Research Paper Presentation
Assessing the use of technology and Khan Academy to improve educational outcomes

Presenter: Carlued Leon
MANAUS Consulting
11620 Wilshire Blvd, Suite 610
Los Angeles, CA 90025
p: +1 213.599.7442
e: cleon@manausconsulting.com

Authors: Carlued Leon, Global Research Manager, MANAUS
Tamar Koosed, President, MANAUS

Keywords: technology; Khan Academy; math performance; program evaluation.

This research paper was originally prepared for the Fundación Sergio Paiz Andrade on May 27, 2016. The paper was significantly shortened to comply with ISTE’s submission guidelines. The full version of the report can be accessed at www.manausconsulting.com.
Table of Contents

EXECUTIVE SUMMARY.................................................................................................................. 2
INTRODUCTION.................................................................................................................................. 4
  The Sergio Paiz Andrade Foundation .......................................................................................... 4
  Khan Academy ............................................................................................................................ 4
  Learning Equality ........................................................................................................................ 4
  Endless Mobile ............................................................................................................................ 4
  Pilot Project in Sacatepéquez ...................................................................................................... 5
EVALUATION METHODOLOGY...................................................................................................... 6
  Sampling .................................................................................................................................... 6
  Data Analysis ................................................................................................................................ 6
  Control Variables ........................................................................................................................ 7
  Methodological Limitations ........................................................................................................ 8
EVALUATION FINDINGS ................................................................................................................. 10
  Technology Combination Impact ............................................................................................... 10
  Internet vs. No Internet .............................................................................................................. 10
  Discussion on Impact Size ......................................................................................................... 11
CONCLUSIONS .............................................................................................................................. 12
REFERENCES ................................................................................................................................. 13
EXECUTIVE SUMMARY

This report presents the results of an independent evaluation of the Sergio Paiz Andrade Foundation’s (Funsepa) pilot program in Sacatepéquez. The pilot aimed to take an innovative approach to Funsepa’s traditional programming and enhance student academic performance in mathematics. Specifically, Funsepa wanted to test the benefits of incorporating Khan Academy’s tools, which include online and offline resources such as practical exercises, instructional videos, and a self-paced learning dashboard, into the organization’s existing programs (i.e. computer labs and teacher training).

The foundation also aimed to understand whether the use of Khan Academy with different technologies (computers and tablets) and Internet connectivity had a differentiated effect on academic performance. The evaluation specifically tested four sub-interventions that made up the pilot: 1) schools with 16-computer labs with Khan Academy (with and without Internet); 2) schools with 30-computer labs and Khan Academy (with and without Internet); 3) tablets and Khan Academy (without Internet); and 4) one-Endless Mobile computer and Khan Academy (with and without Internet).

MANAUS, a third-party evaluation consulting firm, utilized a quasi-experimental design that combined both quantitative and qualitative methods. This approach incorporated quantitative data, gathered through surveys and standardized math exams, along with qualitative information, gathered through focus group discussions. It also employed data extracted from the Khan Academy platforms to understand how students used this tool. This mixed-method strategy allowed the evaluation team to gain a holistic understanding of the pilot’s benefits to academic performance.

The assessment used three evaluation groups comprising 30 schools: 1) the pilot intervention group, including 14 schools with the sub-interventions described above (technology and Khan Academy, with/without Internet); 2) the Funsepa group, comprising six schools that received Funsepa’s traditional program (technology and teacher training, but not Khan Academy); and 3) a comparison group of 10 schools with no access to technology or Khan Academy. To arrive at findings, the evaluation compared the scores of students in the pilot intervention group and in the Funsepa group, separately, against those of students in the comparison group.

Funsepa

Funsepa is a nonprofit organization that supports the use of technology as a tool for improving education to contribute to the social and economic development of Guatemala. FUNSEPA’s core program, Tecnología para Educar (TPE), focuses on the provision of computer equipment to public schools and teacher training to complement traditional methods of teaching.

