minute observations during month 1 and month 3 of the program.

Tablet analytics were collected by our two software partners from their respective servers and shared with the IRC for further analysis.

For qualitative data collection, a team of two qualitative and design researchers from the Airbel Impact Lab traveled to Bangladesh to provide guidance on collecting the first round of observations, interviews, and focus groups. The team trained two IRC Bangladesh staff to continue conducting these observations for the following two weeks. Data was collected in the Rohingya language and translated into English by the team of translators.

4.4 Analytical approach

Members of the Airbel Impact Lab conducted three sets of data analysis. We analyzed quantitative data using STATA 15. We looked at means and frequencies for descriptive statistics. In order to confirm that we observe learning gains in students' literacy, numeracy, and hope and agency during the pilot intervention we identified average score changes in outcomes of interest, and documented changes in the proportion of children at different levels of performance. Given the small sample sizes in the study and the great loss of data at endline due to COVID-19, the findings include information from children who had both baseline and endline data, as well as children who participated in the program from the beginning but only participated in the research at endline.

We analyzed qualitative data, including observations, interviews and focus groups, by identifying themes that recurred across qualitative data sources and aligned with our topics of inquiry including program experience, perceived educational value and quality, and alignment with social and community values.

We also conducted a cost driver analysis. We took actual expenses incurred during the program and attributed them to specific line items in our budget to understand the actual amount spent for each budget category. We calculated the cost-per-child to understand cost-efficiency of the feasibility study and built a scale model to estimate cost-per-child of the program at scale.

5.0 Results

1. Participant outcomes

What baseline-endline changes did we observe in learning (literacy and numeracy) and SEL (hope and agency) outcomes?

In the sections below we present gains made by children in literacy, numeracy, and SEL outcomes. Please note that baseline-endline gains are not a good measure of program impact because the growth observed may be due to other factors such as development. In order to assess impact, we would need to compare treatment and control groups with similar characteristics to be able to account for the changes we would have observed in the absence of the intervention. Please note also that the information below compares baseline and endline scores using samples that include data from children who had both baseline and endline scores, as well as students who participated in the program from the beginning but only had endline data points. The information below, however, is useful in that it can help us confirm—in a short amount of time during the pilot—whether we are observing changes in skills in the expected direction.

Literacy

In this section we present the baseline-endline changes observed among only those children who used Kitkit School because they had access to both literacy and numeracy software. There is no data provided for children who used Can't Wait to Learn because they only had access to numeracy software. At baseline, we observed that 71.60% of children in the sample were not yet able to read words correctly (ASER Levels 0 and 1), and 28.40% were only able to read words and not sentences (ASER Level 2). After 4 months of participation in the AL program, we observed an increase of .21 ASER levels on average. Specifically, we observed that the percentage of children who were unable to read words (Levels 0 and 1) decreased 24.06 percentage points to 47.54%, as they moved to more advanced literacy levels. We also observed that 40.81% of children were able to read words correctly (Level 2), 4.04% of children were able to read a short Grade 1 paragraph correctly (Level 3) and 7.62% were able to read a Grade 2 passage correctly (Level 4).

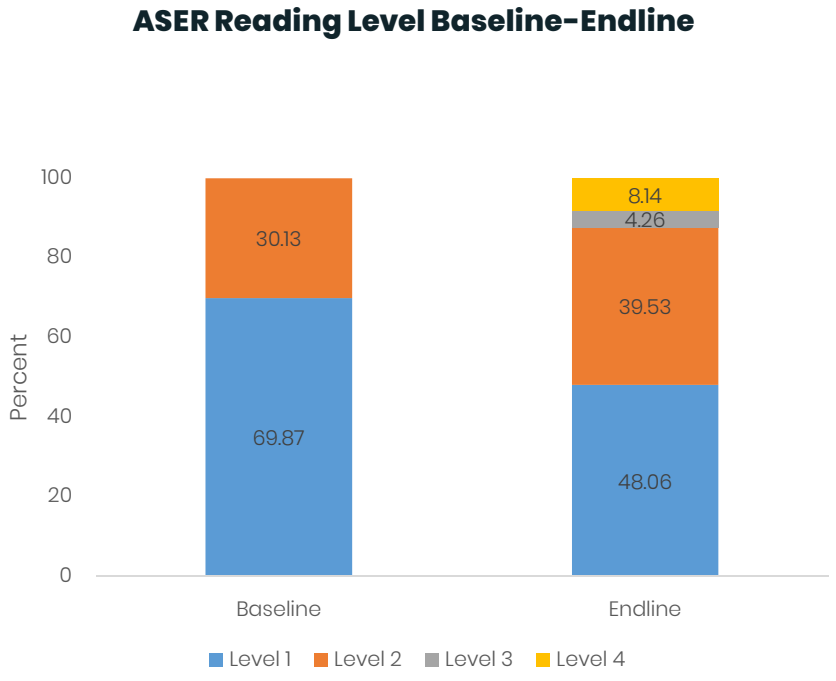


Figure 1

Note: Students in Kitkit school spent half of their time with AL tablets engaged in literacy activities and half in math activities. Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

What are the baseline-endline changes in literacy levels for each delivery model?

In the home-based learning sites where children used Kitkit school to learn and practice literacy skills, we observed improvements in foundational literacy skills. Unfortunately, we were unable to document changes in the literacy skills of children using Kitkit School in center-based sites as endline data collection was interrupted due to COVID-19 before we could visit centers. On average, children in home-based sites improved their ASER literacy level by .13 levels, and children in center-based sites improved by .39 levels. The figure below shows the proportion of students according to ASER reading levels at baseline and endline by delivery model. We observe that at baseline, the proportion of children in levels 0 and 1 decreased from 71.15% to 40.81%, as 24.62% made progress to more advanced reading levels. Specifically, at endline, we observe that 40.81% children were able to read words correctly (level 2), 4% were able to read a grade 1 reading passage correctly (level 3) and 7.62% were able to read a grade 2 reading passage correctly (level 4). Please note that given the small sample size used in the pilot and the great loss of data at endline due to COVID-19 related disruptions, the findings below should be interpreted with caution.

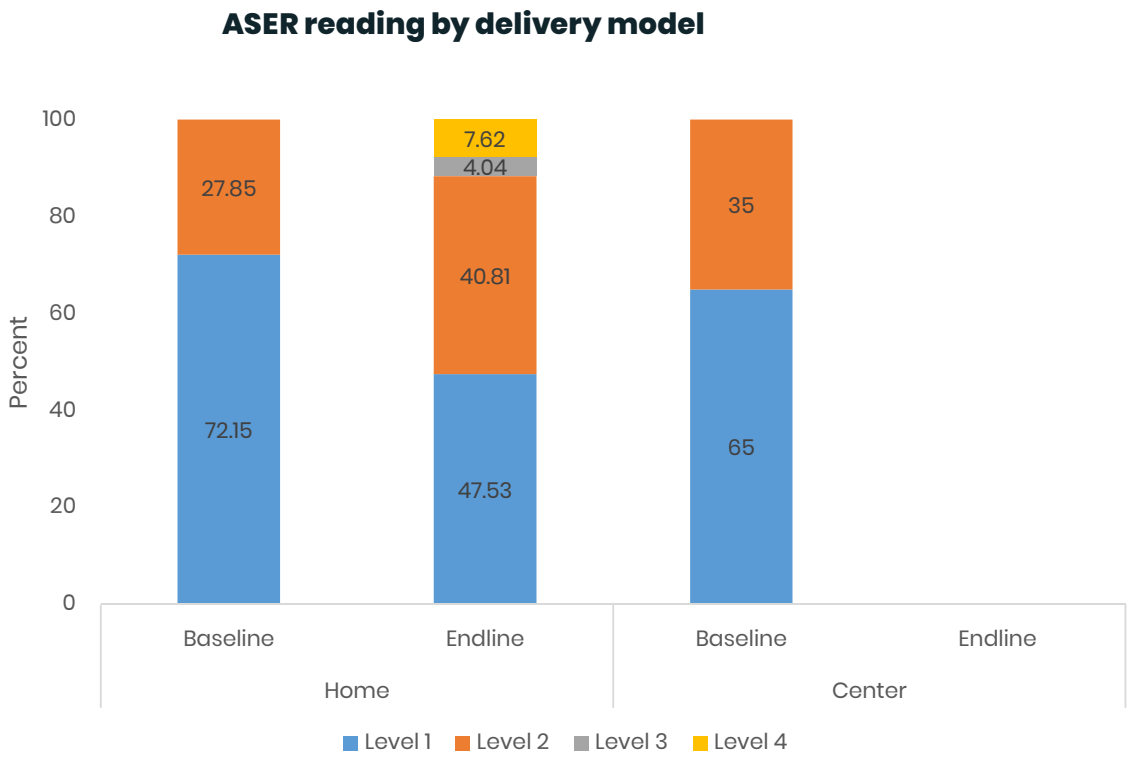


Figure 2

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

Numeracy

After 4 months of participation in AL, we observed positive changes of .33 ASER math levels on average. Specifically, we observed that at baseline, 81.77% of children in the sample were not yet able to correctly identify double-digit numbers (ASER Levels 0 and 1), and 18.23% were able to correctly identify numbers from 11-99 (ASER Level 2). Figure 4 shows that, after participating in an AL program for 4 months, the percentage of children in Levels 0 and 1 decreased by 30 percentage points from 81.77% to 51.55%, as children progressed to more advanced math levels. Specifically, we observed that at endline, 46.51% of children were able to correctly identify numbers from 1 to 99 (Level 2), a small percentage (0.78%) of children were able to conduct subtraction (Level 3) and 1.16% of children were able to conduct division (Level 4). Please note that ASER does not adequately capture changes in children's math skills as the tool does not assess children's ability to conduct addition, a math skill significantly easier than conducting subtraction and division.

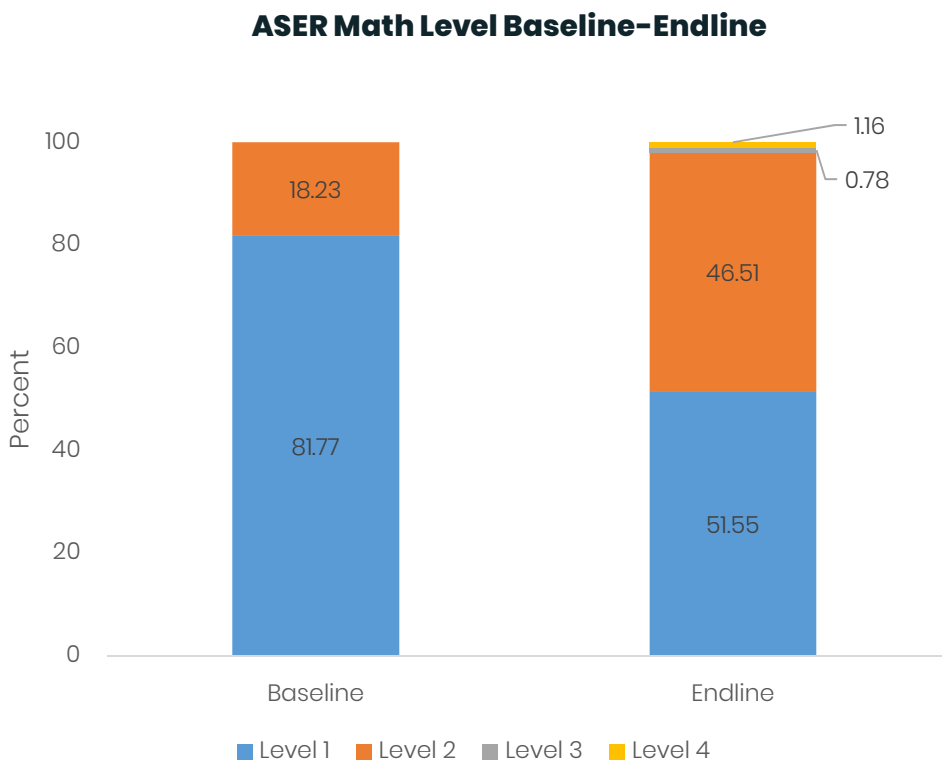


Figure 3

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

What are the baseline-endline changes in numeracy levels for each delivery model?

The data suggest that children in both the home- and center-based sites showed positive changes in the baseline-endline learning gains in math. On average, children in home-based sites improved their ASER math level by .33 levels on average, and children in center-based sites improved by .33 levels. The figure below shows the proportion of students according to ASER math levels at baseline and endline by delivery mode. The data below should be interpreted with caution given the lack of random assignment to home and center-based settings and the small samples for each subgroup.

