Project Title: Evaluating the impact of digital tools to teach math and science in Chile

IDRC Project Number 108369-001

Research Organizations involved in the study: CIAE, Universidad de Chile

Location of Study: Santiago, Chile

By: Roberto Araya, CIAE, Universidad de Chile

Report Type: Final Report

Date: Date the report is submitted: June 30, 2019
1. Executive Summary:

This project is financed in part by IDRC and in part by the Inter-American Development Bank (IDB). The project stems from the IDB research line determining the effectiveness of various educational programs, in particular, Computer Aided Instruction (CAI) in math and science for vulnerable elementary school students. The goal of the project is to produce evidence to inform policies on the use of information and communication technologies (ICTs) to improve the quality of education in Science, Technology, Engineering, and Math (STEM) in schools.

The project is a Randomized Control Trial (RCT) study and it lasted 30 months. At the beginning of 2017, 24 low- or middle-low socio-economic status (SES) schools from Santiago, Chile were selected, each with two classes in fourth grade. One class was assigned to the control group and the other to the treatment group. In total, 24 classes were selected as treatment and 24 as controls. In treatment classes, students performed two weekly math sessions in a computer lab with an innovative technology program that uses gamification. To assess the effects of this program, baseline, midline and endline math and language assessments were carried out. Additionally, data from the SIMCE test (the Chilean national standardized test) were obtained to measure effects in an external, broad measure of math and language learning.

Researchers from the IDB were responsible for measuring the effects of the program. Results indicate that ConectaIdeas generated a large statistically significant improvement in Math learning. The preferred specification shows an effect of 0.27 standard deviations. The estimated effects on Math under alternative specifications range from 0.22 to 0.29 standard deviations. Even though the program aimed to improve learning in Math, it could have generated spillovers to other subjects. Nevertheless, estimates on Language are close to zero and not statistically significant.

The program also affected several non-academic outcomes. On one hand, it increased students’ preference towards using technology for math learning and promoted the idea among students that study effort can raise intelligence. On the other hand, the program increased math anxiety and reduced students’ preference towards teamwork. These results suggest that gamification could be an important tool to boost student learning, but that it may bring unintended consequences.

In parallel to the described quantitative evaluation, a qualitative evaluation was carried out by an independent group of researchers, in order to study the attitudes and perceptions of students, teachers, and principals. It concluded that all the actors interviewed had a positive perception of the experience. The project generated a process of change that was observed with interest by principals and teachers, and that children found a play space for learning mathematics.
2. The Research Problem:

The main objective of this research project was to investigate how technology can be used to improve learning outcomes in low SES schools in Latin American countries. More specifically, the project aimed to measure the effects of the ConectaIdeas program on math learning. To that end, learning levels of students receiving the program are compared to those from a control group.

Besides this main objective, there are a number of other aspects that were important to explore to provide a comprehensive assessment of the program. First, we analyzed heterogeneous effects across students and schools. That is, we explored whether the effects of the program vary across important students’ characteristics such as gender, socioeconomic status, and baseline learning levels (among others). We also explored whether the effects of the program vary along school characteristics.

Second, we used the rich data collected by the platform to provide an objective and detailed assessment regarding when, how much, and where the technology was used. In particular, we analyzed the average number of online sessions that students participate in, how long where these sessions and where these sessions took place (home or school). We also studied the intensity of use of different features of the implementation (students that are monitor, assistance to students, and sessions with peer review).

Third, we analyzed whether the program generated unintended negative effects on learning in language. This is important considering that language skills are foundational in the sense that they are used to develop a vast array of other skills.

Additionally, through a qualitative study with semi-structured interviews and classroom observations, we studied the opinions, perceptions and practices of use of the platform by teachers and students, as well as principals’ opinions of the intervention.

The research process led us to a revise our view on students outcomes so that to include also the effect of the program on socioemotional skills and attitudes, such as attitudes towards math, pleasure with math, whether students consider that intelligence is fixed or malleable, self-efficacy, and attitudes toward collaboration and working in teams. We have collected data on all these aspects. Moreover, the SIMCE also includes a student questionnaire that collects data on some of these dimensions. Additionally, we collected information to study some effects of the ConectaIdeas program on computational thinking skills.