1 Funsepa rolled out the tablets sub-intervention using only the offline platform (KA Lite) because the operating system of the tablets did not support the online platform.
2 Endless Mobile is an organization that provides computers that utilize a Linux-based operating system that can be plugged into an affordable monitor and keyboard, creating an inexpensive and user-friendly desktop environment. Given its potential cost-effectiveness, Funsepa wanted to explore the effect of using Endless Mobile as part of the different technology combinations of the pilot intervention.
The evaluation found that combining technology with Khan Academy produces a higher positive effect on student math performance. Relative to the comparison group, the pilot intervention leads to an average increase of 10 points in math scores, out of a maximum possible score of 100 points, which is double the increase of five points produced by Funsepa’s traditional program.

The evaluation arrived at these findings by controlling for other factors that could have an influence on academic performance, such as gender, socioeconomic status, class size, teacher’s math score, grade repetition, availability of computers or tablets at home, and frequency and time of technology use at school, among other factors.

When comparing the different sub-interventions against the comparison group, the evaluation found that the provision of tablets and Khan Academy has a larger effect on student math performance than the other sub-interventions. On average, the combination of tablets with Khan Academy leads to a 10-point increase in math scores, while the use of computers with Khan Academy leads to an average increase of eight points.

In terms of Internet connectivity, the study found that sub-interventions with no Internet produced a larger effect on math scores, an eight-point increase, than sub-interventions with Internet, a six-point increase, when compared to the comparison group. However, schools in the sub-interventions with Internet faced challenges with the Internet connection. It is thus possible that these sub-interventions may have led to higher math outcomes if the Internet connection had been reliable throughout the implementation of the pilot.

Results for the sub-interventions and for internet connectivity are statistically significant and controlled for factors such as gender, socioeconomic status, class size, teacher’s math score, grade repetition, having technology at home, frequency of technology use at school, and exposure to the interventions.
INTRODUCTION

Funsepa’s pilot intervention in Sacatepéquez integrated the use of technology, Internet access and digital content from Khan Academy, an open educational platform, into classroom instruction. With this pilot, Funsepa aimed to enhance student learning and academic performance in mathematics among primary school students. The evaluation sought to assess the extent to which the pilot intervention contributed to improve math achievement among students, as well as to gauge differences in math scores across the various technology and Internet connectivity combinations employed in the pilot program. The evaluation ultimately aimed to identify the best strategy for Funsepa to improve student educational achievement through its programs. To this end, the study utilized a quasi-experimental methodology that combined quantitative and qualitative techniques to measure the benefits of the pilot.

The Sergio Paiz Andrade Foundation

The Sergio Paiz Andrade Foundation (Funsepa) is a nonprofit organization established in 2004 with the mission of contributing to the social and economic development of Guatemala through the use of technology as a tool to improve education. Funsepa implements its activities through a core program, Tecnología para Educar (TPE). TPE provides computers to public schools in Guatemala as well as training to public school teachers in the use of computers and their effective incorporation into traditional teaching methods.

Khan Academy

Khan Academy is a nonprofit organization that provides free digital educational materials online. Its resources include practical exercises, instructional videos and a personalized learning dashboard that allows students to study at their own pace in and out of the classroom. Educational tools are offered in a wide range of subjects. Specifically, math ‘missions’ guide students through different mathematical concepts using an adaptive system that identifies the individual’s strengths and learning gaps.

Learning Equality

In cases where Internet access is not available, users can employ KA Lite, an open-source software platform that makes Khan Academy’s videos and exercises available on offline devices. KA Lite, created and maintained by the nonprofit organization Learning Equality, has been widely deployed in developing communities around the world.³

Endless Mobile

Endless Mobile is an organization that provides affordable access to technology worldwide. Endless utilizes a Linux-based operating system that can be plugged into an affordable monitor and keyboard, creating an inexpensive and user-friendly desktop environment.⁴ Given the potential cost-benefits of Endless Mobile computers, Funsepa wanted to leverage the evaluation and explore the effect of using Endless Mobile as part of the different technology combinations in the pilot.

