**Title Full Paper**

# Data Collection for Literacy Tablet Reading Project in Rural Ethiopia

**Angela Chang[1], Lidet Tilahun[2], David Nunez[1], Cynthia Breazeal[1]**

[1] MIT Media Lab, Massachusetts Institute of Technology. Cambridge, MA
[2] One Laptop per Child, Addis Ababa, Ethiopia

**Presenting author's email address:** anjchang@mit.edu, LidetT@africa-union.org,

**Session Number:** TH3-SE03-02

**Keywords:** literacy, big data, tablets, rural, education

**Brief biography of presenting author:**

Angela Chang researches how multisensory interactivity improves literacy. Her postdoctoral research was to develop custom literacy applications for children.

Lidet Tilahun directs One Laptop per Child's International Outreach. She advocates for innovative learning to influence large-scale change at the country level and on the global development agenda. Lidet, managed the "Learning to Read" project in two remote villages in Ethiopia testing whether children can teach themselves how to read using only tablets filled with applications designed for self-learning

**Abstract:**

This paper describes the data collection solution employed in the deployment of `literacy tablets" to children in two remote rural villages in Ethiopia. To conduct this research, touchscreen tablets preloaded with educational application were used as educational delivery mechanisms for literacy.

This work aims to investigate how children form collaborative "learning cells" using technology designed especially for learning. This project fosters interactions between children by establishing a supportive environment where literacy learning can flourish and providing a learning platform. Those two conditions are most lacking in remote rural areas.

We detail the data collection software used in a year-long field test. Our modified Android operating system and the data collection framework allowed us to configure data collection probes for privacy and secure backups. The tablet software records usage data that is used to assess and iteratively adapt the reading platform to a child's developing educational trajectory.

We report the educational interests and usability issues encountered. We correlated our findings with primary observations from field reports to the site. We discuss the difficulty of assessing long-term learning trajectories based solely on data, and discuss our reliance on observations on the socio-cultural use of the tablets. Finally, we reflect on how these lessons will influence future deployments, by focusing on potential opportunities for remote data collection in rural terrain. We also discuss the implications for unsupervised teaching through such technology regarding developing digital learning tools for these remote populations.

**Text Full Paper**

## 1 Introduction and Purpose

Millions of children never learn to read, it is estimated that around 100 million children live in poor, remote areas where there is no access to schools and where everyone around them is illiterate. There are at least another 100 million children who live where schooling is so inadequate that they too, fail to learn to read. Few of these children will ever become contributing citizens of our world.

There are and always will be, places in every country in the world where good schools will not exist and good teachers will not want to go. Teacher training is, therefore, not a sufficient solution to this problem, particularly in the most remote areas. We take a fundamentally different approach to this issue. Our team combined scientific knowledge about how children learn to read with new and evolving technological capacities to create a *literacy tablet* prototype (Chang et al. 2013). These tablets are designed to help remotely located children who have no schooling acquire literacy to further UN Millennium goal #2, *universal primary education for all*.

We developed a solution that provides a supportive and engaging learning environment with four critical dimensions:

1. Immersive in all aspects of a child's life;
2. Engaging children with the early basic principles of learning to read;
3. By rewarding and challenging them to continue to develop as creative and critical thinkers;
4. Supporting how children naturally learn from each other by providing the opportunity for organizing themselves into a community of learners.

The research project in Ethiopia was launched in collaboration with the Ministry of Education; Ministry of Information Communication and Technology of the Federal Democratic Republic of Ethiopia, the African Union and United Nations Economic Commission for Africa. Parts of this research was financed by One Laptop Per Child (OLPC) and National Science Foundation (NSF).

## 2 Design and Methods

Modern mobile technologies have previously been used in novel ways to deliver education and information to rural areas (Kumar et al. 2010, Cervantes et al. 2011, Mills-Tettey et al. 2009, Wesolowski et al. 2012). We present a proof of concept of a *literacy tablet* (touchscreen tablets as a new method of educational delivery). We detail our process of using locally sourced infrastructure solutions and data based feedback to minimize outside intervention by (and exposure to) foreigners into an isolated community (Mitra & Rana 2001)).