The project contributed to understand the effect that guided instruction can achieve when it is implemented with appropriate technological support and gamification strategies. It is a first step to guide public policy on how to introduce technology to improve the learning of math in vulnerable schools.
3. Methodology

We implemented a rigorous research method in order to assess the effect of educational programs: a randomized controlled trial (RCT). The plan was to use an RCT to assess the causal effect of the ConectaIdeas program. The project team was tasked with recruiting 24 public schools located in Santiago (to simplify logistics) and that had at least two fourth grade sections. Moreover, the schools should had been classified by the Ministry of Education in the two lowest socioeconomic status categories (out of the five) to test whether ConectaIdeas could also close socioeconomic achievement gaps.

The recruitment process began at the end of January 2017 with the identification of 22 school districts that had schools satisfying the criteria described above. The district directors were contacted first by email and then by phone calls. The ConectaIdeas team then visited 11 districts directors that replied and expressed interest. After that, information sessions with district directors and school principals were conducted in 9 school districts. In the final step, the team conducted school technical visits to verify technical requirements. The technical visits were scheduled to be conducted in 31 schools in 6 school districts. However, after visiting 4 school districts (La Pintana, Maipu, Quinta Normal, and San Bernardo) the team confirmed 24 schools that met all the mentioned requirements. Importantly, the recruitment procedure did not involve individual schools making decisions to self-select into the program.

The makeup of the study sample was quite similar to the sample of low socioeconomic status schools in Santiago with two exceptions: enrollment in the study schools was larger (due to the two-section restriction) and their students performed even worse in math and language. In fact, students in the study sample underperformed the average student in the country by 0.60 standard deviations in language and 0.68 in math. We adopted a within-school, section-level randomization design. Within each of the 24 participating schools, we randomly assigned one of the two fourth-grade sections to the treatment group. These sections participated in the ConectaIdeas program. The other sections were assigned to the control group and received traditional math instruction. For the three schools in the sample that had more than two sections, we only included the first two (i.e., A and B sections) in the evaluation. The randomization was conducted before baseline data were collected, and schools were informed of the treatment status of each section after the baseline was collected in March 2017. There was perfect compliance of program assignment to treatment. That is, all sections assigned to treatment participated in ConectaIdeas and none of the control sections participated in the program.

We estimated the effects of the program using analytical techniques under two main specifications. The first specification involved estimating the following equation:
\[ y^{post}_{ics} = \alpha_1 + \beta_1 \times Treatment_{cs} + \phi_s + \epsilon_{ics} \]  

(1)

where \( y^{post}_{ics} \) is outcome variable in the post period (e.g., Math test score measured in the national standardized exam) for student \( i \) in section \( c \), in school \( s \). \( Treatment_{cs} \) is an indicator variable that equals one if the section was assigned to the treatment group and zero if not. \( \epsilon_{ics} \) is the error term, which should be uncorrelated with the treatment assignment because of random assignment, and \( \phi_s \) are school fixed effects. Our main coefficient of interest, \( \beta \), estimates the average treatment effect of the program on the outcome variable (e.g., the average difference in Math test score between students in the treatment group compared to those in the control group).

Because learning is strongly correlated over time, controlling for the baseline test scores can increase statistical precision. Moreover, doing so can account for potential differences in baseline test score levels. Consequently, the second specification that we used is similar to the first one but also controls for the baseline value of the outcome:

\[ y^{post}_{ics} = \alpha_2 + \beta_2 \times Treatment_{cs} + \gamma_2 \times y^{pre}_{ics} + \phi_s + \epsilon_{ics} \]  

(2)

where \( y^{pre}_{ics} \) is the baseline test score in the respective subject. That is, when estimating effects on Math, we control for the baseline Math test score and when estimating effects in Language we control for the baseline test score in that subject. This is our preferred specification because it controls for potential baseline differences in outcomes and because it should generate more precise estimates.