³ Additional information can be found on https://www.khanacademy.org and https://learningequality.org/ka-lite/.
⁴ Endless Mobile’s (https://endlessm.com) and external interview with Endless Mobile’s CEO (https://endlessm.com/press/).
Pilot Project in Sacatepéquez

In 2012, results from an independent evaluation of Funsepa’s programming found that the academic performance of students in schools that received the foundation’s core program was better than that of students in schools without Funsepa’s programming. Given the demonstrated success, Funsepa sought to test an innovative approach to its programming that combines the use of technology, Internet connection and access to open educational material offered by Khan Academy and KA Lite.

Funsepa implemented the pilot in 14 randomly selected schools, with each school receiving access to Khan Academy, through the online or offline platforms, along with different combinations of technology and Internet access. All grades in each school participated in the pilot intervention, reaching over 3,600 students. To minimize the costs associated with data collection activities, the evaluation sampled a subset of these students, as detailed in the Methodology section of this report. Table 1 illustrates the different intervention combinations.

<table>
<thead>
<tr>
<th>Pilot sub-interventions</th>
<th>With Internet (Khan Academy)</th>
<th>Without Internet (KA Lite)</th>
<th>Total schools in subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 Funsepa computers + Khan Academy</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>This subgroup included students in schools that were provided a 16-computer lab with access to the online or offline Khan Academy platforms.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 Funsepa computers + Khan Academy</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>This subgroup included students in schools that were provided a 30-computer lab with access to the online or offline Khan Academy platforms.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tablets + KA Lite</td>
<td>N/A**</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>This subgroup included students in schools that were provided with one tablet per student in each classroom with access to the offline KA Lite platform.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-Endless Mobile computer + Khan Academy</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>This subgroup included students in schools with one Endless computer in each classroom with access to the online or offline Khan Academy platforms.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** Funsepa rolled out the tablets subgroup using only the offline platform because the operating system of the tablets did not support the online platform.

---

5 In 2010, Funsepa commissioned MANAUS Consulting to conduct the evaluation of its core programs. More information on the findings of this evaluation can be found in document: FUNSEPA Monitoring and Evaluation Final Report, July 2012.
**EVALUATION METHODOLOGY**

The evaluation utilized a quasi-experimental design that combined quantitative and qualitative methods to measure the effect of the pilot intervention on students’ math achievement. The mixed-method approach specifically incorporated quantitative data, gathered through surveys and standardized math exams, along with qualitative data, gathered through focus group discussions. The evaluation also used data extracted from the Khan Academy and KA Lite platforms to better understand how students used these tools. This mixed-method strategy allowed the evaluation team to gain a holistic understanding of the pilot’s benefits by giving context to quantitative findings to explain overall results.

**Sampling**

The study employed a two-stage randomization strategy to select the evaluation sample. In the first stage, the evaluation team randomly selected schools into each of the evaluation groups. In the second stage, evaluators selected students at random within the schools for data collection. At the school level, 14 schools were randomly selected from all public primary schools in Sacatepéquez that had no computer labs or access to any other technology to receive the pilot intervention. These 14 pilot schools were further subdivided into seven pilot subgroups to receive computers, tablets or Endless Mobile computers, and different modalities of access to Khan Academy (online or offline). Ten schools were similarly selected to receive no technology or access to Khan Academy and served as the comparison group. Six schools received Funsepa’s core program, TPE, but did not have access to Khan Academy (hereinafter referred to as the Funsepa group). These six schools were randomly selected from among schools that had recently received Funsepa’s core program.

At the student level, the sample included students from the third and sixth grades only to reduce evaluation costs and to minimize school disruption at the time of data collection. The evaluation team calculated sample sizes based on a 95% confidence level to arrive at a statistically representative sample. The team then randomly selected students from the class rosters to be surveyed and tested. The desired sample of students per school was evenly distributed across grades. Where there was more than one section per grade, the target school sample was divided accordingly across classrooms. After data collection activities, the final sample included 2,356 students.

