A Cognitive Tutoring Strategy for Math Teaching and Learning in Latin America

ABSTRACT

Cognitive tutoring technology has been extensively used in the United States, where it has proven effective for increasing students' mathematics skills. However, reports of the usage of this technology in other regions are scarce. Given the urgent need to improve public education in Latin America, we have adapted and experimented with this technology in three countries (Chile, Mexico, Ecuador) to learn if it is effective for mathematics learning at the middle school level (5th - 8th grade). We discuss difficulties, assessments, and lessons learned in the process of training teachers in the use of the new pedagogical strategies supported by this technology.

Keywords: Intelligent tutoring systems; interactive learning environments; teaching/learning strategies; country-specific developments; evaluation of CAL systems; teacher's training.

1. Introduction

It is widely recognized that the incorporation of information and communications technologies (ICT) in the teaching-learning process is an innovative way for leapfrogging many of the obstacles currently plaguing educational systems across the world: it is a means of improving access to education and learning in a shorter time span than traditional approaches, and doing so within traditional budgetary limits (Luckin, Blight, Manches, Ainsworth, Crook, & Noss, 2012). Several technology driven initiatives to improve education in Latin America (LA) have been introduced with varying degrees of success (De Ferranti, Perry, Gill, Guasch, Malloney, Sánchez, & Schady, 2003; Scheurmann & Pedró, 2009; Chong, 2011). In general, the main focus of these initiatives has been two-fold: providing both basic technology infrastructure and computer literacy (for teachers and students) through training.

Even though there have been advancements in the access to information and electronic educational materials, the promises of ICT to achieve improved learning have not been fully accomplished as yet (Luckin, Blight, Manches, Ainsworth, Crook, & Noss, 2012; Wellington, 2005; Tamin, Bernard, Borokhovski, Abrami, & Schmid, 2011). This is so because the majority of the reported experimentations have placed more emphasis on the technological innovation rather than the transformation of the teaching and learning practices. In fact, performance in national and international tests (e. g. PISA, TIMMS) in the Latin American (LA) region has not increased as expected, and in some cases has even deteriorated (Hanushek, & Woessmann, 2009).

Certainly improving education is a multi-dimensional problem and technology alone does not make a difference (Luckin et al., 2012; Wellington, 2005). Moreover, many authors have indicated that, in technology-driven initiatives, the focus should be on changing the human and social systems rather than on the technology itself (Warschauer, 2003). In the social sciences, there is an extensive research literature on "social technical systems" (Goodman, 2005; Goodman, & Haran, 2009): the basic idea in this body of research is that it is not technical or social systems by themselves that will improve effectiveness. Rather, it is the integration of social and technical systems, through joint optimization, that will lead to effectiveness.

For example, there is a worldwide study of auto plants (Goodman, 2005). This study shows that the most automated plants were not the most effective. Rather it was those organizations that linked the new technology and new social arrangements that were most successful. There are hundreds of studies of failed technological interventions because they failed to adapt the social and technological systems (Goodman, & Haran, 2009). In the specific case of educational systems enhanced by technology, the teacher, classroom, school, district authorities, and parents are part of the social system. The challenge is to both redesign these social arrangements (e.g. how teachers teach) and the technology system, in order to fit and improve students' capacity in the chosen area.

Many reported experiences on introducing and assessing technologies at primary, secondary, and tertiary educational levels in different countries indicate that teachers and the teaching strategies they use remain at the heart of the educational system and no technology can substitute for it (Fritz, & Carrie, 2009; Goodman, 2002). Following these experiences, our research focuses on the teaching-learning strategies that can be enhanced by the technology (assuming that the technology has been correctly deployed[1]). This involves substantial change in the teacher's attitude, motivations, activities, and plans. The teachers need training and time for planning the new classroom-lab strategies. It involves major changes and it is a complex task. The school itself also needs to change: support from higher authorities is essential.

[1] We will see later on that this assumption is, regretfully, not correct in some cases of our experimentation in LA public schools.

1.1. Cognitive Tutor Technology

Following the theoretical principles developed by Anderson (Anderson, 2002; Koedinger, Anderson, Hadley, & Mark, 1997), cognitive tutoring software (CT) was built at Carnegie Mellon University and is maintained and operated by Carnegie Learning Inc.[2] It runs on a UNIX server platform with Windows clients. Each student has a personalized "problem-solving" space (account), with just-in-time feedback and detailed tracking of his progress (Ritter, 2011). Cognitive tutoring follows a personalized self-paced approach, allowing students to sequentially tackle progressively more difficult tasks and freeing up teacher's time to work with students that experience difficulties in their progress. It is a computer-based, interactive technology that tracks students in real time as they answer questions, ask for help and solve problems. It provides personalized feedback and hints when errors/questions are made in key points (Koedinger, & Aleven, 2007).