*Process in Ethiopia:* From February 2011, the program took two months to launch due to administrative and bureaucratic hurdles. In the end, we secured government support at all levels and enlisted the support of the United Nations ECA (UNECA) and the African Union. Representatives from the Ministry of Education (MOE), the Ministry of Information & Communications Technology, the African Union and UNECA had visited the site prior to our launch. Bringing everyone together in just two months was challenging because we are not a registered and staffed non-governmental organization (NGO).

The Ethiopian Government officials pointed us to two remote areas where they have not yet been able to open a school due to the terrain and the low number of school aged children in the area. The MOE agreed that doing this experiment in English was the best possible scenario, because English is the language of instruction.

*Process in US:* The team of researchers consisted of application designers and human computer interaction engineers from Massachusetts Institute of Technology (MIT), neuroscientist specializing in reading from Tufts University Center for Reading Research, and engineers from One Laptop per Child (OLPC). OLPC was in charge of technology deployment, while Tufts was in charge of the literacy content. MIT was responsible for the implementation of the remote data science aspects.

In September 2010, the US team began prototyping the *literacy tablet* to provide a learning experience that would scaffold a child's learning through applications, games, and video material (Ananian et al 2012, Alonso et al 2011). The intent was to expose learners to English words, stories, letters and letter-sounds. As users experienced the tablet, it would track their usage and the data would be reviewed by researchers to develop new material for the tablet. This feedback loop would enable learners to find fresh content that built on prior concepts. Eventually, the system would help children progress from exposure to the ABCs toward eventual mastery of English (writing their own thoughts).

The first prototype, finished in January 2011, consisted of readily available literacy applications. The Tufts team had examined commercially available literacy apps that targeted developmental goals. The result was a curated assortment of 322 children's apps and a custom literacy app called TinkRBook (Chang & Breazeal 2011). The TinkRBook app developed by MIT was proven to help very young children begin literacy by prompting collaborative exploration (Chang et al. 2012). The other apps were children's books, songs with lyrics, video shorts with subtitles, games and activities that the team approved as culturally and pedagogically sound.

## 2.1 Study overview

Interested in providing their children with access to this new technology, the villagers had voluntarily built a communal hut for collecting and charging the tablets (Figure 1). The hut also served as a place for the kids to congregate. OLPC assisted by providing solar panels, (available from Addis Ababa) and building instructions. To protect the tablets from the dusty environment, leather protective covers were locally sourced by a craftsman in Addis. One local adult was voluntarily trained to help collect the tablets daily, charge them, and redistribute them back to the children. The community buy-in process was celebrated with a village gathering where the OLPC liaison distributed the tablets.

**Figure 1a.Solar charging hut built by community** **Figure 1b) Group learning, image from tablet anonymized**

In order to make sure they understood that the tablet contained learning materials and that children may learn by using it, we spent three days meeting with the parents and community leaders explaining the purpose of the project and how important it is to allow children to use the devices. The Ministry of Education representative read and explained the consent forms before the parents signed them. We made sure each parent signed the consent form before we handed the children the tablets.

Because the children have never been in a classroom setting or organized playground, we had to make sure they understood that they were welcome to come every day to play and experiment with the literacy tablets, and also charge them at the solar station. Minimal instruction was given on operation of the literacy tablets (in kindred spirit with Mitra et al. 2004). The villagers understood that the tablets were for the children, but not uniquely assigned to each child. The children received brief instructions on returning the units to the hut for charging. Thereafter, the children had the freedom to explore the tablets at will.

Figure 2 depicts the system overview for data collection. Each week, technicians from Addis Ababa (1 hour drive away), travel to the site to take out the SD cards and FedEx the data overseas for analysis. For the data analysis team, there is a 3 week lag between deployment and data return. This field logistics team consisted of two students from the University of Addis Ababa who were studying computer science. The data collection occurred for 43 weeks in Wonchi and 39 weeks in Wolonchete, from Jan. 2012 to Feb. 2013.