**Data Analysis**

The information collected through primary and secondary sources was systematically analyzed to arrive at findings. Data from standardized tests and surveys was analyzed using Stata. The main analysis evaluated the overall impact on student achievement by measuring the statistical difference between students in intervention schools and students in comparison schools on the average standardized test score across time. The analysis was conducted in stages to examine changes in core coefficients as control variables were added or removed from the regression specification. In addition, the evaluation analyzed differences in test outcomes among intervention subgroups to provide Funsepa with a good understanding of which sub-intervention leads to the greatest increase in math scores.
Control Variables

The evaluation controlled for a series of factors that can influence academic performance beyond receiving the pilot intervention. The evaluation generally controlled for the following factors:

Gender. This variable controls for whether the student is a female or a male. Some studies have found that gender importantly explains variances in academic achievement, reporting significant differences between male and female student performance in subjects like science and math (DeBaz, 1994; LoGerfo, Nichols, and Chaplin, 2006; Guiso et al., 2008; Bedard and Cho, 2010, Niederle and Vesterlund, 2010).

Class Size. This variable controls for the number of students in the classroom of each student in a given year. Research shows that smaller classes result in a higher level of student academic performance (Hou, 1994; Franklin et. al., 1991; Goldfinch, 1996; Vanderberg, 2012). Other studies found that small classes also have a positive effect on factors beyond academic performance, such as student retention (Lopus & Maxwell, 1995).

Household socioeconomic status. This variable controls for the socioeconomic characteristics of the household of the student, as measured by the Progress out of Poverty Index (PPI). The PPI is a poverty measurement tool that scores answers to 10 simple questions on household characteristics and asset ownership to compute the likelihood that those in the household are living below the poverty line. Children from low socioeconomic households develop academic skills more slowly than children in higher socioeconomic groups (Morgan, Farkas, Hillemeier & Maczuga, 2009). Socioeconomic level has also been identified as a cause of the “digital divide,” the gap between students who have access to digital technology and students who do not (Mason & Dodds, 2005).

Availability of technology at home. This variable controls for whether students are exposed to computers or tablets at home. Past research has found a positive association between technology usage and academic performance in math and science (Fletcher, 2003; Galuszka, 2007). Thus, students who have access to technology at home are more likely to benefit from the use of the technology and Khan Academy, as they may be more comfortable working with technology in general.

Favorite subject at school. This variable controls for whether math is the subject that student likes the most. Students who like math are more likely to perform better in the subject (Raza & Shah, 2011; Schenkel, 2009; Waxman & Houston, 2012). It is also possible that students who like math may take more advantage of the use of the Khan Academy platforms than students who prefer other subjects.

Grade repetition. This variable controls for whether the student has repeated any grade at least once. Available literature shows that grade repetition can have important effects on academic achievement, including long-term implications associated with grade-repeaters eventually falling further behind. Research has found that grade repetition has adverse effects on student self-esteem, peer relationships and attitudes towards school. Grade repetition also has negative effects at the school level, as high levels of grade repetition can lead to increased class sizes and classroom management problems due to large age differences among pupils in the same classroom (UNESCO, 2006).
Length of exposure to intervention. This variable controls for the number of months that the school and its students have been exposed to the pilot intervention, as the different pilot sub-interventions did not all start at the same time. Students in schools where Funsepa rolled out the pilot intervention earlier are more likely to have benefited from it more and hence, have higher academic performances than students in schools where the pilot intervention was introduced later.

Frequency and time of technology use. This variable controls for how often (days per week) the student used the technology and the time (in hours) s/he spends working with the technology and Khan Academy. Students who spend little time working with the technology and Khan Academy are more likely to underperform in math than students who work with it more frequently or for longer time. The analysis included this control variable in the form of an interaction between frequency and time.

Teacher’s math performance. This variable controls for the performance of teachers on the standardized math exam, as a proxy for teacher quality. Students in classrooms with teachers who demonstrate subpar knowledge of primary education level math content are more likely to perform poorly in the math test. Research has found that when it comes to performance on reading and math tests, a teacher has two to three times the impact of any other school factor, including services and facilities. Evidence also suggests that a teacher’s impact on student achievement remains consistent even if the teacher changes schools and regardless of whether the new school is less advantaged than the old one (Rand, 2012).