Cognitive tutors have shown considerable potential, and evidence in the open literature indicates that they are effective in improving mathematics and science problem-solving skills (Ritter, Anderson, Koedinger, & Corbett, 2007; Koedinger & Corbett, 2006). Specific mathematics cognitive tutors have been used in large school systems (primary/secondary level) in the United States, including Los Angeles and Chicago, as well as in rural areas; more than 300,000 secondary school students have benefited from such interventions (Arroyo, Woolf, Royer, Tai, & English, 2010). A critical lesson learned is that cognitive tutoring requires extensive and constant support from the teaching staff and/or parents: if they are absent during the educational process, the use of the technology may not produce the desired results within the expected timeframe.

As shown in Figure 1, CT software components include the following: database system for the management of math contents; Web-based interface that displays user interactive screens; problem-generator module for the assembling of different types of problems with randomly generated data and solver (with hints for the student when needed); tracking module which registers students' progress; meta-cognition module ("skill-o-meter") that provides feedback to the student about skills achievements; and a syllabus model for the creation of course plans and agendas. The next section describes the strategies that guide the software implementation.

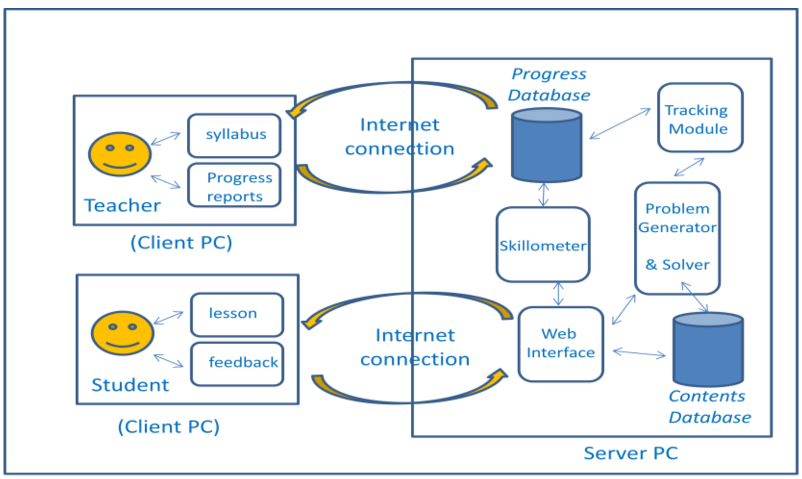

Fig.1. CT Architecture and Software Modules

1.2. Cognitive Tutor Strategies

The main objective of the CT software is to provide each student a unique, enriched environment where he can interact with the system in the solution of specific problems. Multiple graphical representations can be explored by the student for creative thinking practice (Wegerif, MacLaren, Chamrada, Scheuer, Mansour, Miksatko, & Williams, 2010; Feenstra, Aleven, Rummel, & Taatgen, 2010; Rau, Aleven, & Rummel, 2009).

A problem is presented and a solution is requested. Instead of jumping to the final solution, the software provides step-by-step scaffolding. This divide-&-conquer strategy asks specific questions, from more easy to more complex, so the student can advance at his own pace in the solution of the problem, with just-in-time automated aid if necessary.

The first question in each problem presented to the student is always related to the appropriate reading of the problem narrative. This can be particularly helpful in LA public schools, where student's problem solving abilities are many times hindered by their lack of reading skills. The next questions (posed by the software) guide the student in the solution of the problem. The student thus learns "by doing" (constructively) the various steps to produce the final solution of the problem.

Similar to traditional books, the CT software contains different modules/chapters with the contents for math classes for primary and secondary levels. The teacher can define the syllabus for the course, choosing and mixing chapters according to his

[2] Cognitive tutoring technology is a trademark property of Carnegie Learning Inc. (http://www.carnegielearning.com).

(or school's) criteria, following a "less-to-more complex" approach. The teacher can choose to force the successful completion of one chapter before the student can proceed to the next one.

In addition to the on-line conceptual book-like contents, the CT software provides a problem-oriented interactive environment where the student can build solutions for different types of problems related to the subject matter. Problem types are presented in increasing order of complexity so that students can overcome simpler situations before getting into more complex ones. Due to the random-based problem generator, the student does not go through repeated problems. At the beginning of each problem solving session, the student is presented with the typical solutions of the different problems types; the student can always come back to this explanation at any moment during the work session.

Each student has his own personalized working space and the system tracks individual progress (there should be one computer per student in the lab). By means of a problem-generator module, each student is presented a slightly different situation (random generation of parameters within problem types). Even though some patterns of problem-solutions could be decrypted (by smart kids), the random mechanism of the problem generator algorithms would avoid automation of responses (Blessing, Gilbert, Oureda, & Ritter, 2009).