We also estimated effects using a difference-in-differences specification. In this case, we compared changes in test scores since baseline and the time of measurement between treatment and control students. In this case, we estimated the following equation:

\[ y_{icst} = \alpha_3 + \beta_3 \times Treatment_{cst} + \theta_i + \epsilon_{icst} \]  

(3)

In this specification, there are two observations per student. One observation corresponds to the pre-period, that is the baseline period, and one observation for the post period. \( y_{icst} \) presents the value of the outcome for the student \( i \) in section \( c \) in school \( s \) at time \( t \) (that can be 0 for baseline and 1 for the post period). Also, \( Treatment_{cst} \) is indexed for time and takes the value of 1 for treatment individuals in the period in which we measure the effects (e.g., endline).
and zero otherwise. Finally, all estimates presented throughout the paper will include heteroscedasticity-robust standard errors that are clustered at the section level (the unit of randomization).

Additionally, a qualitative study was done where an external team conducted unstructured interviews to principals, teachers, and students of 12 treatment courses. The qualitative study was carried out in two stages. The first during the second week of June and the first week of July and the second in the month of October 2017. Principals and teachers were interviewed, 12 sessions were observed in the classroom or laboratory. Likewise, 4 video films were made, which were recorded on video with the prior authorization and consent of the teachers and parents of the students. Thus, we had a total record of 16 observations of classes (4 recorded in video and 12 based on observation guide and ethnographic record of them). Once the fieldwork was completed, 48 interviews of principals and teachers were transcribed for the work of coding and subsequent analysis. To analyze the texts and discourses of the interviewees, a method of content analysis based on the Grounded Theory and Semantic Structural Analysis was used.

4. Project Objectives and Activities:

The Specific Objectives of the Grant Agreement are:

1) To rigorously evaluate learning effects (using a randomized controlled trial) and estimate the costs of an ICT-based mathematics education program in Chile (called "ConectaIdeas")
2) Determine the key factors that need to be available and provide clear protocols to ensure the effective implementation and scaling of similar programs in other low-income contexts
3) Seek to inform and influence other public and private institutions to adopt and scale up the positive components of this solution to improve the learning of mathematics and other subjects related to STEM among low-income students in Chile and other Latin American and Caribbean countries.

Activities carried out in the project:

The first two months of the initial stage of the project – January and February 2017– took place during Summer vacation. During these months the project was presented to the Municipal Corporations who are responsible for educational planning and implementation in each municipality. Then, the project was presented to the principals of the respective schools. Later, the research group visited potential schools to ensure that their teachers were interested, and that their infrastructure facilities and internet connectivity were adequate. Finally, 24 schools in 4 municipalities in Santiago, Chile were recruited. All principals signed a document committing to participate in the project through year 2017, and to facilitate the implementation of the project. In January, the project coordinator was
hired and teachers who would be lab coordinators were selected and trained. In February the selected teachers and the software supervisor were hired.

Demographic information was collected of all fourth grade students at all participating schools. Together with the IDB team, statistical analyzes of the characteristics of student population (age and gender) were carried out. Then, an external organization was contracted to conduct a paper and pencil pretest (i.e. a baseline test) of math and language. The test was handed out to students by 6 teachers hired for the project. Immediately after completion of the pretest, IDB specialists randomly assigned classes to treatment and control in each school. It was verified that the characteristics of the individuals in the treatment and control groups were balanced. Then, during the first weeks of the project, the team was reviewing the conditions of connectivity and infrastructure. Four mobile services and high capacity routers were acquired, and teachers were trained in how to use them. In late April the data of the baseline test was obtained.

In April the project team agreed with the IDB team on the content and structure of the information to be gathered throughout the project implementation. To this end, a dropbox folder was created to store project logs and reports. It was agreed to generate a specifically structured weekly logbook with an automatic daily report of the Conectaldeas software. This report specifies what sections of the curriculum were studied and what activities were carried out by students in each session. The joint work of the project research team and the IDB research team has been very fruitful. Joint meetings were held every Tuesday, where the progress of the project was discussed and analyzed; the weekly reports were reviewed; the analysis of the stored information were shared; and the tasks for the upcoming week were specified.

Due to the legal requirements of the approvals and acquisitions in the public administration of Chile, it was not possible to acquire all the services and equipment necessary for the implementation of the project. The fact that the onset of the project – January and February 2017 – coincided with the summer vacations, and that it requires several months to place and execute public tenders and make purchases, for the purpose of starting the project on time in early March (which is when the school year in Chile begins), the project team decided to use services and equipment available at the university (these services and equipment are designated to other ongoing projects). Much fewer tablets were required than it was initially budgeted.