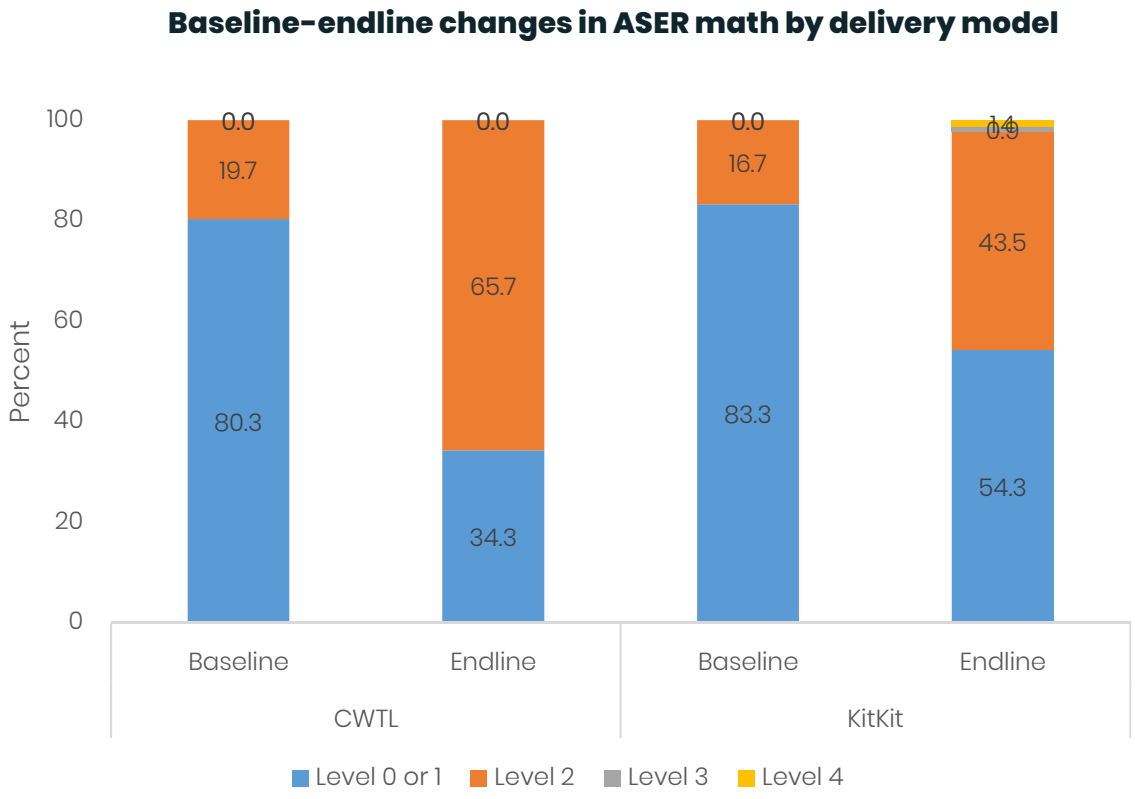


Figure 4

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

What are the baseline-endline changes in numeracy levels for each software?

Children using Can't Wait to Learn spent the entire time they had access to tablets learning and practicing their numeracy skills. In contrast, children using Kitkit School spent less than half their time on numeracy and the remaining time on literacy and other games in the tools section of the application. Given the difference in dosage, the results associated with Kitkit School and Can't Wait to Learn should not be directly compared. We observed a positive change of .46 ASER numeracy levels on average for children who used Can't Wait to Learn. Figure 5 shows the proportions of students using Can't Wait to Learn at each ASER numeracy level at both baseline and endline. We observe that at baseline, the great majority of children using Can't Wait to Learn were not able to identify double-digit numbers (Levels 0 and 1). At endline, 46% of children moved to Level 2 and successfully identified numbers 10-99.

As mentioned, children using Kitkit School split their time on the tablet between numeracy, literacy and other tools that support self-expression through the arts. In this case, we observe an improvement of. 33 ASER numeracy levels on average. Figure 6 shows the proportions of students using Kitkit School at each ASER numeracy level at both baseline and endline. We observe that at baseline, 83.3% of children using Kitkit School were not able to identify double-digit numbers (Levels 0 and 1). At endline, 29% of children moved to Level 2 and successfully identified numbers 10-99. Additionally, we also observe a very small percentage of children (<1%) who progressed to ASER Levels 3 (subtraction) and 4 (division).

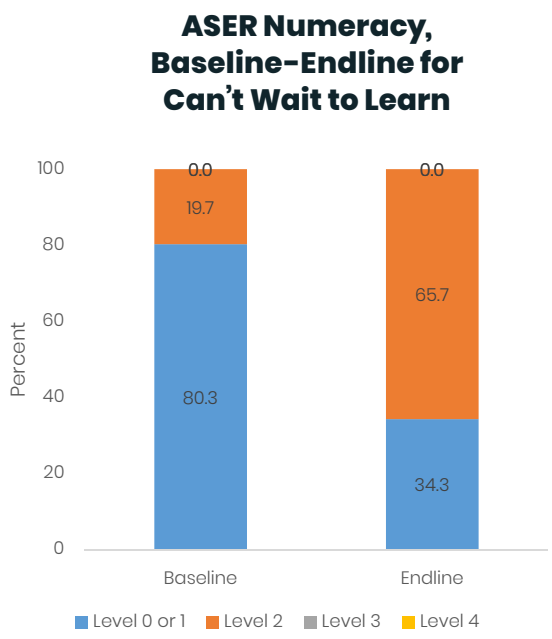


Figure 5

Note: Students using Can't Wait to Learn Software spent the entire time with AL tablets engaged in math activities. Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

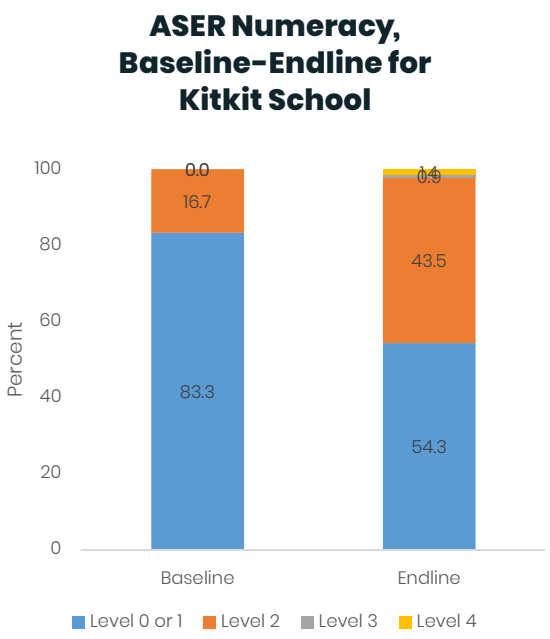


Figure 6

Note: Students in Kitkit School spent half of their time with AL tablets engaged in math activities and the other half in literacy activities. Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

Hope and Agency

We observed improvements in children's hope and agency after participating in 4 months of AL. Figure 7 shows that at baseline, 6.35% of children reported low levels of hope and agency, 33% reported medium levels, and 60.4% reported high levels. At endline, the percentage of children with low levels of hope and agency decreased by 5 percentage points, and the percentage of children reporting high levels increased by 9 percentage points.

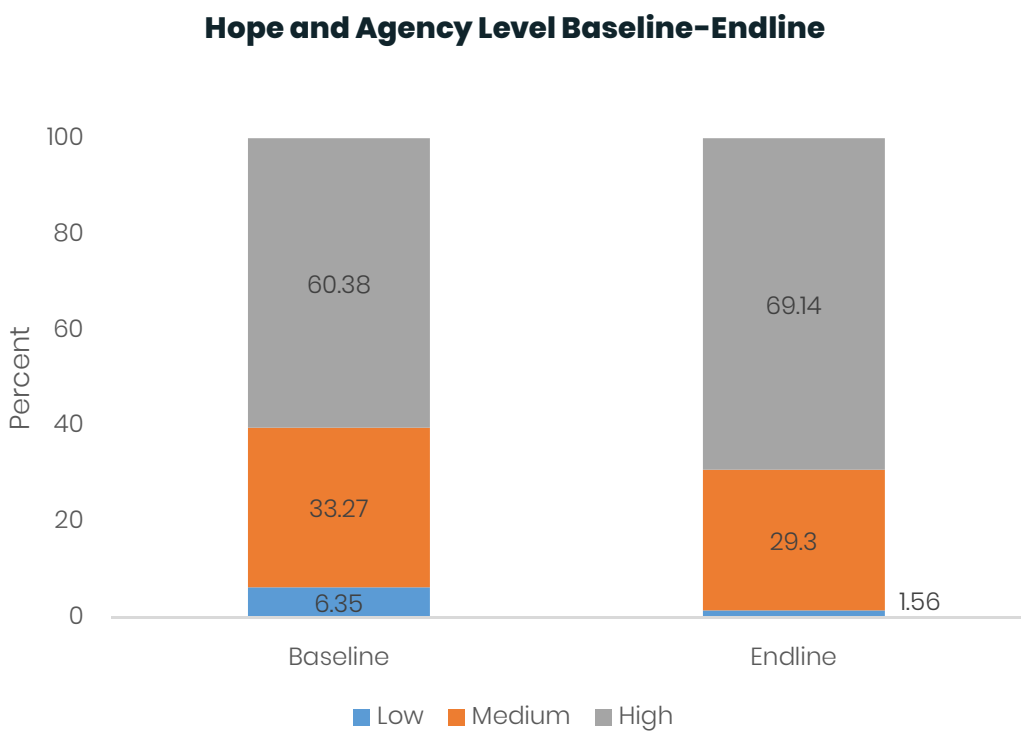


Figure 7

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

What are the baseline-endline changes in hope and agency levels for each delivery model?

We observed improvements of .66 likert scale points in children's hope and agency in the center-based setting and improvements of .10 likert scale points in the home-based setting. Figure 8 shows that in the home-based settings, 6.44% of children reported low levels of hope and agency and 58.84% reported high levels, compared with 5.13% and 79.49% reporting low and high levels in the center-based settings. At endline, we observed a reduction of 4.75 percentage points for children with low levels of hope and agency in the home-based settings, and a reduction to zero in the center-based setting. We also observed that the percentage of children with high levels of hope and agency increased by 7.83 percentage points in the home-based setting and by 20.21 percentage points in the center-based setting. Note that findings need to be interpreted with caution given the lack of random assignment to home- and center-based settings and the small sample sizes for each subgroup.

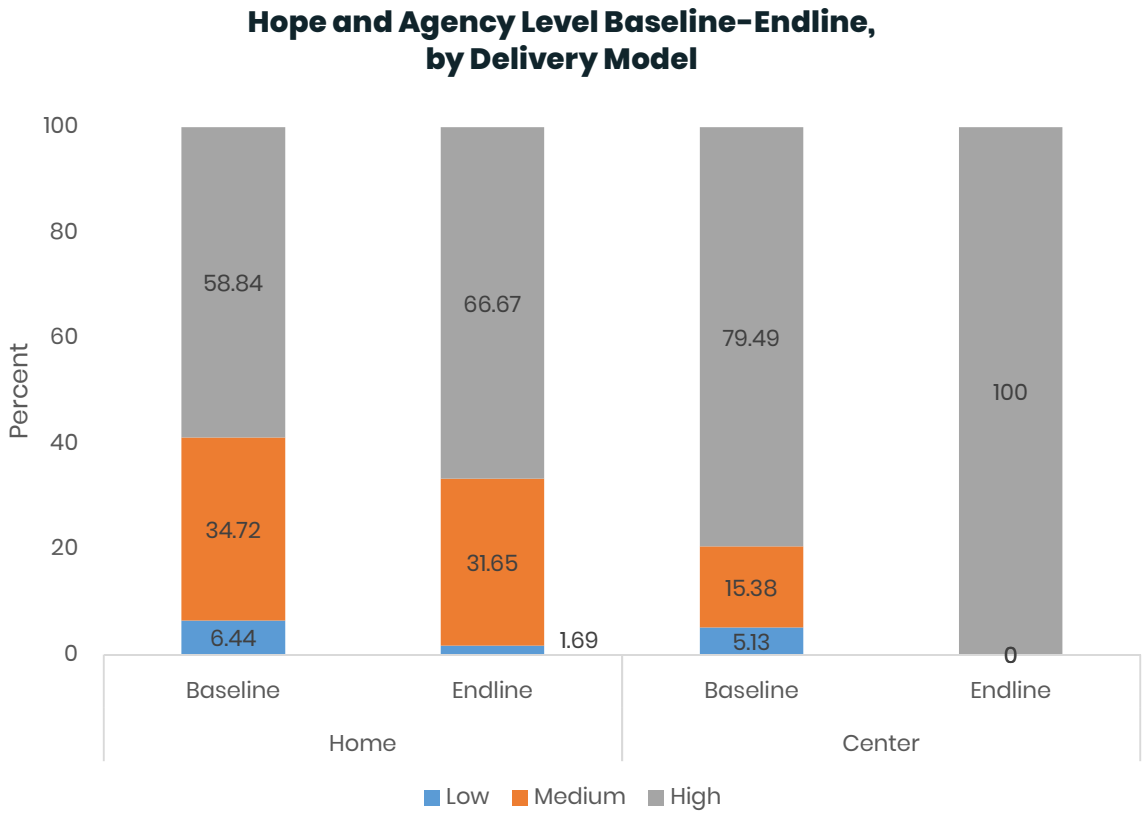


Figure 8

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

What are the baseline-endline changes in hope and agency levels for each software?

We observed improvements of .74 scale points in the hope and agency of children using Can't Wait to Learn software. Figure 9 shows that in the Can't Wait to Learn group at baseline, 5.68% of children had low levels of hope and agency, 31.82% had medium levels and 62.5% had high levels. At endline, we observed significant progress, with 94.29% of children who used Can't Wait to Learn reporting high levels of hope and agency.

We also observed improvements of .08 likert scale points in the hope and agency of children using Kitkit School. Please note that these findings need to be interpreted with caution given the small sample sizes. Figure 10 shows that the percentage of children with low levels of hope and agency decreased by 5.22 percentage points, from 7.03% to 1.81%, while the percentage of children with high levels of hope and agency increased by 6.96 percentage points, from 58.2% to 65.16%.