Wonchi is a village high in the mountains, while Wolonchete is located in a semi-desert region. Both villages are located approximately one hour drive from Addis Ababa. The first site, Wonchi, is a rural village west of Addis Ababa located in the highlands atop a crater lake. The second site, Wolonchete, is located in a semi-desert in the heart of the Great Rift Valley. Each village contained 20 families where no one in the village could read. Due to the geographical location of each village, the children (ages 4-11) are unable to have access to education. Most children stay home until age 11 or 13 years of age until they are old enough to walk to the nearest town to attend primary school. However, because of the hazards in traveling, and the requirements of the farming schedule, older children rarely attended school. The adults in these villages happened to have no education, and farmed most of the day. Except for having no schooling, the children are well cared for in these peaceful villages run by village elders

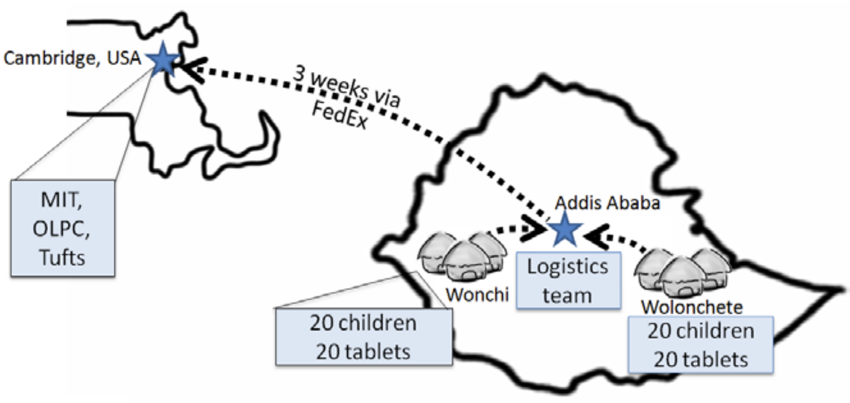


**Figure 2.Data collection system diagram**

## 2.2 Data Collection Details

Our custom interface ran on Motorola Xoom tablets running the Android 3.x Honeycomb operating system. The stock software on the devices was heavily modified to allow for the data collection and security requirements for our experiments; the team engaged the community of Android enthusiasts to help determine the most efficient way to access the deepest levels of the operating system via a process of unlocking and reconfiguring the operating system (i.e. "rooting the tablet"). We required this level of access to enable full-time data collection, allow system level file manipulation, and engage safety and lockdown protocols to prevent users from accessing functionality of the tablets outside of the scope of this project. Figure 3 shows the tablet architecture.

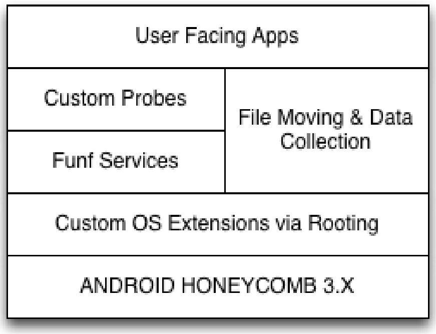


**Figure 3. Tablet architecture**

On top of the Android operating system, the team built a large collection of infrastructure software to support the deployments of tablets to test sites. The system supports our process for remote data collection and transmission to our data server, allows the team to configure data collection probes remotely, makes secure and encrypted backup files for our data, and enables users to switch privacy settings on and off (Aharony et al. 2011).

For the basic data collection infrastructure for our deployment, we used the Funf Open Sensing Framework (Funf 2014), an extensible sensing and data processing framework for mobile devices developed by our colleagues at the MIT Media Lab. Funf provides support software on which we built our customized data probes; it provides templates for writing software which can query the device sensors, log user activity, and record device status.