The project went according to the planned schedule. We finally obtained the individualized results from the SMCE national test on November 2019. During December 2018, and January and February 2019 the team produced the analysis presented in this report. We plan to publish the results as an IDB working paper in July 2019.

The main activities were the following:

Diagnostic assessment
We first studied the performance of the students participating in the study on the pretest. The tests used were designed by Aptus, a well-known Chilean organization specialized on testing students’ academic achievement on topics covered by the Chilean curriculum. These were paper and pencil tests. After the test results were known, IDB analysts selected on each school one class as the treatment class, and the other classes on the school were the control classes. The selection mechanism was a random process. 24 classes formed the treatment group and 24 classes formed the control group. The collected baseline data showed a small difference in learning outcomes in math between students in the treatment and control groups (those in the treatment group slightly underperform those in the control group). This aspect was considered during the estimation of treatment effects by considering the baseline learning levels of the students in the estimation of program effects.

Sessions

We implemented the project on 1,446 sessions. This means that each class had an average of about 60 sessions with Conectaldeas.

Request for help

We tracked the collaboration between students. This strategy was implemented through different protocols. One protocol involved that students that correctly finished 10 or more exercises in a row were preselected as potential monitors (teaching assistants). Then the lab coordinator could select some of them as monitors. Monitors then had the job to help peers explaining the concepts in order to be able to do the exercises. Lab coordinators coordinated the support to students using the Conectaldeas platform. Students were instructed to do their request for assistance using the Conectaldeas platform. In this way each request was registered and its corresponding information was saved on the server. In practice, not all request for assistance were done using the system. However, according to the lab coordinators, most of them were done with the platform. During the year, on the 24 classes on treatment, there were 17,274 registered requests for help on a total of 1,161 sessions.

One interesting research question that emerged was related to students’ preference to seek help between teachers or peers, and whether the preference evolved during the year. In particular, we examined whether there was a difference on the amount of request for assistance between first and second semester, and whether the pattern of preferences changed. We found that 8,293 of those requests were done in the first semester (from April to July) and 8,981 requests were done in the second semester (from August to November). Thus, there was a similar amount of requests in both semesters. However, on the first semester there was an equal preference of request for peer and teacher assistance (4.9 requests per hour
effectively available), but in the second semester there was a much higher preference for request for assistance from peers (6.1 requests for peer assistance per hour effectively available versus 3.8 request for teacher assistance).

Just after receiving the assistance, the student could assess the quality of the assistance in a scale from 1 (very bad) to 5 (very good). The assessments show a slight improvement from the first to second semester. They also showed that there was no much difference when the assistance was given by a teacher or a peer. For requests that involved asking for a teacher and then receiving assistance from a teacher, during the first semester the average score was 3.7 and during the second semester the score was 3.9. When the help was requested by the student and the assistance was given by a teacher, the average score was 3.9 in the first semester and 3.9 in the second semester. When the help was requested by the student and the assistance was given by a student, the average score was 3.6 in the first semester and 4.0 in the second semester.

Gender factor

We tracked the implementation of the program according to gender. On the initial baseline math test, girls obtained an average score of 37.2% which is less than 40.9%, the score obtained by boys. The difference was statistically significant. In the control group girls obtained 2.7 points less than boys, whereas in the treatment group girls obtained 4.7 points less than boys. However on the midline test girls in the control group obtained 2.5 points less than boys but on the treatment group girls obtained 3.9 points less than boys. This means that the performance gender gap was reduced slightly more on the treatment group than in the Control group. Additionally, for students below the median, on the treatment group the initial gap of 2.8 points was reduced on the midline test to 0.

Gender differences on students on the top 20%

We tracked the implementation of the program on the best students. We used the baseline test to identify students in the top 20%. Those students that scored on the upper 20% on this test in each class were defined as top students. At the baseline test in the control group 37.99% of the top students were girls and at the treatment group 39.22% of the top students were girls. At the midline test the percentages changed to 37.29% and 35.35% respectively. The drop in the treatment group was not statistically significant.