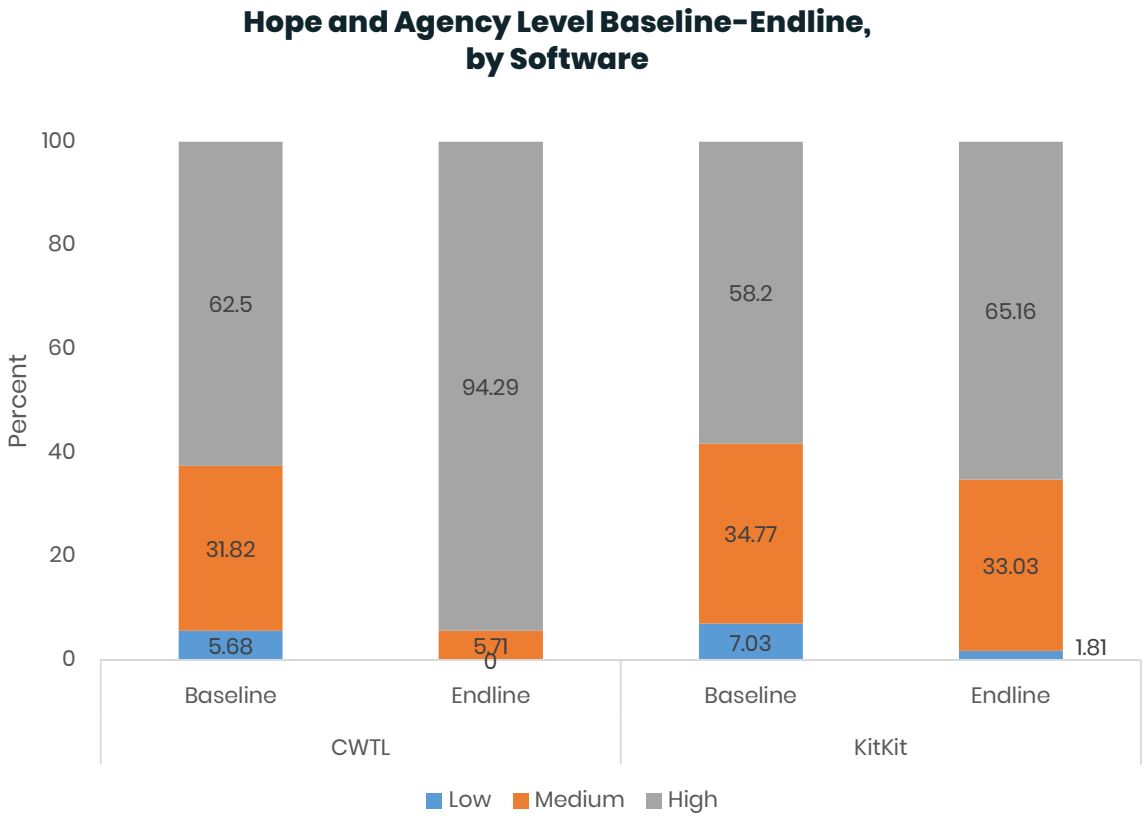


Figure 9

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

2. Implementation Fidelity

What levels of attendance, dosage, progress and engagement and overall levels of facilitator proficiency do we observe in the AL sessions?

What levels of attendance did we observe?

Facilitators manually collected child attendance data throughout the pilot, however, due to the COVID-19 lockdown in Bangladesh, IRC staff have not been able to access and analyze these records. Below we use a combination of self-reported attendance from the endline survey and tablet analytics to assess child attendance for the program.

In our endline survey, 76.36% of children reported attending at least 3 out of 4 learning sessions per week.

Attendance was also measured using tablet analytics collected by Kitkit School and Can't Wait to Learn. Attendance was measured using two different methods for each of the software due to the unique structure and limitations of the data.

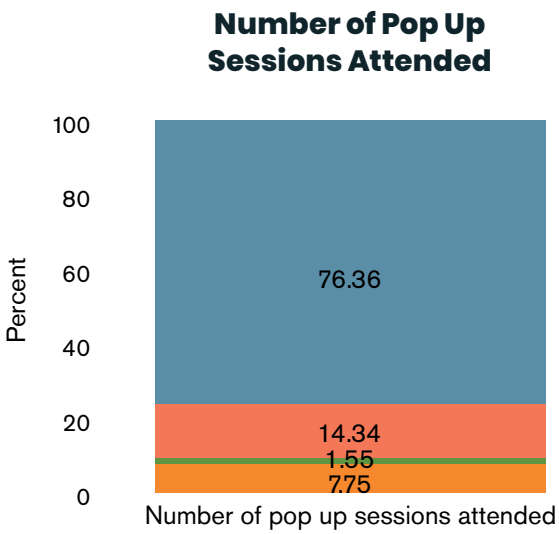


Figure 11

KITKIT SCHOOL

During this pilot, because of issues we faced with students' unique logins for Kitkit School leading to different children accessing the same profile, we were unable to track activity and attendance for individual students. This problem persisted from the program start date, November 17, 2020, until January 19, 2020, when log-in issues were remedied.

However, we were able to deduce attendance using tablet analytics from Kitkit School by comparing the number of intended unique learning sessions each day to the actual number of unique learning sessions that took place each day.

Using this method we found students attended 56% of learning sessions throughout the program on average in Camp 8E.[15] This may underestimate attendance, because 16% of children using Kitkit School in Camp 8E were enrolled in the program late, starting several weeks after launch. The late start is not accounted for in this analysis.

CAN'T WAIT TO LEARN

We were able to deduce attendance using tablet analytics from Can't Wait to Learn by calculating the number of dates each student profile was active within our implementation window (November 17, 2019 to March 12, 2020) and comparing that to the number of dates learning sessions were meant to take place.

Students using Can't Wait to Learn attended 35 out of 58 total planned sessions (60%) on average. Most students attended 40 out of 58 total planned sessions (69%).

15. Camp 8E only. Camp 22 data was not usable due to a network issue for synchronization.

What amount of time did children spend engaged with different software programs?

Children were intended to receive 45 to 60 minutes of learning on the tablet depending on the group in which they were enrolled. All home groups and one center group aimed for children to receive 60 minutes of learning on the tablet. Two center groups aimed for children to receive 45 minutes of learning on the tablet. Learning was limited to 45 minutes in two center groups because other programming separate from Pop-Up was already taking place in the center and time slots for Pop-Up learning sessions were limited.

CAN'T WAIT TO LEARN:

To better understand the number of minutes children spent actively using the tablet per session, we calculated the average minutes played for each session attended for all students.

Students using Can't Wait to Learn in the home experienced an average dosage of 54 minutes learning on the tablet per session or 90% of the intended dosage. Students using Can't Wait to Learn in the center experienced an average dosage of 51 minutes learning on the tablet per session or 113% of the intended dosage.

Can't Wait to Learn	Intended Dosage	Average Dosage
Home	60 minutes	54 minutes
Center (SHLS)	45 minutes	51 minutes

These findings indicate high implementation fidelity and align with our observations during the qualitative research

KITKIT SCHOOL

We are unable to accurately report on the dosage for Kitkit School. We were unable to track time spent learning on the tablet per session due to network synchronization challenges and some children accessing and learning on the same profile. However, initial analyses of Kitkit School data showed that participants spent approximately equivalent time on each subject in Kitkit School. This implies that the dosage for Kikit-math would be about half of the total dosage and less than half of the numeracy dosage received by children learning on Can't Wait to Learn, which was math-only.

How far did children progress through the game? What learning pace, time on task, and success rates did we observe among participant children?

KITKIT SCHOOL[16]

Progress

Students progress through the Kitkit School software by hatching "eggs" or course levels and growing their creatures in literacy and math "coops." There are 11 egg levels each for literacy and math, including an introductory course focused on cognitive and digital skill-building as well as the final post-course. Each egg consists of a range of sessions (approximately 30 on average) with sequenced activities to develop foundational skills and confidence.

Kitkit School's curriculum includes 280 sessions with more than 1,200 activities for literacy and 260 sessions with approximately 1,200 activities for math. Each session includes 3 to 6 activities. Children solve multiple problems in each activity. Most egg levels have 3 mini-quizzes within a series of sessions. Activities are completed and celebrated with a star, moving children toward mastery with scaffolded reinforcements. Sessions are completed with coins awarded, and children's creatures grow in their coops. All courses have one final post-course, the "egg quiz," after which the creature is crowned.

Children must score 80% or higher on the egg quiz to unlock the next level. If they score below 80%, they must revisit the activities and retake the quiz until they pass. In the prototype software version built for this pilot program to provide insights on initial learner readiness, children were able to access the first 3 egg levels in literacy and the first 5 egg levels in mathematics without having to pass the egg quiz. In future deployments of this software, we would restrict egg access and require that all students pass each egg quiz in order to advance to the next level. This will ensure that children progress sequentially and cannot advance without gaining necessary foundational skills, as designed.

We were not able to track progress for individual students using Kitkit School due to issues with network synchronization and children sharing or switching between profiles. Instead, we studied children's progression through the learning software by looking at the number of attempts on quizzes (Figure 12, Figure 14) compared to the number of completed quizzes (Figure 13, Figure 15), over the course of the 5 months in which the program was implemented. Completion means that a student completed all the questions. This is different from a "pass" or "success" rate, which is getting 80% or higher correct responses.

We found that students' attempts to solve the quiz focused on the open levels (1-3 for literacy and 1-5 for math). While some advanced to higher levels in the software, student progress dropped off after the open levels (Figure 13, Figure 15). While the pilot version of this software was designed to provide insights into the children's initial learning capacity and different learner profiles within the camp, many children found ways to repeat activities they enjoyed and/or circumvent some of the more challenging activities. Nevertheless, high engagement overall and persistence with quiz attempts evidence children's learning despite the operational and data-related issues that compromised our analysis

11. Camp 8E only. Camp 22 data was not usable due to a network issue for synchronization.

Literacy progress

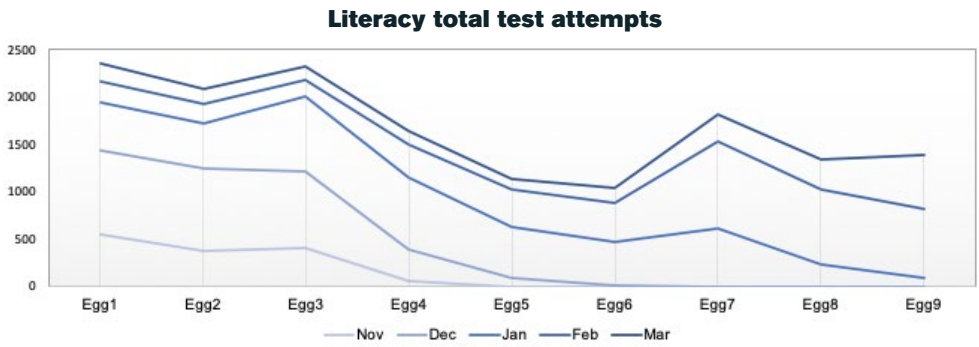


Figure 12

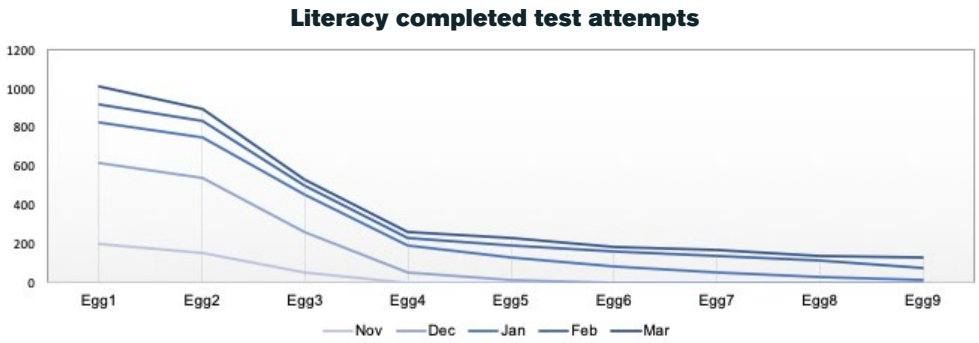


Figure 13

Math progress

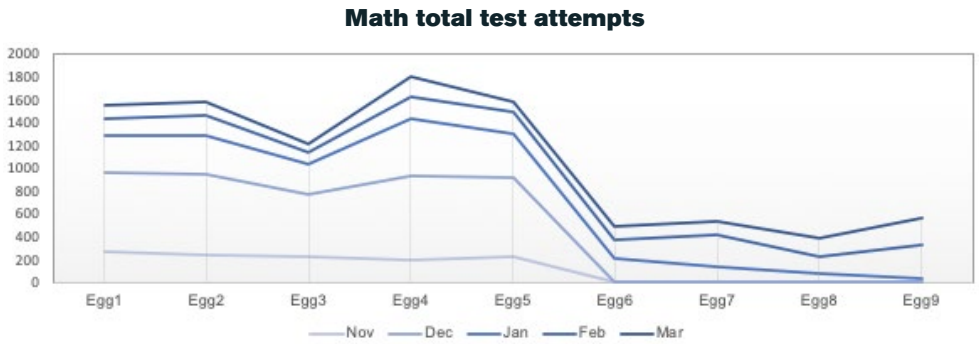


Figure 14

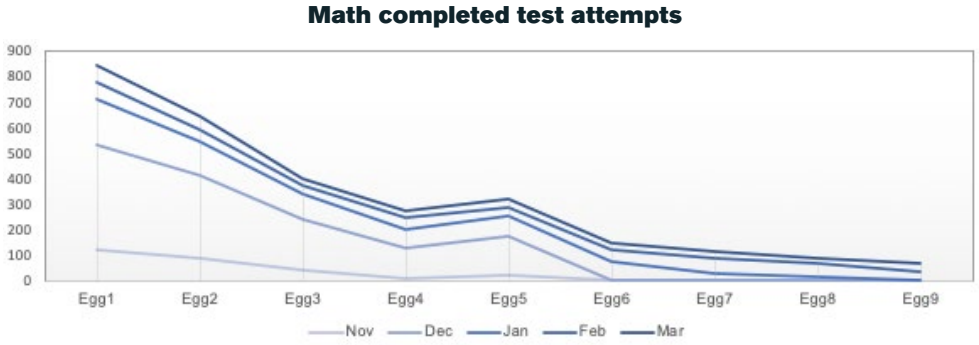


Figure 15

Learning pace

Students spent an average of 1.5 minutes on each completed activity for literacy and 1.7 minutes on each completed activity for math. When children left the game without completing it, they spent an average of .7 minutes per activity for literacy and .9 minutes per activity for math.