NOTE: Variables age and grade were highly correlated. Controlling for both variables generally made standard errors larger, making the model less precise. For this reason, the model was run using both variables together and also separately to choose the best model. The team ultimately decided to use the model that only includes age to report findings.

Methodological Limitations

Quasi-experimental studies and the difference-in-differences statistical analysis technique are among the most rigorous research methods for program evaluations. Nonetheless, the methodological and logistical limitations outlined below can affect the validity of the findings presented in this report:

The design and implementation of the pilot sub-interventions varied significantly in terms of level of access to the technology. Because school administrators typically assign visits to and time in the computer lab evenly among all grades in the school, access to the technology was more limited for students in the 16- and 30-computer lab sub-interventions. In contrast, the tablets and Endless Mobile sub-interventions were less affected by this issue because the technology was usually available within the classroom. The design of the tablets sub-intervention also differs in that each student received his/her own device. Given these differences in the design and implementation of each of the different sub-interventions, the findings of the evaluation may be misestimating the real effect on academic performance of specific technology combinations.

Schools in the Funsepa group may have different characteristics than schools in the intervention and comparison groups. Schools in the Funsepa group had to submit an application to participate in TPE. This means schools in this group chose to receive computer labs and teacher training. In contrast,
schools in the pilot and comparison groups were selected fully at random, regardless of whether they had a previous interest in receiving Funsepa’s program or not. Because of this, schools in the Funsepa group may have characteristics that differ significantly from schools in the pilot and comparison groups, such as better administrators who understand the benefits of technology for education. Though these issues do not affect the estimation of the effect of using Khan Academy on the math performance of students in the pilot group, as these estimations are derived from comparing the pilot group with the comparison group (and not the Funsepa group), any comparisons between the effect of the pilot intervention and that of Funsepa’s traditional program should be interpreted cautiously.

**School size varied importantly across sub-interventions, affecting the degree of exposure to Khan Academy.** Though the evaluation randomly selected schools into the different sub-intervention groups, Funsepa modified the subsample of schools allocated to the tablets and Endless Mobile subgroups due to logistical issues and implementation costs. Schools in the computer sub-interventions were significantly larger than in the other sub-interventions, in particular, the tablets sub-intervention. Specific characteristics associated with school size—such as school management or school principals’ capacity to closely oversee student progress—can also have an effect on student academic performance and the evaluation did not control for such characteristics. The findings of the study may be misestimating the real effect of the different sub-interventions on math performance.

**The quality of the Internet connection varied significantly across the sub-interventions that combined technology and the online Khan Academy platform.** Sub-interventions incorporating Internet faced significant challenges during the implementation phase. Given the overall infrastructural challenges in Guatemala, Funsepa had difficulties installing Internet in the schools randomly selected into this subgroup. In focus group discussions with teachers, a common constraint mentioned was the unreliability and lack of speed of the Internet connection, without which these students could not use Khan Academy. Teachers reported that issues with the Internet connection were relatively frequent and some indicated that their students were unable to use Khan Academy for up to a month. As such, the findings of the evaluation may be underestimating the effect the online Khan Academy platform could have on student math performance had the Internet connection been consistent.

**The use of teacher performance on the standardized math exam provides a limited indicator to assess teacher quality.** The evaluation tested teachers’ math knowledge as a proxy for teacher competency. However, quality of teaching involves many teacher characteristics beyond content knowledge. Recent research indicates the best way to assess teachers’ effectiveness is to look at their on-the-job performance, specifically what they do in the classroom, regardless of where they went to school, whether they are licensed or how long they have taught for (Rand, 2012). Though teachers’ math knowledge is correlated with their effectiveness in teaching math and controlling for it is better than not controlling for any teacher characteristic, the evaluation may be misestimating the pilot intervention’s effect by only controlling for teachers’ knowledge of the math content.
EVALUATION FINDINGS

The evaluation found that combining technology with Khan Academy and KA Lite produces a higher positive effect on student math performance than the traditional Funsepa intervention of providing the technology without the Khan Academy and KA Lite platforms. Relative to the comparison group, participation in the pilot intervention leads to an average increase of 10 points in math scores, out of a maximum possible of 100 points, which is double than that produced by Funsepa’s traditional intervention, a five-point increase.