Gender differences on students on the bottom 20%

We tracked potential changes on the gender composition on the students on the bottom 20%. During the baseline test, 48.43% of the bottom students in the control group were girls and in the treatment group 61.58% of the bottom students were
At the midline test the percentages changed to 45.24% and 48.99%, respectively. The drop in the treatment group is large and statistically significant. This means, that the Conectalideas program reduced the proportion of female on the bottom 20% group of students.

Attendance to Conectalideas sessions in the school

We tracked the attendance rate to Conectalideas sessions of each student in the treatment group in his school.

Attendance rate to school

We tracked the attendance of each student of the 48 classes for each day of the school year. With this information we analyzed the effect of the program on attendance.

Number of exercises

We tracked the number of exercises on the Conectalideas platform. We analyzed for each student on every day the number of exercises done at school and at home.

Number of exercises done at home

We tracked the exercises done at home separately, and analyzed the effect of intensity of exercises done at home.

Curriculum coverage

We tracked the number of exercises done by each student in each strand (Numbers, Geometry, Statistics, Algebra) and each specific objective of the curriculum. This information allowed us to analyze the relative impact of curriculum coverage on the learning gains.

Open-ended questions

Besides regular multiple options exercises lab coordinators posed open ended questions that promote reasoning and argumentation. During the year 1,946 open ended question were posed and 37,433 answers were received, and 14,003 comments to peers were received. This means about 48.71 open ended questions per class, and 81.08 open ended questions if request for comments are also considered. Each student answered on average 48 open-ended questions.

Length of answers to open-ended question

One of the goals of Conectalideas is to develop 21st century skills in addition to math content knowledge. One of these skills is written communication, which includes writing with reasoning and argumentation components. Since fourth-grade
students write very little, ConectaIdeas tracked the lengths of the answers and immediately informed the teacher. The indicator measured was the number of words included in the written responses to open-ended questions. During the first semester the average length was 6.5 words and on the second semester was 7.8 words. This is a sizeable increase. If we do not include answers to request to comment peers’ answers, then the average length was 7.3 words in the first semester and 8.9 words in the second semester. Answers on the Number strand were the shortest with an average of 7.5 words, and the ones on Algebra strand were the longest with an average of 8.9 words. We are interested on investigating the use of reasoning connectors. For example, consecutive connectors. These are the following: 

"[x,por,pr] [que,ke,k,q,qe]", "ya que", "de[v, b]ido", "dado que", "pues", "pues bien", "puesto [que, q, ke]", "enton[s,es]", "asi pues", "por ello", "a casusa", "por ende", "en con[s,ecuen][c,s]ia", "por con[c,s]iguiente", "de modo que", "por lo tanto". In the first semester they were used on 74% of the answers, and on the second semester they were used on 92% of the answers.

**Tournaments**

Another motivational strategy of ConectaIdeas is the synchronous inter-school tournaments. Four tournaments were carried out, two in each semester. All students from every class were expected to participate. And in fact, participation was high: 432 students in the first tournament, 444 in the second, 425 in the third, and 304 students in the fourth tournament (3 schools with classes in the afternoon are not considered in the statistics). After or during the tournament one or two questions were posed to the students to assess the experience. Students assessed the experience on a scale from 1 (very bad) to 7 (very good). The average for the four tournaments were: 6.8, 6.3, 6.4 and 6.1 respectively. Out of 425 students, 67% answered that they preferred to learn math with the spiral game in an inter-school tournament (this was the game used in the inter-school tournaments), 19% preferred to play the spiral game played with a classmate, 4% preferred to play alone, 2% preferred on the computer but not on a game, and the rest preferred traditional lessons.

**Award events**

One of the main components of ConectaIdeas is based on leveraging social motivation strategies to sustain high levels of use of the online learning platform. As mentioned, twice a semester synchronized tournaments were held. Finally, on November 4, 2017, a public award event took place to recognize the efforts of the students. In this event, each class received a cup. During the award event students played an inter-school math game. Then in the next week students assessed the event. Average score of the event obtained was 6.4 in a scale 1 (minimum) to 7 (maximum).
Results

This section answers the main question of the study: did ConectaIdeas affect student learning? To answer this question, IDB analysts computed the program effects on math academic achievement measured using data from the 2017 Chilean national standardized exam (called SIMCE). Results indicate large effects on math achievement for all the specifications presented. In the specification that did not control for baseline math achievement, the estimated effect was 0.22 standard deviations. In the preferred specification by the IDB analysis, which controls for baseline math achievement, the effect was slightly larger at 0.27 standard deviations. In either case, the estimated effects were statistically significant at the one percent level.