We see that children spent about the same amount of time on literacy content as they did on math content. It should also be noted that facilitators in Camp 8E decided amongst themselves during a COP meeting that they would implement time limits on each subject, asking their students to play math games for approximately half the session and literacy games for the other half of the session.

Total time on literacy and math games and videos

	Game time (minutes)	Video time (minutes)	Total time (minutes)
Literacy	153,095	44,674	197,769
Math	170,536	22,062	192,598

Time on task

Students using Kitkit School spent the most time engaging with the game compared with the books and videos (Figure 16). The majority (72%) of children's time was spent in the Learning App (game-based curriculum) followed by 25% on learning videos, and 3% on books available in the Library.

Playtime by contents

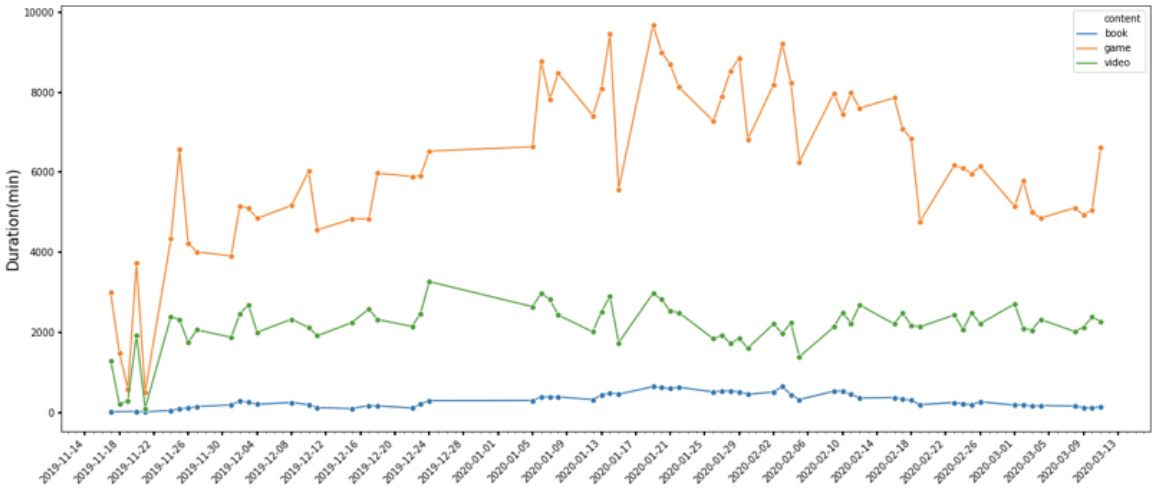


Figure 16

Students using Kitkit School spent the most time accessing content through the Kitkit School channel compared with the library. The average ratio of Kitkit School to Library was 86 to13 (Figure 17).

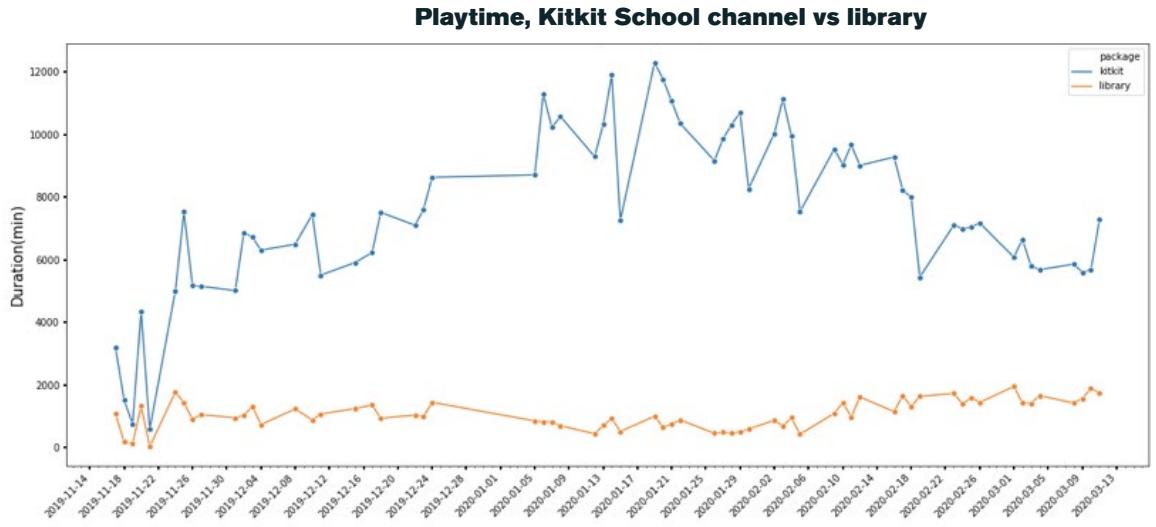


Figure 17

Time on task by content (game, books, videos) in Figure 16 and time on task by channel (Kitkit School, library) in Figure 17 do not include total time spent on the tablet because the time spent on Kitkit School's suite of tools for creativity and self-expression was not tracked in the tablet analytics

Success rates

In light of multiple challenges with tablet synchronization and user login issues, success rates at the individual user level were not analyzed for this initial pilot. To better understand user behavior and readiness as children progressed through the curriculum, the Kitkit School team looked at completion rates for literacy and math games.

Students using Kitkit School experienced an average completion rate of 82.2% for literacy games and 79% for math games over the course of the pilot. This suggests that across all levels, children were deeply engaged with and persisted in completing the activities they encountered.

CAN'T WAIT TO LEARN

Progress

Students progress through the Can't Wait to Learn software by levels. There are 91 total levels with multiple "minigames" included in each level. The amount of minigames by level varies and is predefined by educational experts. Each minigame includes a series of mathematical problems that help teach a concept. Students advance to the next minigame only when they have mastered the minigame they are currently working on. After completing all the minigames in one level, children are automatically presented with a video to teach a new concept.

A total of 261 children were registered and tracked using Can't Wait to Learn for the duration of the program. Out of 91 total learning levels in mathematics, students completed 44 levels, or 48% of the total mathematics content on average over the course of 16 weeks of learning and 58 learning sessions. Forty-four students (17%) completed all 91 levels. The most common levels reached were levels 21 through 30, which represent 23% to 33% of the total math content (Figure 18).[17]

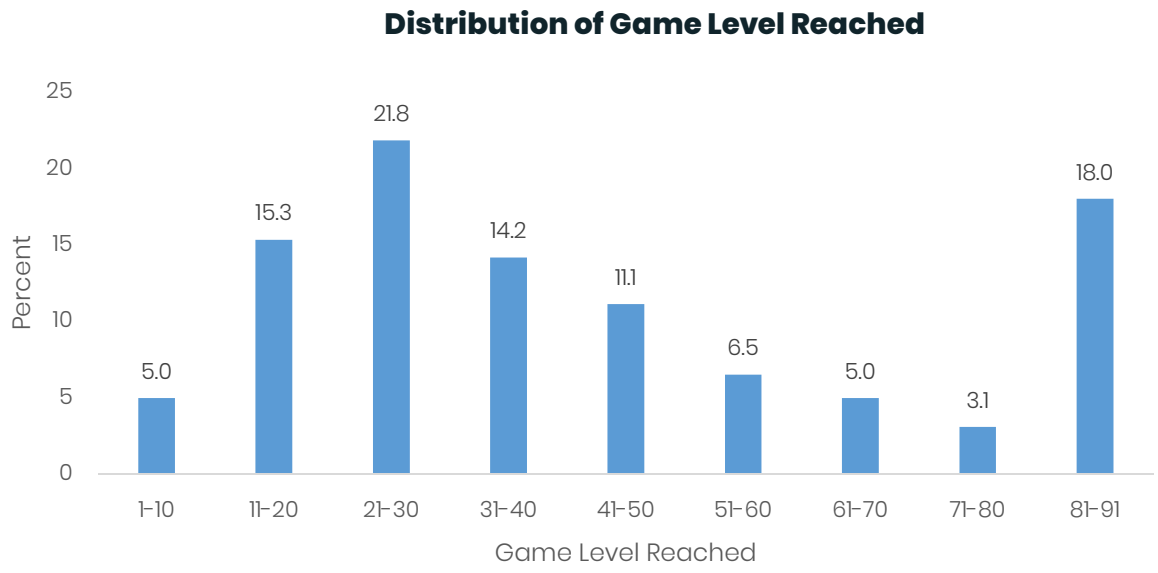


Figure 18

17. The Can't Wait to Learn software version used in this pilot program included all math content currently available through the software, but with less repetition than other Can't Wait to Learn deployments due to the pilot having a shorter timeline.

Learning Pace

Students spent 38 seconds on average on each minigame. Learning pace per minigame did not change significantly as children progressed through the game (Figure 19).

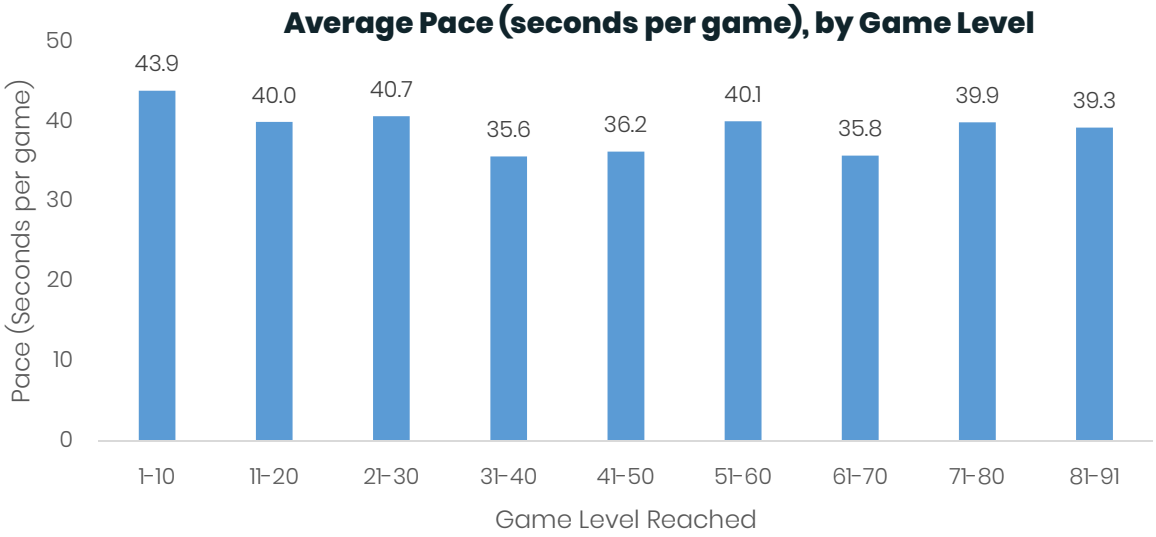


Figure 19

Time on task

We were unable to report time on different tasks for Can't Wait to Learn. Can't Wait to Learn only covered mathematics, which prevented disaggregation of playtime by subject. Additionally, playtime by channel (minigame versus video) was not disaggregated in the tablet data.

Success rates

Students played a total of 2,766 minigames over the course of the program and won 1,238 (45%) of those games. The win percentage is larger (51.5%) for students who advanced the farthest in the game to levels 81 through 91, compared with students who did not advance very far in the game to levels 11 through 20 (41.3%) (Figure 20).

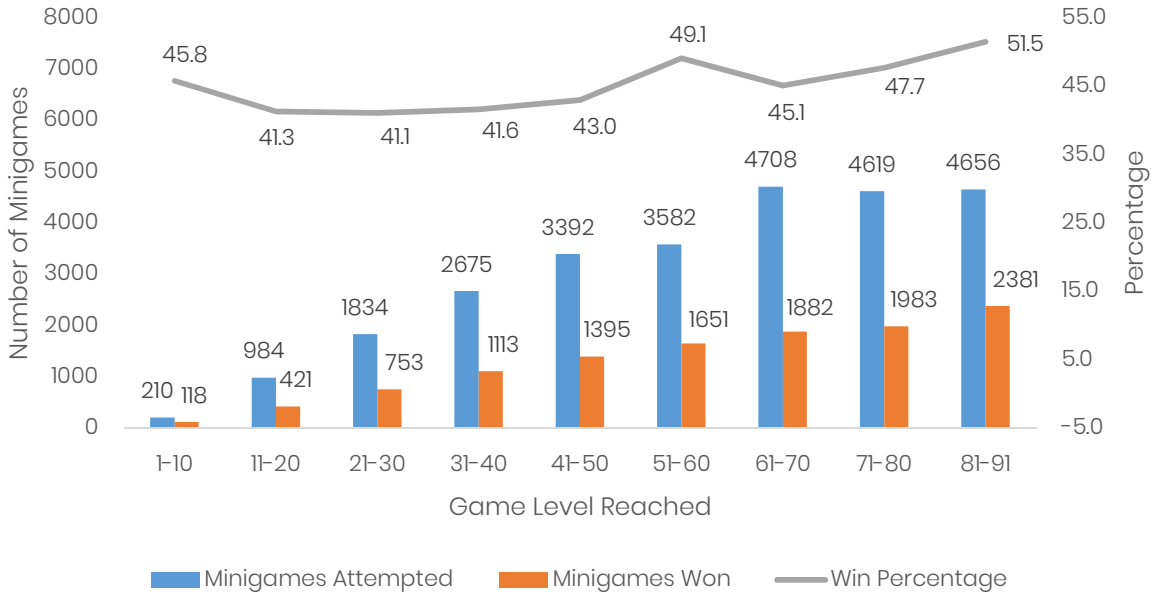


Figure 20

What overall levels of implementation quality do we observe in the AL sessions?