For our deployments we can collect data from several probes using new software we developed:

- Time spent in tablet apps - a timestamped log entry when a user opens an app and periodically (every 30s) note a log of all running apps. From these data points, we can infer how much time users are spending in given apps.
- Periodic device information - data about battery levels, temperature of the device, GPS location, and screen activity

- Camera - the front-facing camera can record a snapshot of the user either on a periodic basis or on given events (ex. an app launches)
- Microphone - audio from the tablets' microphones collected as snippets for further processing to learn more about what users are saying as they are using apps.
- TinkRBook – a custom literacy app to collect usage data at a high fidelity. For example, we can log every time a user places or removes a finger on the screen and every time the user changes pages. We can also track the exposures to conceptual nodes; for example, we know how many times and how often the user is exposed to a word in written or spoken form, including comprehension probe words relating to colors and numbers.

Additionally, we wrote some support code for enabling remote data collection, albeit with high latency. Funf data is stored in encrypted SQL Lite databases on the tablets internal storage. Our software periodically then copied these databases to the external Secure Digital (SD) cards we installed in each device.

On a weekly basis, the logistics team in Ethiopia would visit the village sites to exchange the external SD Cards with empty cards and return to a central office to process this data. The team was instructed to copy the databases from each SD Card to a highly structured filesystem on a computer. This data was then copied onto a single external hard drive which was then shipped back to Massachusetts for further processing. Due to the number of tablets and the week-long delays between site visits, we would often receive hundreds of database files on each shipment. We used a collection of batch processing scripts to merge, decrypt, and scrub the data for analysis. This yielded a collection of data in JSON format that we could directly analyze.

While visiting sites, the field team could also complete simple upgrades to the tablets via scripts and content bundles we provided remotely. For example, we were able to patch the system software to eliminate bugs or upgrade the literacy software to provide new content.

For security and tablet stability, we locked down access to configurations of the tablet to prevent users from doing things like accessing the data collection software or system settings.

# 3 Results

## 3.1 Hardware data

Initially, the hardware data was the most studied due to issues initial issues with data collection. We noticed that some of the tablets would completely deplete battery life; this caused the device clock to reset to a non-valid date. Using the built in Global Positioning System (GPS) hardware in the tablet which receives multiple satellite transmissions of the current time to triangulate the position of the tablet, we automatically reset abhorrent device timestamps to maintain data continuity.

In the laboratory, an engineer would subsequently charge up the battery when the battery was depleted. The children, however, have no concept of battery life. When the tablets automatically shut off because the batteries are running low, the children would keep trying to turn it on. After 4 tries, the tablet clock would reset to Jan. 1970, and all of the future data timestamps would be invalid. From only looking at the data, it was impossible to tell what was happening. Eventually, an OLPC engineer flew on site to watch the children and he realized that the children had no idea why the tablet would turn off. Without the concept of battery life, they would keep trying to turn it on. The on-site visit also confirmed that the solar hut was working correctly, and that the process to charge the batteries at night worked as expected. In order to extend the battery life even more, the camera probe was turned off whenever the image was too dark. The changes were made upon discovering thousands of blank pictures taken by the camera when the cover was closed, or in extreme darkness.

## 3.2 Tablet usage

Due to the launcher app, we initially could tell all the apps that were opened by the children every week (figure 4). Every day for the first month, they opened every one of the 323 apps.
Initially, their favorite applications were related to learning. Our data shows that they most frequently opened phonics apps, movies relating to African culture (Tinga Tinga Tales) and the custom TinkRBook application (figure 5)

Because of the battery reset error resulting in unreliable timestamps, it was impossible to determine the total time spent on the apps reliably in the first month. We were able to make some reasonable assumptions about the corrupted data to track which apps the children opened. They spent approximately all their free time (6 hours) playing with the tablets. From our logs, someone is likely to be using a tablet at any hour of the day. Over the months, tablet activity decreased, probably as the novelty effect wore off. However, there was still intensive engagement. On average, they opened approximately 50 apps per day, and played the apps they liked for longer periods. The children still ran down the batteries and needed to charge them every day.