Even though the program focused exclusively on math, it could have generated spillover effects into language. For instance, the program could have motivated students to study more overall, or it could have discouraged students in language class if they shifted study time to math instead. Results indicate that the program did not affect language achievement. In both specifications, the estimated effects on language are small and not statistically significant.

Benchmark

A common strategy to benchmark program effects involves comparing them with expected learning progression during a year. Hill et al. (2008) documented that students in the U.S. improve their learning in about 0.54 standard deviations during fourth grade. Assuming that student academic progression in Chile is similar to the U.S., we can think that students that participated in ConectaIdeas advanced substantially more (about 50%) than their counterparts in the control group (0.27/0.54).

Another way to benchmark the effects of ConectaIdeas on math achievement involves comparing to those from other technology in education programs. The magnitude of the effects of ConectaIdeas are well above to those documented in the U.S., above those documented in China but below those documented in two studies in India (Banerjee et al., 2007 and Muralidharan et al., forthcoming). However, caution is needed when comparing the effects of ConectaIdeas to those from these evaluations not only because of potential differences in contexts, but, especially due to potential difficulties in comparing across evaluations that used quite different types of exams. In particular, exams in evaluations implemented in India and China have typically not included state or national exams. Alternatively, we could compare the effects of ConectaIdeas from those documented from other evaluations that have also analyzed effects on the Chilean national standardized exam. In particular, Bellei (2009) documented increases in math and language of about 0.06 standard deviations from a program that extended the school day from
4 to 7 hours a day. In turn, Bassi et al. (2016) found that a program that provided lesson plans to teachers and teacher materials improved math and language test scores in about 0.07 and 0.09 standard deviations, respectively.

Heterogeneity

We explored whether the effects of Conectaldeas varied across different populations defined by gender, mothers’ education and baseline academic achievement. Effects were slightly larger for boys than for girls (0.29 versus 0.24 standard deviations) while effects by subsamples defined by mothers’ education and baseline academic achievement were almost equal.

Effects on Students’ Perceptions

Beyond measuring effects on academic achievement, we also measured effects on a number of dimensions. Results indicate positive statistically significant effects on two areas that are well aligned to prior expectations. To start with, the basis of gamification involves producing a more engaging and attractive experience and indeed 79% of students in the treatment group report preferring doing math sessions in the computer lab instead of in the regular classroom. In contrast, only 59% of students in the control group report preferring doing math sessions in the computer lab. This difference translates to a positive effect of 0.40 standard deviations in students’ preferences towards doing math lessons in the computer lab. In addition, one of the Conectaldeas features involved presenting personalized ads to students to motivate the adoption of a growth mindset. And we document a positive effect of 0.10 standard deviations in this area.

In contrast, there are two areas in which we do not find statistically significant effects, though some effects could have been expected. The first one is on intrinsic motivation, that is, the inherent enjoyment of learning math per se. Because Conectaldeas emphasizes doing math exercises to increase scores and fare better in individual and group competitions, it may reduce math intrinsic motivation. However, we do not find evidence supporting this expectation. In fact, the effect on intrinsic motivation is positive though not statistically significant. The second one is on math self-concept or the self-perception that students hold on their own abilities to solve math exercises. Because Conectaldeas produced large increases in math achievement, we could expect positive effects on this area. Yet, we do not find statistically significant effects.

In turn, there are two areas in which we find statistically significant effects that can be considered as undesirable. In particular, we found positive statistically significant effects on math anxiety of 0.13 standard deviations that could be linked to the social
comparisons and individual and group competitions that are built in ConectaIdeas. We also document negative statistically significant effects on preferences for teamwork of 0.21 standard deviations. This result can be surprising considering that ConectaIdeas promoted within-class collaboration by setting up group competitions. One potential explanation for this unexpected result is that some students may notice the disadvantages of working in teams (e.g. the weaker link between own performance and final outcomes) when participating repeatedly in the ConectaIdeas team competitions.