Facilitators' proficiency was assessed with regard to behaviors such as providing positive encouragement, being attentive to the needs of children, playing games and practicing calming activities, using positive discipline strategies to respond to misbehavior, and helping children find their own solutions. Data from classroom observations indicate that 24% of the facilitators were rated with some emerging evidence of proficiency when conducting sessions. Approximately 76% showed either good or exemplary levels of proficiency. Children's behaviors were assessed with regard to the degree to which they appeared to be focused on the activities, took good care of the tablets, and felt comfortable asking for help. We observed that more than 86% of children behaved in proficient or exemplary ways, with only 13.8% of them showing emerging evidence of good behavior. Finally, the classroom environment was assessed with regard to whether the space was clean, spacious, and free from distractions. The data indicates that more than 90% of the learning environments had good or exemplary evidence of these characteristics.

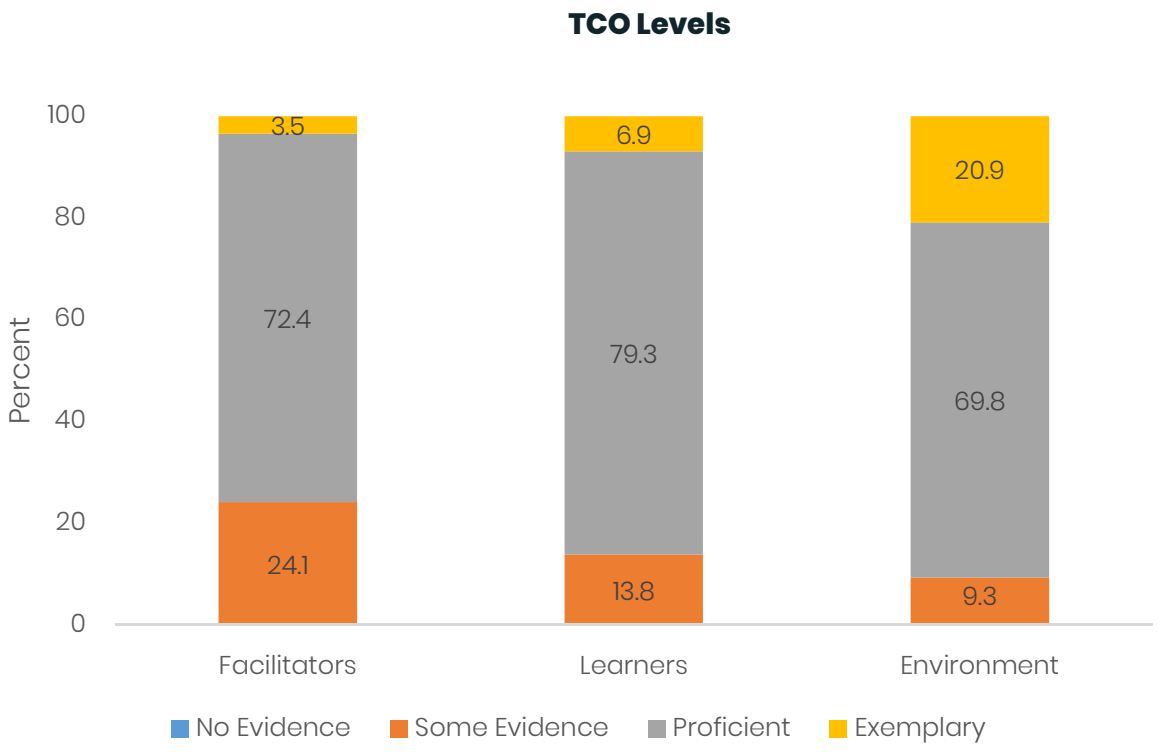


Figure 21

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

What levels of implementation quality do we observe in home-based and center-based delivery models?

Below, we present some comparisons between the quality of the learning sessions in the home- and center-based settings, which should be interpreted with caution, given the very small sample size of students in each of the groups, especially in the center-based settings. On average, the quality of the learning sessions was similar in center- and home-based settings. Figure 22 shows that the majority of facilitators in both settings delivered learning sessions that met standards of proficiency, with less than 25% of them in the home-based settings and 20% of them in the center-based setting, exhibiting only emerging levels of proficiency. Similarly, we observed that the majority of children in both settings showed good levels of engagement, but we observed more variety of performance in the home-based settings, perhaps due to the fact that the sample was significantly larger for this group. Finally, the quality of the environment in the center-based settings was better than in the home-based settings, but this may be due to the fact that the pilot had very few center-based sites to observe.

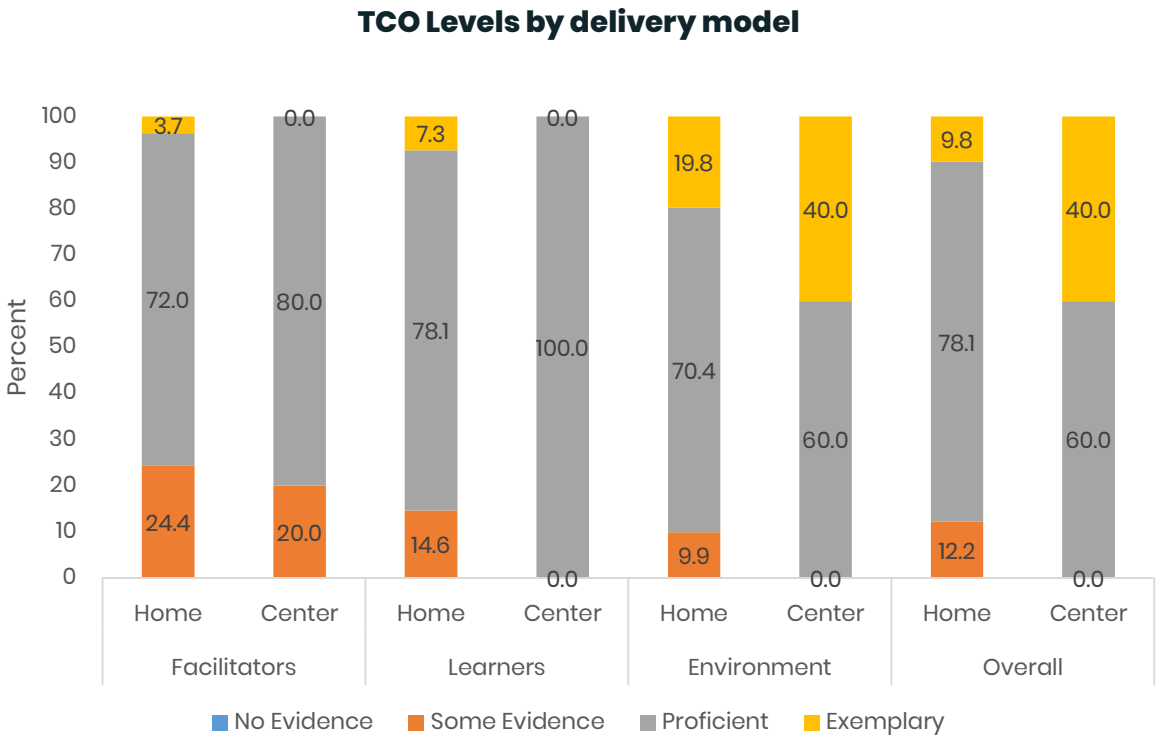


Figure 22

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

What levels of implementation quality did we observe for each software program?

Finally, we observed the quality of the learning environments in the sites that used different AL software. Once again, the data should be interpreted with caution given the small sample sizes used in the pilot study. We did not observe differences in the quality of the learning sites by software, but we saw slightly higher averages in the quality of the facilitation provided in learning sites that used Can't Wait to Learn, and slightly higher levels of engagement with facilitators among students in the Kitkit School learning sites.

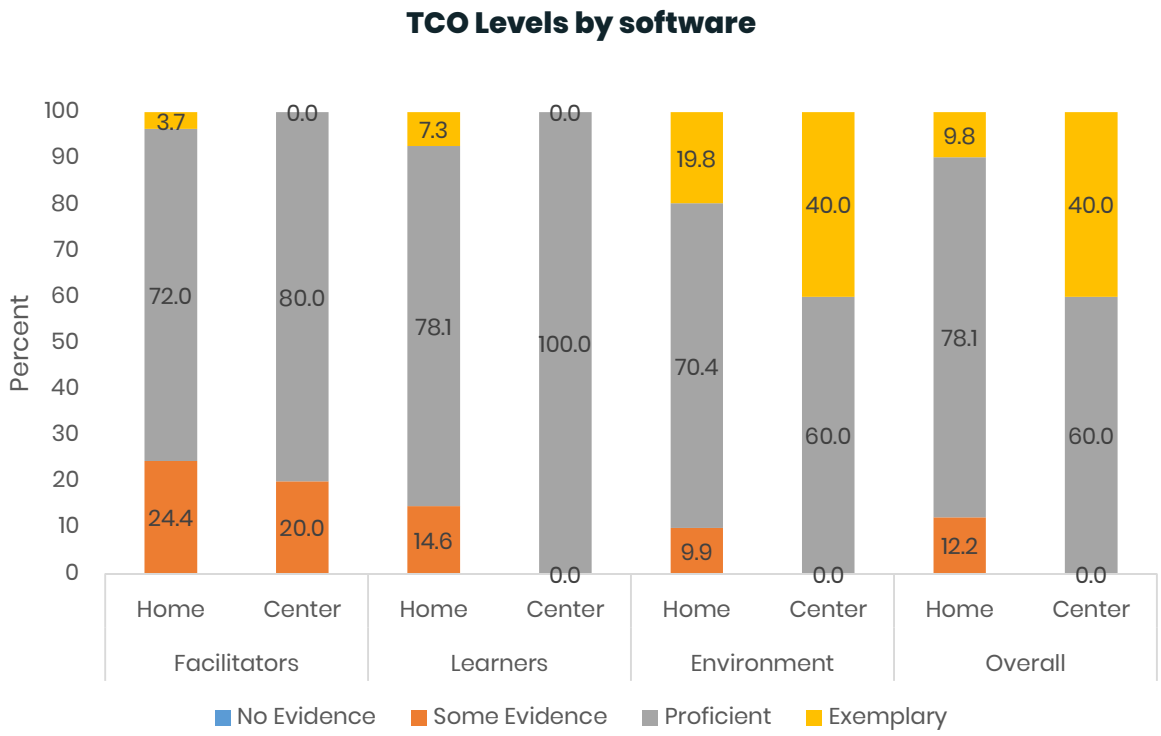


Figure 23

Note: Findings need to be interpreted with caution given the small sample size and the great loss of data at endline due to COVID-19 related disruptions.

Additional qualitative notes were collected alongside each indicator in the TCO:[18]

Facilitators often showed attention to children by making rounds and checking their activity during the session, however, they might have been unable to help the child if the child was stuck or struggling. Enumerators reported on their notes from their classroom observation tool:

- **"She looked after the children's activities by making rounds to each child. But she does not know how to understand."**
  — Enumerator observing a session in a home in camp 8E

- **"She looks after each child and provides the proper instructions when required."**
  — Enumerator observing a session in a home in camp 8E

Facilitators ensured that children's hands were clean before using the tablets, and they often implemented a warm-up activity called "The Train Game" that they learned during their kickoff training. Enumerators reported on their notes from their classroom observation tool:

- **"The Train Game was played at the beginning and they checked whether the nails are neat and clean or not."**
  — Enumerator observing a session in a home in camp 8E

- **"At the beginning of the session she ensured the cleanliness of each child."**
  — Enumerator observing a session in camp 8E

Qualitative observations demonstrated that most facilitators did try and implement a game at the beginning of the session, but that it could be repetitive, too fast, and not done "right" per the observer's judgment. Enumerators reported on their notes from their classroom observation tool:

- **"Facilitators do not play games regularly as the children were not comfortable with the game."**
  — Enumerator observing a session in camp 22

- **"Facilitators try to play a game but have a hard time doing it right or go very fast."**
  — Enumerator observing a session in camp 22

While the facilitator's role was to guide the children to solve the problems themselves, this was one of the most challenging aspects of the program for them. Overall, we observed mixed abilities from facilitators to support children without giving them the answers or without doing the exercise themselves.

- **"Facilitators guide the children to resolve their own problems, but sometimes they still go quickly and try to do it herself."**
  — Enumerator observing a session in camp 22

- **"Children ask for help and she gives them instructions as required."**
  — Enumerator observing a session in camp 8E

18. Some quotes have been edited for grammar and clarity

Our TCO tool also had a section for unprompted observations from the enumerators. Among some of these observations, we highlight the following records from the enumerators:

- **"There is no time management. The facilitator needs training. "**
  — camp 8E
- **"Session started before the session time. The children are found to play with tablets at 14:00. Facilitator does not look after all of the children equally."**
  — camp 8E
- **"The morning session started at 9:50 am instead of 9:00 am. They are late due to attending moktab in the morning."**
  — camp 8E
- **"Facilitator should recap at the end that what they have learnt today."**
  — camp 22
- **"Facilitator very interested in conducting sessions and the children love her. The neighbors praised her."**
  — camp 22
- **"They should be more aware of the use and care of tablets and headphones and Facilitators should be trained on session management."**
  — camp 22

From these qualitative assessments and observations, we note that, overall, there was a very mixed level of facilitation implemented by the 25 facilitators. Some were able to stick to the light-touch model and basic care that was the objective of their training, some were frustrated at their inability to actually teach, and some were not fully able to provide the light-touch structure required by the program. Overall, the idea to implement an SEL activity at the beginning and the end of the session showed mixed results: IRC staff did not have the capacity to teach more than one game (The Train Game) to the facilitators, which understandably became repetitive and disengaging for all stakeholders when implemented for 16 weeks in a row.