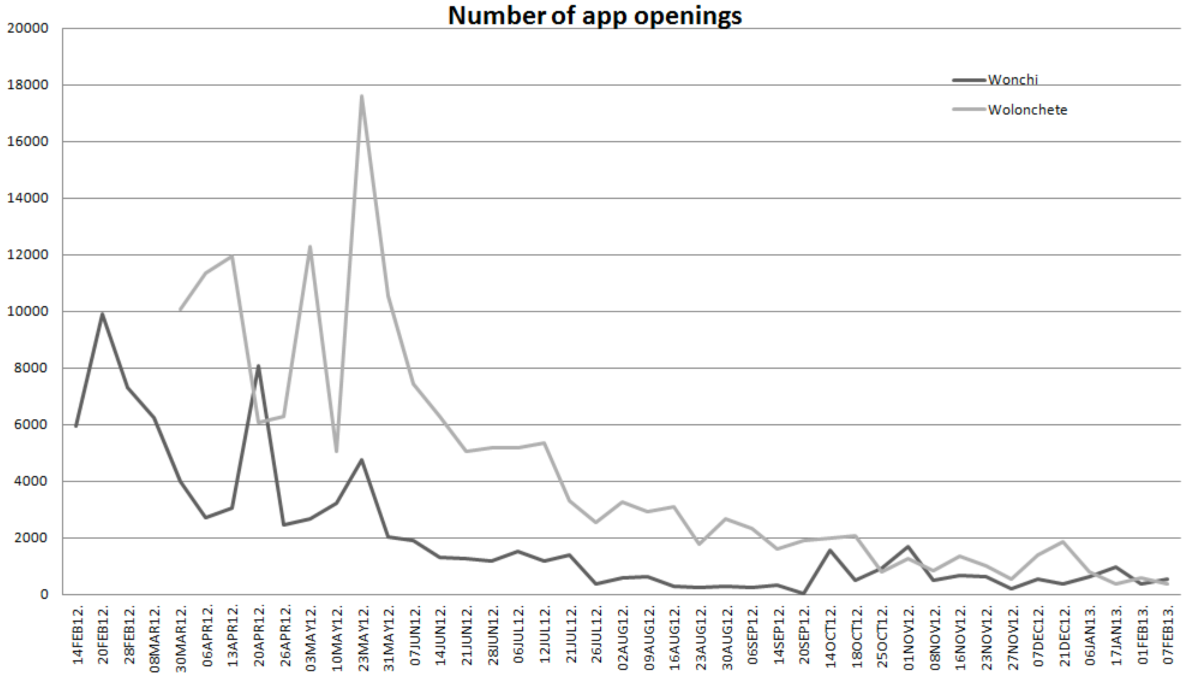


**Figure 4. Number of app openings every week**

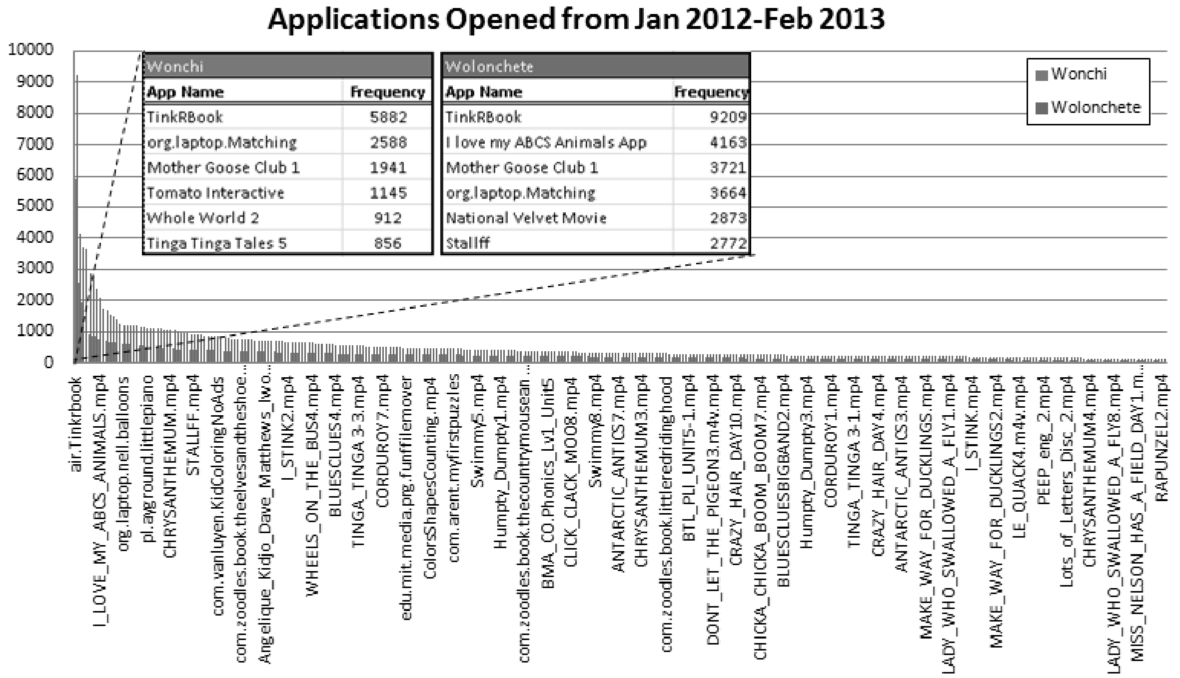


**Figure 5. App openings rated by popularity**

## 3.3 Application usage

From within the TinkRBook, we could tell what items users were most comfortable interacting with. We could tell which activities they liked most within the scenes of TinkRBook. Figure 6 shows that the children preferred playing with color mixing, then sound creation, then counting. They preferred playing verbs in the TinkRBook app least of all.

From this data, we are able to change and create new apps, and install them when the logistics team arrived weekly to download the data from the tablets. The lag between analyzing data and creating a new app was long. It took approximately a month to develop and update the tablets. One app was a color matching app that tested whether children recognized the images of their favorite words. Another app gave pronunciations of the words from TinkRBook. The children responded enthusiastically to any new apps installed.

We could tell what words the children used in their story. For example, in week 2 (Figure 7), the children changed the colors of the duck character to orange. In their TinkRBook, to create an orange duck, users must mix red and yellow within the application. Similarly they had to mix the primary colors, in order to see the words “green, purple, orange, and brown within the narrative text.  Thus, their usage behavior was correlated with learning new words.