5. Project outputs

Research

We expect to publish in July 2019 an IDB working paper titled “Does Gamification in Education Work? Experimental Evidence from Chile,” with authors Roberto Araya from CIAE, Elena Arias Ortiz from IDB, Nicolas Bottan from Cornell University, and Julian Cristia from IDB.

We have gained experience in the management and adaptation of technology for vulnerable schools where internet connectivity is unstable. We have learned to use mobile internet and higher capacity routers to connect all the computers or tablets in a classroom.

We have established a collaborative network for research with IDB, the Ministry of Education of Chile and Peru, and GRADE, a non-profit research center in Peru that advises the Peruvian Ministry of Education.

We have achieved significant milestones in educational programs assessment knowledge and in networking with schools and district superintendents.

Capacity

During this period we carried out several dissemination activities. We did presentations of the results in all the four participating municipalities. Additionally, we organized a large-scale event, attended by more than 800 teachers, on December 6, 2017 at the Municipal Gym of the San Bernardo Municipality where we presented the project and initial results.

http://www.ciae.uchile.cl/mail_list/email/362.html

Also two presentations to the Minister of Education were carried out. Moreover, the main dissemination event in the country was the conference titled “Educar con Evidencia” organized by the Ministry of Education and held on August 23, 2018 in Santiago, Chile. The conference was attended by 200 teachers and principals. It was
led by the Director of the Curriculum and Evaluation Unit of the Ministry of Education. Some of the participants were: Robert Slavin and Nancy Madden from the Center for Research and Reform in Education at Johns Hopkins University, Julián Cristia and Samuel Berlinski, from the Inter-American Development Bank (IDB), David Evans from the World Bank, Santiago Cueto from the research center GRADE from Peru, Francisco Gallego from Pontificia Universidad Católica and J-PAL, Susana Claro, Pontificia Universidad Católica, and Roberto Araya from CIAE University of Chile.

http://sitios.mineduc.cl/educar-evidencias/

We also presented the results at the Tata Institute for Social Sciences (TISS) on the 'Connected Learning at Scale: An International Symposium' held in Mumbai, India, August 8 and 9, 2018.

https://www.tiss.edu/view/5/projects/connected-learning-initiative-clix/connected-learning-at-scale-an-international-sympo/

On March 2019 we presented the findings at the Society for Research on Educational Effectiveness (SREE) Spring 2019 Conference, Tensions and Tradeoffs: Responding to Diverse Demands for Evidence, that will take place in Washington, D.C. from March 6-9, 2019.

https://www.sree.org/conferences/2019s/pages/participants.php

In Research on Improving Systems of Education (RISE) annual conference www.riseprogramme.org, Julián Cristia presented the project and its results. It can be seen between 1:41:24 and 1:57:18 in https://youtu.be/-lN3LSRU3xA?t=6080

The project also contributed to provide career advancement opportunities to women. In particular, the coordinator of the project is a woman, and five of seven hired team members are women.

Policy and practice

After learning about the project, the Ministry of Education of Chile got interested in conducting pilot programs and measuring their effect, before launching them at large scale. The Ministry started this new strategy with the organization of the “Educar con Evidencia” conference on August 2018. Currently the Ministry is requesting the assistance of CIAE of the University of Chile to evaluate four programs following the same evaluation strategy.

6. Problems outcomes
Main outcomes

The main scientific outcome was that we obtained rigorous empirical evidence, following the highest standards on program evaluation, that a careful implementation of a guided instruction program based on information technology together with a motivational strategy of gamification produced very good results in student learning.

Additionally, the project helped to produce capacity building in the design of randomized controlled trials, in the administration of educational projects on a large scale, and in the development of laboratory coordinators for the use of technologies in teaching. The CIAE team is now being asked by the Ministry of Education of Chile to conduct other similar evaluation studies using RCT methodologies.

In schools, the project started to operate in late March 2018. It immediately produced an impact on schools and the project team. All of the schools of two Municipalities continued with the program on 2018, and not only for fourth grade but also for third, fifth and sixth grade courses. Additionally a pilot with 8 schools was implemented in Lima. Preliminary results show positive effects. These results were presented on a workshop at GRADE, Lima, Peru, on December 13th and 14th, and in visits to the DRELM of the Ministry of Education in Peru.