> "Other schools give food and slippers, but we have tablets."
> — Facilitator

3. Participant experiences

What are the experiences of children, facilitators, caregivers, and community members with the Pop-Up program?

What are the perceptions of children, facilitators, caregivers, and community members about i) the relevance and usefulness of the program and ii) the quality of facilitation and the learning environments?

Caregivers

The caregivers we spoke to placed great value on their children's education and had decision-making power over which programs they sent their children to. When evaluating the whether to enroll their children in the Pop-Up program, caregivers tended to consider four important criteria:

i. The quality of the curriculum and teaching implemented in the program/center,

ii. The respect of cultural and gender norms,

iii. The program's ability to ensure discipline, to be orderly and neat, and

iv. The program's respect of religious practices and norms.

Based on these criteria, caregivers saw great value in the Pop-Up program and believed that it provided a very high-quality education for their children, compared to other programs available in the camps despite some scheduling conflicts with childrens' religious studies.

What is valued by caregivers	How Pop-Up aligned with these values
Children get a quality education	+ Caregivers saw that children were motivated to learn and take notes, and that the curriculum covered English and mathematics. + Caregivers were able to witness the quality of the program, given that it took place close to their homes. + Caregivers noted that children were learning through the tablet, and not just playing games.
Gender norms are respected and women and girls stay close to home	+ Girls were able to learn in a location close to their home, or in their own home. + Mothers were able to observe the program and saw for themselves that cultural gender norms were respected. For example, facilitators ensured that girls and boys were sitting at safe distances from each other. + Facilitators were all young adult females from the community and were able to work for Pop-Up because of the home-based design of the program.
Children learn how to be discipline, orderly, and neat	+ Facilitators enforced basic hygiene rules to protect the tablets, and they made sure that children had clean hands and handled the material properly. + Caregivers were able to witness that children were sitting properly, with space, and that they were quietly learning since each child had an individual tablet and headphones. + Because Pop-Up took place in facilitators' homes, hosts usually cleaned up and prepared for the session and the space was perceived as conducive to learning.
Religious studies are highly respected and prioritized	− When Pop-Up sessions were scheduled in the afternoon, they conflicted with the Madrassas schedule, and children had to either skip Pop-Up or skip their religious studies. − When Pop-Up sessions were only scheduled in the morning, they i) conflicted with programming from other learning centers and ii) failed to maximize usage of the tablets, thus reducing the cost per child.

Below are quotes that provide different views on these themes

On the quality of education, caregivers commented further:

MOTHER

"We tell our friends, 'Yes, your children are learning games, but our children are learning through games. Our children don't even want to leave the school.'"

MOTHER

"Other schools are explaining with black boards and children get bored. Here they listen on tablets and hear something new each time."

MOTHER

"[The other learning centers] give pens, biscuits, slippers, and cold drinks at different centers, but the learning is poor quality."

On respecting gender norms and staying close to home, caregivers, children, and facilitators commented further:

MOTHER

"It's better for Pop-Up to be at home. If it's in the center, it's hard to observe."

CHILD

"My Mom visits my learning session one time each week. Before Pop-up I went to [another] learning center and she never visited because it was too far."

CHILD, WOMEN'S CENTER GROUP

"My parents are happy that I am learning with only girls."

FACILITATOR

"I feel happy because it's at home. If I go outside, it won't be comfortable."

On being disciplined, orderly, and neat, caregivers commented further:

MOTHER

"Our facilitator is doing well and maintaining the gaps between students (where they sit)."

MOTHER

"In this program the children learn deeply, there are no disturbances, and the facilitator supports well."

MOTHER

"It is a clean environment and when we see this it makes us happy."

MOTHER

"Most of the time, they have learned to clean themselves. The facilitators teach them how to be clean."

On the importance of religion, caregivers noted how important it was to ensure that children still have time for religious studies, which happens daily in the afternoon.

MOTHER

"Timing of the [morning] sessions is good because the children can still go to the mosque."

MOTHER

"If they can learn English [from Kitkit School] and Arabic [from the mosque], it will be good because both are important."

MOTHER

"One bad feeling is that in the afternoon, they have to go to the mosque to learn the Quran. It's better if we change the time [of the Pop-Up session]."

Caregivers also noted specific benefits for children using the program: They demonstrated an increase in motivation to learn, an increase in confidence, and new skills related to technology.

MOTHER

"After playing with the tablets we are seeing children's brains opening up. They show an interest in learning."

MOTHER

"Before we would beat them to go to school and they wouldn't. Now they go without us asking."

MOTHER

"Children are more confident about blocks[14] in camp. They can read the camp signs (in English)."

MOTHER

"Before Pop-up children wouldn't go to the market. Now they do - they are becoming brave because they can read and write."

14. Navigating the different neighborhoods in the camps

Facilitators

Facilitators perceived the program as relevant and useful for children in this context. It was consistently described as superior to other learning centers and other education programs in the camp.

"There is no formal education at [other] learning center. It was more like a child friendly space." — *Facilitator*

Facilitators, similarly to caregivers, also valued quiet and orderly learning environments. The fact that children were learning individually with headphones was appreciated.

Facilitators also valued that Pop-Up chose to hire Rohingya community members (as opposed to hiring Bangladeshi teachers as in other centers).

"Learning instruction at the other learning center is poor and there are Bangladeshi instructors [not Rohingya instructors]." — *Facilitator*

Community members

Community leaders perceived the program as highly valuable for the children and their caregivers.

"I decided to welcome the program in my community to make our children educated. In Myanmar, our children couldn't get enough education. So, from this program children can learn something." — *Community Leader*

Across interviews, participants shared the belief that Pop-Up's use of technology made the program special and contributed to the perception that the program was modern and effective. Community leaders we interviewed noted that learning English and math was important for the children, but that learning the Burmese language was also necessary. Additionally, they reported that hiring facilitators from the Rohingya community was a very positive aspect of the program. Community leaders also valued the fact that learning took place in small groups in the houses with interactive materials that were adapted to the needs of the community and the situation of the camp. Overall, we found that the very local, Rohingya-led aspect of the program contributed to its success and wide adoption by the community.

"Children are learning the names of days and months, which is in English. And also they are learning how to calculate which is very helpful for their daily life." — *Community Leader*

"From some families I received feedback that it is a good program that their children can learn very easily on the tablet." — *Community Leader*

"In my experience I can say that children are learning Math and English by playing games on the tablet. And also, I see that there are some changes in children's behavior, like now children are not fighting each other." — *Community Leader*

In conclusion, children, caregivers, facilitators, and community members perceived Pop-Up as a valuable program that brought quality education to the community. They each appreciated the home-based model for its ability to i) enable small-group learning that was visible by caregivers, ii) enable women and girls to stay close to home and iii) enable quiet and clean learning environments, despite the small spaces in each home.

Each of these stakeholders had their own recommendations for improving Pop-Up. We will be integrating these suggestions into the next phase of the project, to ensure it is a high impact, cost-effective program that our clients love to use.

What were facilitators' and community leaders' perceptions about the benefits and challenges of participating in the community of practice?

The IRC recruited 25 facilitators to provide the human facilitation needed for this program. Their role was to ensure that children were getting the right amount of exposure to the tablet, that the tablets were moving from point A to point B, and that learning sessions were happening in the best environment possible (without child protection issues, without disturbance from neighbors and so on). Our 25 facilitators were young female adults from the Rohingya community, with various levels of education and skills. Their average age was 22 years old, and many of them were mothers and older sisters to children in the project. Due to cultural norms within the community, most young women must stay at home and are unlikely to have jobs. For this reason, Pop-Up sessions were designed to run either in the home of the facilitator, in their parents' home, or in a center located in the same sub-block as their home.

> "I have two children in the program. I became a facilitator to help them learn and to help other children learn."
> — Facilitator

Facilitators felt a great sense of ownership and responsibility when their own children, family, or neighbors were in their session. They described their job and basic responsibilities to include distributing the tablets, welcoming and encouraging children, playing games, and recapping the session.

Our training and recruitment indeed focused on these main responsibilities, which were very intentionally limited. We wanted to set up a program that did not depend on talented teachers, and that could be rolled out with light touch human support. As a result, we found that facilitators were fully able to perform their basic responsibilities, yet also became frustrated about their inability to teach, instruct, or help children when they get stuck. They believed that more knowledge and training about how to teach the content would allow them to do their job better.

Facilitators were also able to precisely follow the model that was established during the kickoff training:

"Our program has a session structure that I always follow: 15 minutes play at beginning, 1 hour on the tablet, 15 minutes questions and close out."

However, facilitators also noted that they needed more guidance from the IRC to be able to support children who got stuck in the games, to detect the ones who were guessing or falling behind ,and to ensure that this experience was benefiting all kids, no matter their ages and learning levels.

"If I can distribute the tablet properly, then I've done my job well… but there are a lot of things I don't know. It's important for me to know so that I can teach them." — *Facilitator*

The IRC staff also organized regular meetings with facilitators to provide them with ongoing support and to understand what was working well and what aspects of the program needed to be improved. These meetings were called Communities of Practice (COP) and took place 5 times over the 16 weeks of the pilot in facilitators' homes. Each COP was facilitated by an IRC education staff and focused on ensuring high-quality implementation. The IRC staff would ask facilitators about the challenges faced over the previous two weeks, check in about children's attendance and home visits, and then provide refined instructions on implementation.

Some of these instructions included:

- **Ensure children are playing on the right login ID and complete the login process for the children to ensure data accuracy.**
- **Be welcoming to onlookers and behave politely to maintain good community relations.**
- **Put headphones on children and tell children to keep their tablets on the ground to prevent breakage.**
- **Check children's cleanliness to help create good habits.**

In some cases, the COP also presented the opportunity for facilitators to learn from each other's experiences in different groups, and to exchange tips.

"One facilitator shared that her students were in the counting level, but the instructions were hard. I learned how to instruct the game when the students in my session reached there."
— Facilitator

Overall, we found that the facilitators' perception of IRC's support for ongoing practice—specifically the COP—was mixed. Facilitators desired more support and wanted to advance their skills to become more capable of teaching and helping students. Although understandable, this desire was not aligned with the program's objective of providing a learning experience that was independent from skilled teachers. We will revisit the purpose of the COP, to clarify expectations with the facilitators and to enable them to provide the right amount of guidance to children who need the most support.

What aspects of the learning experience did children like and dislike and which ones were easy and difficult for children to navigate?

Overall, children's experience with the program was highly positive. They enjoyed going to the sessions, arriving early and attending regularly:

"The children come everyday. We don't have to tell them to come. They play and listen and follow instructions." *— Caregiver*

They enjoyed learning mathematics, **"The numbers song is my favorite. It's a song with nice music,"** and hearing the instructions in Rohingya, **"I hear interesting things in Rohingya language."**

Learning through the tablet was perceived as a serious and modern approach to education. Children enjoyed the interactive content and the way that they received instructions.

"My brothers go to school but their education is bad…" "Pop-Up has a computer, it gives instructions for what to do, I can see what I learn and I can play."

Additionally, children enjoyed the flexibility of the program; a non-negligible factor in contexts where children often have to miss school or leave at different hours to support families with work.

> "I don't like that if I'm late to school, they don't allow me to go inside. With Kitkit, if I'm late, I still get the tablet to play."
> — Child

Figure 24 shows endline data, which reflects an overall positive and engaging experience for children. Specifically, we observe that over 95% of children report that attending sessions, watching videos, playing games, and learning with other children were often or always considered fun. Books, stories, videos, and games available on the tablets were also considered interesting in the vast majority of the cases.

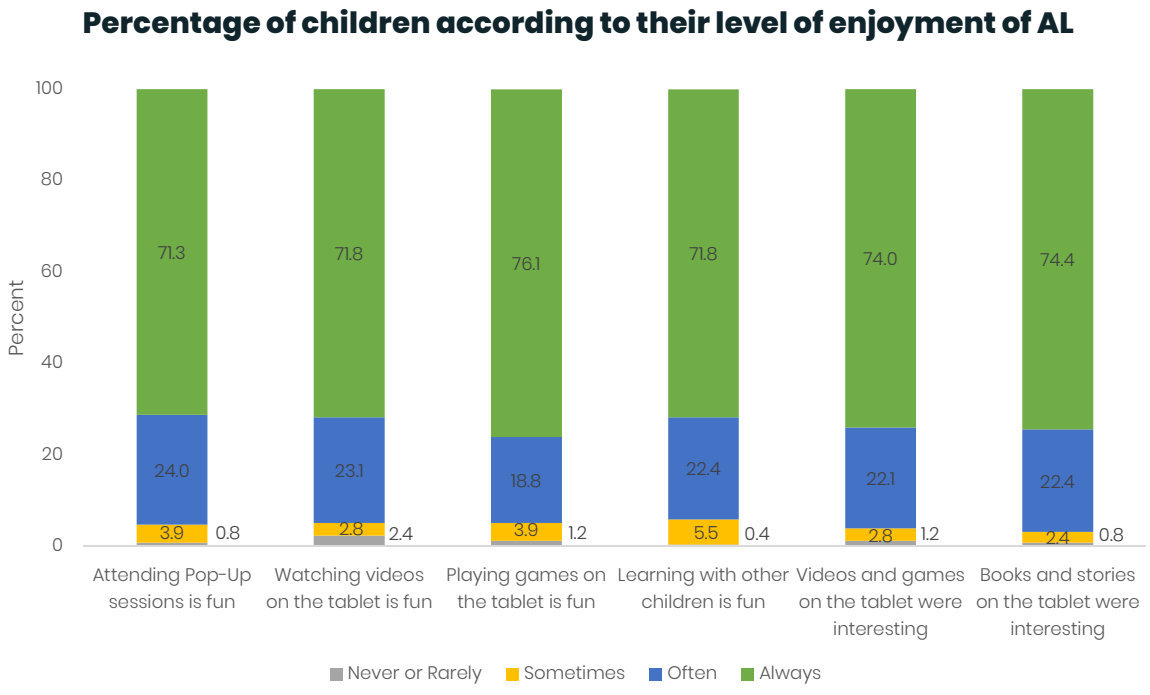


Figure 24

Additionally, children reported that sometimes they experienced difficulties understanding instructions on the tablet or moving through different games and levels. Figure 25 shows the percentage of children who report different levels of understanding how to navigate AL software. We observe that over 90% of participants report that learning on the tablet is easy often or always, they understand Rohingya instructions on the tablet, they know how to move to different games and levels often or always, and know how to navigate the table very well often or always.