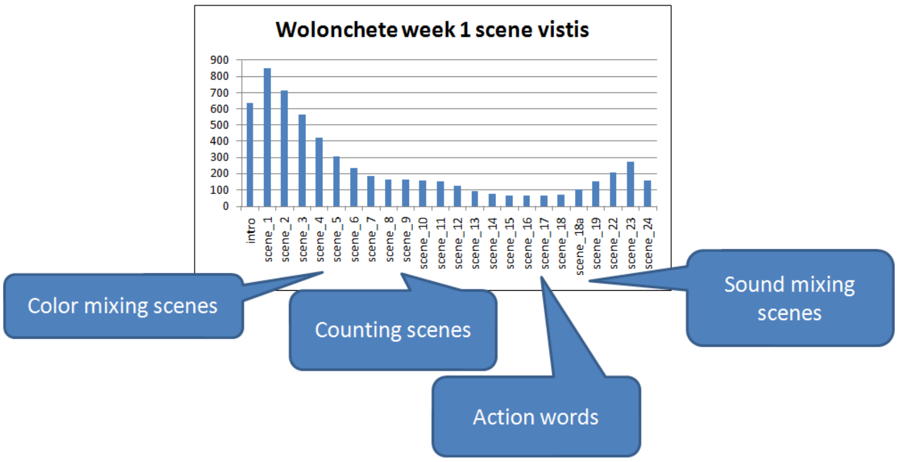


**Figure 6. Logs from the second site showing how often children interacted with complex literacy ideas in first week.**

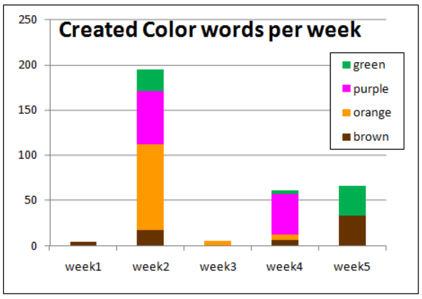


**Figure 7.Color words created over 5 weeks of use**

## 3.5 Site visits

Researchers and technicians, one at a time, performed site visits approximately every month initially to diagnose problems (like the power-on reset error) and provide anecdotal reports on the project from the ground.  These visits brought to light many assumptions we had about learning to read and usability.

- *Orientation:* One unexpected issue with tablet software is that we’re in danger of inadvertently teaching kids the entire alphabet upside down or sideways. Landscape orientation seemed the most stable, natural, so that the child can rest the tablet on both thighs. When there is no teacher to say\ “this is the way you hold a book” or “this is the top of the letter and this is the bottom.”
- *Reverse engineering*: The logistics team reported that the children were able to bypass our password protection for the home screen. The home screen logged which apps were opened. Some children had customized their desktops by creating their own shortcuts to favorite apps. From our data, tablet usage seemed to decline, until the visiting engineer reported that the children were bypassing the launch application in order to access the shortcuts. They were not supposed to be able to do this, but somehow they observed one of the weekly technicians entering the password.
- *Ubiquitous and social learning*: The tablet images showed us that the children were gathering in groups and that the adults were often present (Figure 1b). There was clear evidence of social learning, as the

average number of people detected in the photos was often more than 3. The images showed that the children took tablets everywhere; even going to sleep. When researchers visited the site, they observed children socially teaching one another. They documented choral singing of the ABCs and group learning.

3.4 Project Status

The data collection system operated for almost a year, with few problems besides the initial data timestamp issue. Other hardware problems were easily fixed: over time the SD cards became corrupted; A few tablet screens were cracked and broken, but there were extra tablets that could be put to service.

As the project reached one year, the logistics team was still operating with no change in protocol since first implementation.

One researcher went on site to visit the children, and assessed that they were at a kindergarten level of learning. Publication on the literacy assessment is pending.

## 4 Conclusion

We have presented our proof-of-concept for a method to deliver educational materials to remotely located people in developing countries. When possible, we sought to use local resources.

### 4.1 From the perspective of our Ethiopian coauthor:

Overall, the villagers are really excited and supportive of our project. They have donated the land we used to build the charging hut and have also assigned two people who will watch over the solar powered charging station, and watch over the children, and make sure the children and the tablets are safe.

This area in Ethiopia is pretty isolated. The parents are so excited about the research and see this occasion as a beginning of good things to come to their children. The whole exercise showed us how eager the parents are for their children to be educated and they assured us that they will do everything to make sure the research succeed as long as we don't abandon them at the end, and bring education to their children.

On December 10, 2013 as part of the Science and Technology Information Week – of the African Union - the children from Wonchi one of the villages involved in the research were invited to demonstrate what they have learned on their Motorola Xoom tablets to conference participants.

It was symbolic to see the children of Wonchi running around the lobby of the African Union Commission with their tablets, and taking pictures with the Ethiopian Ambassador to the African Union - Ambassador Kongit Sinegiorgis. Some conference participants raised important questions about the absence of a teacher and how the children learned without the physical presence of an instructor. The Ethiopian education officer present explained in depth that this was a new way of learning and that the children would be transitioning into schools, and thus building the necessary building blocks for future success

If children can learn to read on their own, UN Millennium goal #2, universal primary education for all, will be achieved for these children far sooner than the time it will take to build schools and train teachers in sufficient numbers.