The reception of the training by the lab coordinators was very positive. It generated a fruitful discussion of strategies aimed at raising students’ involvement and development of students’ own explanations. Similarly, the reception of the project by school teachers was also very favorable. Our lab coordinators regularly coordinated with the school teachers on the math exercises proposed during the lab sessions, so that the exercises were in accordance with the curricular contents taught by school teachers every week. They were also involved in the online follow-up during sessions, and they gave support to more needy students alerted by the Conectalideas system.

All students in this project were from low SES groups. The project has benefited these students through giving them the opportunity to use the technology in math classes, to develop written explanations of math problems, and to participate in the peer review process, which is something completely unknown in these low SES schools.

A group has been formed with the capacity to implement guided teaching projects with computer support and with the capacity to evaluate the impact of educational programs. Together with the IDB team led by Julián Cristia we are in 2019 implementing an evaluation in 40 Lima schools. On the other hand, the Ministry of
Education is interested that the CIAE team will help in evaluating the impact of digital technology. This is a project that will be implemented on the second semester of 2019.

What was learned?

We learned that teacher communication is central. There were some initial communication problems with teachers. This was also detected by the external group that does the qualitative study with focus groups. Most of the problem is due to the protocol followed for the definition of a control group and a treatment group.

Another problem was that several control group teachers used some of the exercises and strategies of the treatment group to improve their teaching strategies. For example, the qualitative study reports a teacher that said: “Sí, mucho, en especial, lo aplico, en especial en el otro curso, como ellos (4to B parte de la experiencia) ya lo manejan, aplican mucho esos ejercicios, en la sala del Conectaldea voy viendo los ejercicios, memorizo alguno, le saco foto, y lo aplico”.

Lessons

One of the main challenges was the Internet connection at schools. We used mobile internet services and special routers. However in several cases the services were not high quality. During the year the connection improved, but still was not always available.

There was a very weak maintenance support from schools and districts for the computers and related accessories. Additionally some tablets and computers were stolen at the schools. We solved replacing equipment with tablets from the project. However, we needed less amount of tablets than planned since we used some tablets and equipment from other projects.

Potential implications

The result of this project opens several possibilities. On the one hand, the Ministry of Education of Chile and the participating districts now know a learning improvement strategy based on the appropriate use of technology and gamification-type motivational strategies. On the other hand, the Ministry now identifies a local group with practical experience in evaluating the impact of educational programs. This means that the authorities and the community begin to incorporate the development strategy of implementing pilots and measuring the impact on pilots before launching large-scale programs.
Comparison between intended and actual results

The project aimed to rigorously measure the effect of appropriately using technology to improve learning. That was fully achieved. Additionally, strategies not originally planned were incorporated and adjusted, such as those of more active parent involvement and growth mindset. Both were very well received by the school community and contributed to the goal of the project.

7. Overall Assessment and Recommendations

At the beginning of the project our only suggestion and request was to adjust some parts of the budget, reassigning resources from computer equipment to communication and diffusion of the project.

We have also realized the importance of parents for the social motivation strategy. We started to visit parents meetings at school at the end of the first semester. We learned that they can be an important motivational force and that they can be helped to change attitude to a growth mindset. We recommend to consider this communication strategy and active participation of parents from the beginning of the project.

The support of IDRC was critical for this project. On the one hand, IDRC financial support and advice was fundamental to carry out the project. On the other hand, IDRC suggestion to conduct an independent qualitative study was very important. The qualitative study team alerted us about initial communication problems, which allowed us to handle them in a timely manner.

On the other hand, the partnership with IDB was fundamental. It allowed us to learn to perform a more balanced design of the experimental measurement and to complement and perfect our knowledge of analytical evaluation strategies. This collaboration has made it possible to greatly improve the capacity of the local CIAE team to carry out evaluations of educational programs.

References:


Bellei, C. "Does lengthening the school day increase students’ academic achievement? Results from a natural experiment in Chile." Economics of Education Review 28, no. 5 (2009): 629-640.