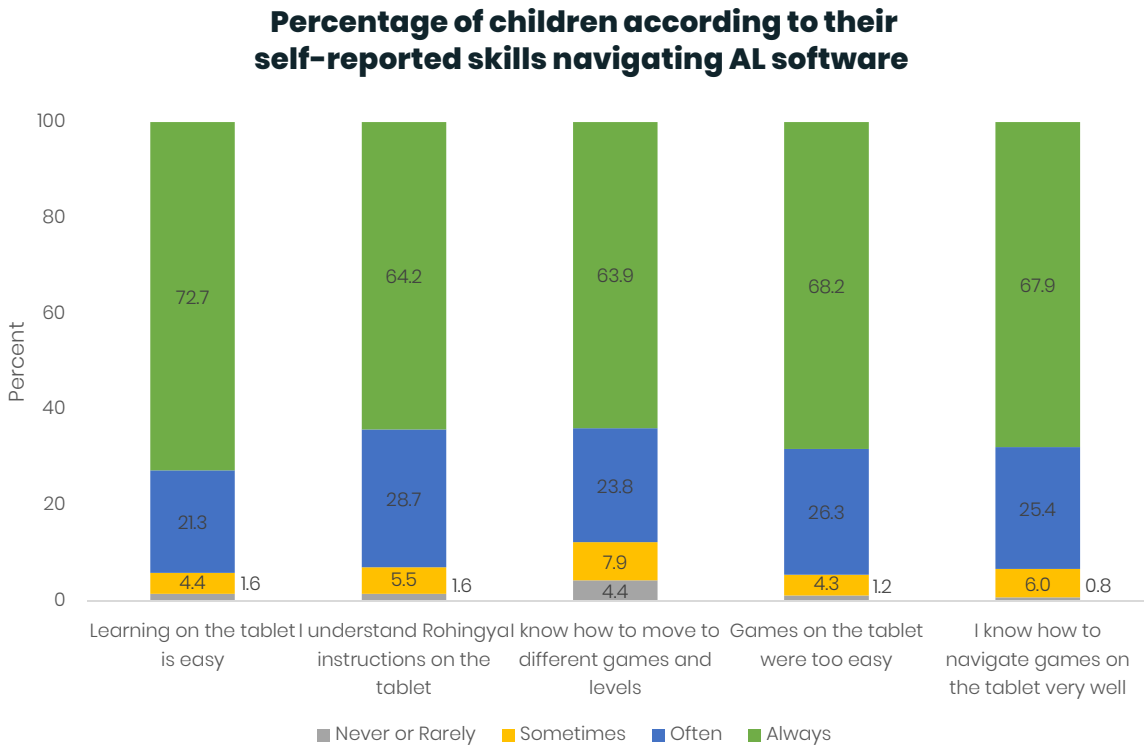
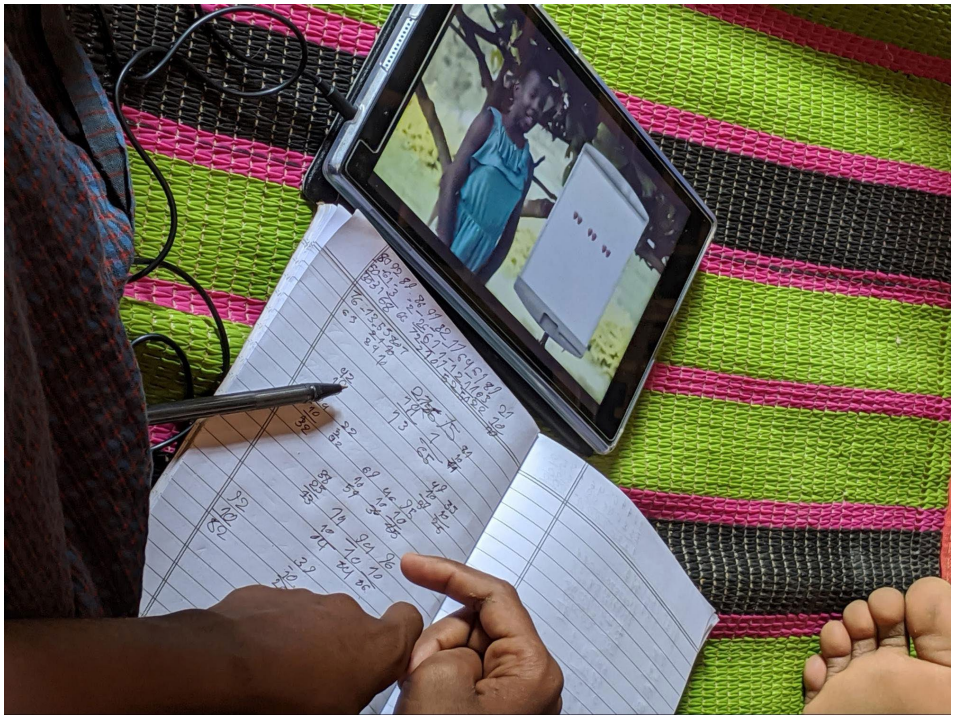


Figure 25

Our observations and user testing sessions with the tablets have pointed out that sometimes the tablet alone was not enough of a learning aid for children. They also used their voices, finger counting, air drawing, or notebooks to help them learn and solve problems. This can be interpreted as a good sign because it shows that children are making efforts to process the activity, get the right result and learn, but it also illustrates the need to provide children with basic supportive materials such as pen and paper, to write down the additions, subtractions, or their newly learned vocabulary.

Finally, we also learned that children found it difficult to sit on the floor with the tablets on their laps. Facilitators, program staff, and other stakeholders have asked for solutions to improve children's posture, but this did not seem to be a major obstacle to participation or attendance.



What were children's perceptions and experiences with different software programs? Which features of each software did children prefer?

Kitkit School

> "Also, during sessions, when one achieves a level, receives an egg, and the egg hatches, the students shout, "I got a snake!" and others are encouraged to also achieve that level so they can get a snake too."
> — Facilitator

Children understood how to use the software and they could interpret what they should do on each screen. All of the children who participated in our testing sessions knew how to navigate the app, how to switch between school and library, how to switch between math and English, and how to enter and exit games.

The initial pilot version of Kitkit School for Rohingya Learners was a compressed prototype or beta version of the software designed to be tested and then adapted according to the needs of the children and context. There were a number of challenges that arose including a login mechanism that allowed children to switch accounts and unlocked egg levels meant to help provide insights on initial learner readiness. As a result, many children accessed levels for which they were not adequately prepared, often guessing or switching between activities or accounts.

- **In some cases, facilitators created a structured schedule in the sessions, to ensure that children would focus and avoid jumping around tasks and topics.**
- **Some children were overwhelmed by the number of options and activities possible, and were observed jumping in and out of activities without completing them fully.**

- **In some groups, younger children (ages 6-7) would spend more time watching videos, while older children (ages 8-10) would spend more time on specific learning games, per the recommendations of their facilitators.**
- **Children who advanced too fast in the higher levels ended up stuck with exercises that were too hard for them and would either keep guessing and inputting random answers or try and revert to lower levels.**

Children understood progression and rewards: they know that eggs and shining hexagons show their level progression and are motivated to unlock levels.

"When one completes a game or a level, he or she yells, 'Hurray!'" – *Facilitator*

Some children were very sequential in completing levels. Other children left levels incomplete before moving on to another and found themselves unable to progress but didn't know why.

"When children go too fast and get stuck at a level, I remind them that they have to play all the old ones in order to open a new level. An open folder should be played." — *Facilitator*

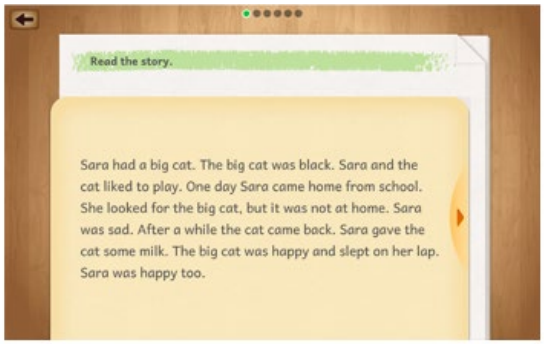
This will be addressed in the updated version of Kitkit School for Rohingya Learners, where the learning path is sequential, less compressed, and more structured.

Kitkit School's most favorable features included:

- When children completed a level and unlocked an egg.
- When children interacted with characters (***"I like the fish"***).
- When children watched and sang along with videos.
- When children heard their own language.
- When children were able to master a game, progress, and get new games.
- When children played with creative tools.

Kitkit School's least favorable features and moments of low engagement included:

- When the level was too advanced and children could not guess or receive help from a facilitator.
- When children saw or read English sentences and books that they did not understand.



Can't Wait to Learn

Children were successful in navigating the game, however, they did not always understand the map and the structure of the different levels. Children who participated in our testing sessions knew that the different houses indicated different levels, but were not sure if they had completed them or not. Some of these children mentioned it did not matter to them, because the level they worked on was automatically selected by the app.

Children liked the videos and understood that they were markers for progression, but did not always recognize other markers. Some children knew that the stars earned indicated progress, other children believed that they finished a level when they saw a new video.

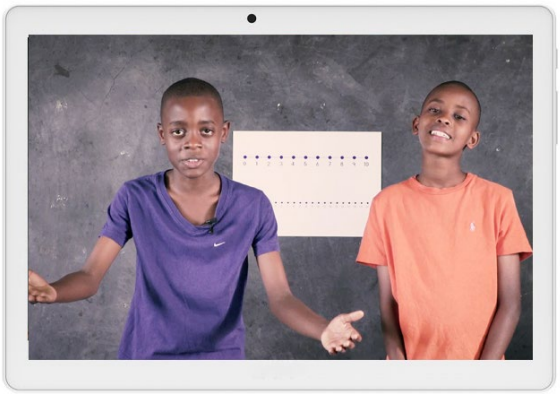
All children knew however when they got the answers right or wrong, thanks to the easy interface with smiley and non-smiley emoticons.

Overall, because Can't Wait to Learn is a structured curriculum with one activity at a time that children must move through sequentially, we found that children spent more time thinking and less time guessing. The games presented few opportunities for taking shortcuts, guessing, or repeating activities.

When children encountered difficulties or high repetition, they learned how to exit the game in order to open a new one to avoid the previous game they were on, at least temporarily. This functionality will be removed in the next version of the game and Can't Wait to Learn is considering the option to enable children to opt out and do something different, in case they feel too frustrated.

Can't Wait to Learn's most favorable features included:

- When children unlocked a video as a reward for completing a level.
- When children were given a mix of audio + visual tasks.
- When children were able to master a game and progress.
- When children heard their own language.



Can't Wait to Learn's least favorable features and moments of low engagement included:

- When the videos were too long.
- When the game was hard and children were not allowed to pass by guessing.
- When children mastered concepts quickly and encountered high repetition.

In conclusion, children enjoyed both learning software programs but had a different experience with each. While we can highlight features that they particularly enjoyed or that need improvement, it is too early to make conclusions about the child experience using Kitkit School, given the unique conditions around this software for the pilot (compressed curriculum and unlocked levels) as well as a host of operational and technical challenges that affected one of the camps disproportionately. While most children seemed to understand instructions given with the Rohingya voice-over, further rigorous research will be needed to understand if children are effectively reaching foundational skills in English and numeracy.

The IRC is interested in continuing to work with software partners to improve each application and ensure that products are suited for these challenging crisis contexts.

4. Cost efficiency

What was the average cost per child during this pilot and how can it be optimized? What are the key cost drivers per child?

A retrospective cost analysis of actual costs incurred during the pilot program from August 1, 2019 to February 29, 2020 was conducted to better understand the cost per child associated with implementing the pilot program and key costs drivers. These figures exclude all research and design costs, including those that were incurred during the period above.

The initial costing analysis included disaggregating costs by delivery model (home and center) and software (Kitkit School and Can't Wait to Learn). However, in practice, there was little difference in the center and the home delivery model. Group sizes did not differ as much as we expected and the ratio of facilitators to students was nearly identical. For these reasons we focused the retrospective analysis on disaggregating by software and understanding the implications of removing localization and headquarter (HQ) costs.