One year after the launch in Wonchi, the children are playing, learning from each other, and engaged in a way that would have been hard to imagine just a year ago. And while the pilot has ended, the solar huts that were built to serve as gathering places remain, the tablets are with the children until they can no longer be used, and the sense of wonder and discovery experienced by each child can be visibly seen on their young innocent faces. However, if we do not continue to provide them with more materials and transition them to learning the Ethiopian curriculum, we have done them a great disservice. As they are far advanced than the children in the traditional school setting and they will definitely not be engaged nor challenged.

To this end, we are in consultation with the Federal Democratic Republic of Ethiopia Ministry of Education, Ministry of Communication and Technology, United Nations Economic Commission for Africa and African Union commission to continue consultation on how to continue the children's learning through technology based education. Investing in children is the primary responsibility of any nation, and Education is the single most important investment, if our African children are to thrive as individuals and global citizens by 2063.

## References


Aharony, N., Pan, W., Ip, C., Khayal, I., & Pentland, A. (2011). Social fMRI: Investigating and shaping social mechanisms in the real world. Pervasive and Mobile Computing, 7(6), 643-659.

Alonso, J. B., Chang, A., & Breazeal, C. (2011). Values impacting the design of an adaptive educational storybook. In *Interactive Storytelling* (pp. 350-353). Springer Berlin Heidelberg.

Ananian, C. S., Ball, C. J., & Stone, M. (2012, June). Growing up with Nell: a narrative interface for literacy. In *Proceedings of the 11th International Conference on Interaction Design and Children* (pp. 228-231). ACM.

Boteanu, A., & Chernova, S. (2012, July). Modeling Topics in User Dialog for Interactive Tablet Media. In Eighth Artificial Intelligence and Interactive Digital Entertainment Conference.

Cervantes, R., Warschauer, M., Nardi, B., & Sambasivan, N. (2011, May). Infrastructures for low-cost laptop use in Mexican schools. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 945-954). ACM.

Chang, A., & Breazeal, C. (2011, June). TinkRBook: shared reading interfaces for storytelling. In Proceedings of the 10th International Conference on Interaction Design and Children (pp. 145-148). ACM.

Chang, A., Breazeal, C., Faridi, F., Roberts, T., Davenport, G., Lieberman, H., & Montfort, N. (2012, May). Textual tinkerability: encouraging storytelling behaviors to foster emergent literacy. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (pp. 505-520). ACM.

Chang, A., Nunez, D., Roberts, T., Sengeh, D., & Breazeal, C. (2013, June). Pre-pilot findings on developing a
literacy tablet. In Proceedings of the 12th International Conference on Interaction Design and Children (pp. 471-474). ACM.

Funf Open Sensing Framework
http://code.google.com/p/funf-open-sensing-framework/source/checkout

Kumar, A., Tewari, A., Shroff, G., Chittamuru, D., Kam, M., & Canny, J. (2010, April). An exploratory study of unsupervised mobile learning in rural India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 743-752). ACM.

Mills-Tettey, G. A., Mostow, J., Dias, M. B., Sweet, T. M., Belousov, S. M., Dias, M. F., & Gong, H. (2009, April). Improving child literacy in Africa: experiments with an automated reading tutor. In Information and Communication Technologies and Development (ICTD), 2009 International Conference on (pp. 129-138). IEEE.

Mitra, S., & Rana, V. (2001). Children and the Internet: Experiments with minimally invasive education in India. *British Journal of Educational Technology*, 32(2), 221-232.

Mitra, S., Dangwal, R., Chatterjee, S., Jha, S., Bisht, R. S., & Kapur, P. (2005). Acquisition of computing literacy on shared public computers: Children and the" hole in the wall". Australasian Journal of Educational Technology, 21(3), 407.

Wesolowski, A., Eagle, N., Tatem, A. J., Smith, D. L., Noor, A. M., Snow, R. W., & Buckee, C. O. (2012). Quantifying the impact of human mobility on malaria. Science, 338(6104), 267-270.