When exploring the effect of the pilot intervention by grade, the study found a similar effect on math performance, with sixth grade students benefiting slightly more. Relative to the comparison group, third graders increase their math score by an average of eight points, while sixth graders increase their scores by an average of nine points.

These findings are statistically significant and controlled for other factors that can have an influence on student academic performance, such as gender, socioeconomic status, class size, teacher’s math score, whether the student had repeated a grade at least once, availability of computers or tablets at home, and frequency of technology use at school, among other factors.

Technology Combination Impact

When comparing the different technology combinations against the comparison group, the evaluation found that the provision of tablets and Khan Academy has a larger effect on student math performance than the other technology sub-interventions. On average, the combination of tablets with KA Lite leads to a 10-point increase in math scores, out of a maximum of 100 points, while the use of computers with Khan Academy/KA Lite leads to an average increase of eight points. These results are also statistically significant and controlled for factors such as gender, socioeconomic status, class size, teacher’s math score, grade repetition, availability of computers or tablets at home, and frequency of technology use at school, among others.

Internet vs. No Internet

The evaluation found that sub-interventions with no Internet produced a slightly larger effect on student performance, an average eight-point increase in math scores, than sub-interventions with Internet, at
six points, when compared to the comparison group. These findings are statistically significant and also controlled for factors such as gender, socioeconomic status, class size, teacher’s math score, grade repetition, availability of technology at home, and time of technology use, among other factors.

However, as explained before, schools in the sub-interventions with Internet faced significant challenges during the implementation phase. It is then possible that the subgroup with Internet may have had similar or higher math outcomes if the Internet connection had been consistent and students had not experienced interruptions in their exposure to Khan Academy.

On the other hand, it is also possible that the sub-interventions without Internet led to a higher effect due to differences in the way the online and offline platforms work. Specifically, the offline platform deletes the progress of a student when s/he responds a problem incorrectly, forcing the student to start all over again. It is possible that this “forced” review of the exercises helps reinforce knowledge on key math concepts and the overall process to approach the exercise, leading to better performance in the math exam.

**Discussion on Impact Size**

The results shown above reflect the impact of the pilot intervention as it was implemented on the ground. However, the combination of technology and Khan Academy/KA Lite may potentially have an even larger effect on student performance. The evaluation identified various elements that possibly hindered a more successful integration of Khan Academy/KA Lite into traditional instruction.

- **Students’ limited exposure to Khan Academy.** Students’ exposure to the technology was limited, with students generally using the technology one to two times a week for less than one hour at a time. Limited access to the technology consequently constrained exposure to Khan Academy/KA Lite.

- **Unreliable Internet connection.** As explained earlier, schools that used the online platform experienced significant issues with the Internet, hindering their ability to access Khan Academy. These sub-interventions included the four largest schools in the sample, so there is a possibility that students may have needed to pair up to use the computer or share the lab with other grades in the school, further limiting students’ exposure to Khan Academy.

- **Restricted use of Khan Academy.** In focus groups, most teachers said they restricted what students could do on Khan Academy/KA Lite. Some indicated they did not want students to get ahead in the curriculum, so they discouraged students from exploring the platform on their own. One of the advantages of Khan Academy/KA Lite is precisely the possibility of self-pacing one’s learning process. This restriction limited student’s capacity to maximize the benefits of the platforms.

---

6 According to KA Lite developer, Learning Equality, the KA Lite version used in the study (version 0.12) requires users to get 10 out of the last 10 questions correct in order to complete an exercise. In newer versions, users only need to get eight out of 10 questions correct.
CONCLUSIONS

The evaluation found that the provision of technology and access to Khan Academy in public schools in Sacatepéquez, Guatemala led to positive effects on student learning and achievement in math. The evaluation also found positive effects across the different technology combinations, with the tablets sub-intervention having a larger effect on student math performance. In terms of Internet access, the study found that the combination of technology with Khan Academy/KA Lite leads to positive math performance outcomes, regardless of Internet connectivity.
REFERENCES