Cost per child

Cost/Child (All Costs)	Cost / Child (By Software)	Cost/Child (No Localization)	Cost/Child (No Localization + No HQ)
$575 **Including software localization for both Kitkit School and Can't Wait to Learn**	**$723** Kitkit School numeracy and literacy	**$284**	**$256**
	$412 Can't Wait to Learn numeracy		

The cost per child including both Kitkit School and Can't Wait to Learn was $575 per child for the initial pilot study and early-stage investments, including software localization, equipment, infrastructure, and administrative costs. The cost per child by software was $723 per child for Kitkit School (numeracy and literacy) and $412 per child for Can't Wait to Learn (numeracy). Localization of the software comprised 51% of the total budget (40% Kitkit School, 11% Can't Wait to Learn). Localization is a one-time cost incurred when launching the software in a new language.

The cost per child excluding localization was **$284 per child**. In this model, equipment—including the hardware, i.e. tablets and other tech set-up materials—made up the majority of the budget at 36%, followed by national staff (32%) and country operating and support costs (16%). Similar to localization, equipment is a one-time cost incurred when launching the program, and tablets can be used for multiple years across multiple cohorts of children thus improving the cost per child over time.

Costs of a new pilot program differ from those of an established IRC program and we expect to see cost efficiencies once the program is tested and established. Among these cost efficiencies include cutting HQ costs. In this scenario, we would expect a **cost per child of $256**. Additional efficiencies are also being explored to reduce this cost further including storing tablets within camp thereby reducing vehicle and fuel costs and increasing the number of sessions per day thereby increasing the number of children who share the same tablet. Future iterations of the program will also capitalize on scale efficiencies as we grow to reach more children.

In the future we also expect some costs to increase including the cost for tablets as 90 of 210 tablets used for this pilot program were donated.

Initial scale projections estimate a **cost per child of $151** when reaching 32,000 children and excluding localization costs and HQ costs which will not be incurred during the scale phase of this project.

Major Cost Drivers

Major Cost Drivers by Percent of Total Spend (All Costs)		Major Cost Drivers by Percent of Total Spend (No Localization)		Major Cost Drivers by Percent of Total Spend (No Localization + No HQ)	
HQ Costs (Staff, Travel, Equipment, Supplies)	5%	HQ Costs (Staff, Travel, Equipment, Supplies)	10%	-	-
Localization Kitkit School 32% literacy and numeracy Can't Wait to Learn 9% numeracy	51%	-	-	-	-
National Staff	16%	National Staff	32%	National Staff	35%
National Travel	1%	National Travel	2%	National Travel	2%
Equipment Tablets 7% Vehicle and Fuel 4% Import cost for donated tablets 2% Tech set-up 2% Other 3%	18%	Equipment Tablets 15% Vehicle and Fuel 8% Import cost for donated tablets 5% Tech set-up 4% Other 4%	36%	Equipment Tablets 17% Vehicle and Fuel 9% Import cost for donated tablets 5% Tech set-up 4% Other 6%	39%
Supplies Tablet storage 1% Program facilitators pay 1% Other 1%	3%	Supplies Tablet storage 3% Program facilitators pay 2%	5%	Supplies Tablet storage 3% Program facilitators pay 2% Other 1%	6%
Country Operations and Support Costs	8%	Country Operations and Support Costs	16%	Country Operations and Support Costs	18%


brac

6.0 Limitations

This pilot study was not designed to document the impact of AL software or different delivery models on children's learning, but to confirm that AL is a feasible solution with the potential to be cost-efficient and cost-effective. The findings that we document, with regards to learning outcomes by software and delivery mode, need to be interpreted with great caution due to different limitations. The baseline-endline changes in children's literacy, numeracy and SEL should not be interpreted as impact, because our pilot study did not use a control group, and did not randomly assign children to different treatment groups, and so there are many reasons that can explain the changes observed other than the intervention. We discourage the reader from comparing software packages or delivery modes, because the baseline samples included in the present study were not only small and not powered to conduct comparisons, but also, because COVID-19 disrupted data collection activities at endline, further reducing our sample sizes.

Additionally, the present study used the ASER tool to confirm positive baseline-endline changes in children's literacy and numeracy outcomes, but ASER does not adequately capture learning gains in a fine-grained way for the purpose of identifying impact. Future studies will need to document the effect that AL software and different delivery models have on children's literacy, numeracy and social-emotional skills, using a research design that includes a control group to capture impact, with an adequately powered sample to conduct comparisons between treatment and control group, and with tools with evidence of validity and reliability for the Rohingya population, that are fit for purpose in the context of an impact evaluation.


17+32=
14-7=1

7.0

Conclusions & Recommendations

The results of the Pop-Up Learning pilot confirm that autonomous learning software is a promising avenue for alternative education in crisis settings and requires further investment into nimble infrastructure, software features, and rigorous research to better understand program results and to be deployable quickly at the onset of a crisis, in a more cost-effective manner. The pilot helped us to obtain preliminary evidence to support hypotheses with regard to different dimensions of the program.

1. With regard to learning, we found preliminary evidence supporting the hypothesis that:

Displaced children who are out of school can acquire foundational academic skills and SEL skills through tablet-based AL.

Children who had access to numeracy games through both software programs—Can't Wait to Learn and Kitkit School—showed positive baseline-endline changes in ASER numeracy scores. Children who had access to literacy games using Kitkit School (children in Can't Wait to Learn did not have access to the literacy curriculum) showed improvement in baseline-endline ASER literacy scores. Children using both software programs also showed positive changes in their hope and agency skills. Students in both home-based sites showed baseline-endline improvements in their reading, numeracy, and SEL skills. At endline, we were not able to collect data from children using Kitkit School in center-based sites, but we observed improvements in the numeracy and SEL outcomes of children using Can't Wait to Learn.

2. With regard to implementation, we found evidence to support the hypothesis that:

Displaced out-of-school children are able to attend home- and center-based learning sites.

When sessions are scheduled in line with cultural norms and values, attendance was very high. Community members specifically appreciated a program that empowered local facilitators, that enabled girls to learn at home, and that was not scheduled at the same time as religious studies. Future research will need to collect and document more reliable data on dosage.

Tablet-based interactive content can keep children engaged while learning foundational academic and SEL skills.

Both software programs are designed in ways that keep children engaged and progressing autonomously.

Low-skilled caregivers or community members are able to conduct AL sessions.

Despite that the majority of facilitators exhibited only emerging levels of proficiency with regard to their facilitation skills, children attended Pop-Up and were able to show improvements in their learning. Facilitators were able to provide a clean environment and ensured children took good care of the hardware.

3. With regard to participants' experiences with Pop-Up, we found evidence that:

Children, facilitators, caregivers, and community members consider Pop-Up as a relevant, valuable, modern, and effective program that improves the learning of displaced children.

Caregivers we spoke to placed great value on their children's education and had decision-making power over which programs they sent their children to. When assessing the Pop-Up program, they considered different criteria such as the quality of the curriculum and the teaching implemented in the program orcenter, the respect of cultural and gender norms, the program's ability to ensure discipline, to be orderly, and neat, and the program's respect of religious practices and norms. With these criteria in mind, caregivers highly valued the Pop-Up program and believed that it provided a very high-quality education for their children, compared to other programs available in the camps. In general, stakeholders appreciated the home-based model of implementation.

Caregivers, family members, or low-skilled community members with low-to-no literacy skills can facilitate informal learning for individual children and groups and promote the use of, and engagement with, tablet-based experiences.

The program recruited 25 facilitators who were young female adults from the Rohingya community, with various levels of education and skills. Their average age was 22 years old, and many of them were mothers and older sisters to children in the project. Facilitators were fully able to perform their basic responsibilities, yet also became frustrated about their inability to teach, instruct, or help children when they get stuck. They believed that more knowledge and training about how to teach the content would allow them to do their job better.

Caregivers can be recruited as facilitators through existing touchpoints in the system (e.g distribution or registration center).

The 25 facilitators were recruited through existing community touchpoints and with support of the community leaders. It was not a challenging task for IRC staff.

Children can safely learn in temporary homes and centers.

We found evidence that the quality of the environment in the center-based and home-based settings was conducive to learning.

Caregivers can be held accountable and be compensated for their work.

Facilitators felt a great sense of ownership and responsibility when their own children, family, or neighbors were in their session. The COP ensured facilitators felt supported and could be held accountable. Facilitators could also be compensated for their work and the community and facilitator's family were highly supportive of the facilitators working for pay from their own homes.

Aid organizations can enable adaptive learning through tablets shared by children.

However, aid organizations must be more nimble on supply chain and storage solutions in order for AL to have the potential of being a cost-effective solution.

We did not find evidence that tablets can be safely stored and preserved in the camp.

The IRC did not take the risk to store the tablets in the camp for this pilot. We are unable to confirm the hypothesis that tablets can be securely deployed close to clients.

Localization into niche dialects and languages can be much faster and cheaper than current practice; new workflows can enable timely localization.

The IRC localized two software programs independently, in nimble and innovative ways. These practices were new to many members of the organization and can be improved for the future. Leveraging the diaspora of the displaced Rohingya community was challenging and ineffective.

4. With regard to the cost of Pop-Up, we found that:

The cost per child including both Kitkit School and Can't Wait to Learn was $575 per child for the initial pilot study and early-stage investments, including software localization, equipment, infrastructure, and administrative costs. The cost per child by software was $723 per child for Kitkit School (numeracy and literacy) and $412 per child for Can't Wait to Learn (numeracy).. Localization of the software comprised 51% of the total budget (40% Kitkit School; 11% Can't Wait to Learn). The cost per child excluding localization was **$284 per child**. In this model, equipment, including the hardware (tablets and other tech set-up), made up the majority of the budget at 36%, followed by national staff (32%) and country operating and support costs (16%). Similar to localization, equipment is a one-time cost incurred when launching the program, and tablets can be used for multiple years across multiple cohorts of children thus improving the cost per child over time.

Costs of a new pilot program differ from those of an established IRC program and we expect to see cost efficiencies once the program is tested and established.

Initial scale projections estimate a **cost per child of $151** when reaching 32,000 children excluding localization costs and HQ costs which will not be incurred during the scale phase of a project.

Below are 15 recommendations and areas to focus on for our next phase of work:

For program implementation:

1. **Store tablets near learning spaces** (whether they are homes or centers) and ensure that the infrastructure is more nimble, despite the security risks. Storing tablets in the most secure location — while mitigating some risk — resulted in inefficient logistics and higher cost, which can be avoided with alternative storage and charging solutions.
2. **Invest in charging systems that are adapted to local constraints.** Invest in various charging stations and materials, such as individual solar panels, power banks, and so on, to enable alternative power solutions for the community. For example, facilitators or households could keep the tablets for several days and maximize their usage if they had the ability to charge them.
3. **Increase facilitator model reach.** An empowered team of community facilitators that values the program is one of the best investments we could make. Working with mothers, sisters, and women from the community was a strength, and we can optimize their reach by increasing the number of sessions they run or the number of students they supervise.
4. **Clarify facilitator roles and positioning in the community.** The dependable human support that facilitators brought is essential, but the role and expectations could be clarified and communicated to community members more effectively.
5. **Test the value of workbooks for blended learning models.** Consider testing the use of paper workbooks to support's children's interactions on the tablet and to make learning tangible for other community stakeholders.
6. **Structure sessions to include SEL activities and stretching breaks.** Sitting on the floor with tablets was at times uncomfortable, so consider specific stretching breaks and playful activities to support children's attention span. This could also be prompted by the software.

For software and product design:

7. **Ensure children are guided through the curriculum.** Given the various learning levels of children served by the program, using a structured and guided curriculum will help with engagement and learning.

8. **Improve First Time Use (FTU) experience and optimize for the absence of training.** Software providers can include stories and various demonstrations through the tablet, to avoid the need for explanations. Children should understand the game structure, the mechanisms for unlocking various levels, and the ways to navigate its features without the need for facilitators to explain.

9. **Include ways for children to ask for help in the software.** When we design programs for the lightest human support possible and the most autonomous experience, it is important to consider ways for children to express their confusion or frustration within the game. Software providers should include a way for children to ask for help and to show that they are blocked.

10. **Propose different ways of learning a concept in the software.** The software could propose alternative methods and activities for children who are struggling with a concept. Currently, options are limited and the child can only go back to the previous level.

11. **Integrate more adaptive features to tailor content to children's level.** Software providers can focus their investment on embedded learning assessments to ensure children learn at the right level.

For research:

12. **Improve tablet analytics and work towards standardized metrics.** In contexts where research is challenging and expensive, implementing organizations should be able to rely on data coming directly from the tablets, as opposed to surveys or other self-reports. Software partners should work together towards standardized metrics. For research purposes, it is also highly important to ensure that the program and technology support tracking the activity of unique users.

13. **Use valid and reliable tools of learning that are fit for purpose to measure learning gains.** In order to detect progress in participants' learning skills, researchers should use valid and reliable assessments that are fit for purpose, and more adequately capture changes in students' learning outcomes. Research efforts will obtain more precise estimates of learning progress with tools such as EGRA and EGMA. Ensure that these tools are localized and enumerators are well trained in their administration and that researchers collect evidence of validity and reliability for the Rohingya population.

14. **Conduct further operational research to test feasible ways to store and distribute tablets and to deliver the program safely during the COVID-19 crisis.** It is imperative to adapt the existing program to respond to the growing educational needs of children in the time of COVID-19. Operational research should be conducted to test the feasibility of adapted implementation models while ensuring the safety of participants and IRC staff.

15. **Build the evidence about what works to implement AL on children's learning and SEL outcomes and about cost-effective ways to implement AL at scale.** After refining the model based on the evidence from the pilot, conduct rigorous research that will include an initial phase of design and implementation research to identify more cost-efficient ways of storing and distributing tablets, and confirm that all ingredients of the program are being implemented as expected. Then conduct a randomized controlled trial to test the cost-effectiveness of different implementation modalities of the program. Given the specific complexities of crisis settings, we need a better understanding of how autonomous learning can work for children at all levels, including those who