If the student provides a wrong answer in a given step, the software provides a hint to aid the student in the solution of that step. Hints are graduated from minimum to maximum help, and the solution itself is given away when the student surpasses a pre-defined threshold of mistakes. Since it is not an assessment environment, hints can be accessed at will at any moment. The accessed hints are also graduated, with the final, "bottom-out" hint again essentially providing the answer. The system keeps track of the number of mistakes/hints of each student. If a student "solves" problems by means of "trial & error," his increased number of mistakes will be reflected negatively in the skills achievement record.[3]

The student gets feedback (positive or negative points in a roster of skills to be achieved) whenever he answers questions within a problem. Once a sufficient number of problems of a specific type are solved (with no errors or hints), the student is given the "achieved" status for that skill. This meta-cognition approach presents the students with the roster of skills he has to achieve at the beginning of each chapter. The student gets immediate feedback on the "skill-o-meter" at any time during his work session (Mitrovic, Ohlsson, & Barrow, 2013). This is depicted in Fig. 2.

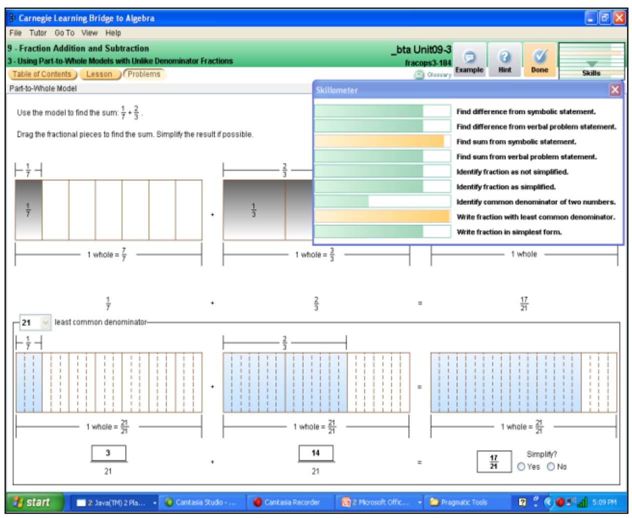

Fig. 2. The "skill-o-meter" as meta-feedback for the student

1.3. The Teacher's Role

Experience in using the CT in the United States indicates that lab work should not completely replace the regular classroom sessions. A distribution of around 60% of the time in the classroom and 40% in the lab has shown good results (Ritter, 2011). Table I shows strategies and activities for teachers and students both in the classroom and the lab.

The teacher's role in the computer lab is to motivate students in the use of the software and to supervise their progress. The teacher does not intervene on the student's computer, but is available for support when requested. The teacher's role here as a motivator and as a guide is of foremost importance. The teacher must have a thorough knowledge of the strategies and tools offered in the CT. In order to do this, he must have participated in the corresponding training and instructional design process.

The teacher's role in the classroom follows a more traditional approach, going through the subject contents in preparation for the lab work. Collaborative group work is recommended here. Using feedback given by the software tool, the teacher can

[3] There is extensive literature with thorough description and analysis of the cognitive tutoring technology (Koedinger, Anderson, Hadley, & Mark, 1997; Koedinger, & Aleven, 2007; Ritter, Anderson, Koedinger, & Corbett, 2007; Arroyo, Woolf, Royer, Tai, & English, 2010; Anderson, 2002; Ritter, 2011; Feenstra, Aleven, Rummel, & Taatgen, 2010); Rau, Aleven,, & Rummel, 2009).

define workgroups with an appropriate mixture of more advanced students and those trailing behind.

An innovative feature of the CT technology is that it provides diagnostic data that can be used by teachers to identify obstacles for learning and possible training interventions (perhaps tied to different educational strategies) to facilitate problem solving (Luckin, 2008). This meta-data is the heart of the system. If it is not used for remedial actions, there will be no substantial progress. Our current work is on the development of a decision support system that helps teachers on the selection of remedial activities.

TABLE I
MCT-LAC System: Classroom & Lab Strategies

MCT System	Classroom	Lab/Software
Class/Lab Time Distribution	60% of time(aprox.)	40% of time(aprox.)
Evaluation	- Facilitated by teacher.	- Formative evaluation. - Facilitated by software, guided by teacher. - Immediate feedback, meta-cognition.
Teaching/Learning	- Collaborative learning. - Teacher facilitates group interactions.	- Software facilitates personalized tutoring - Individually adaptive. - Progress feedback to teacher
Problem Solving	- Group work with paper & pen - Instructional materials	- Software interactions - Personalized, just-in-time hints/help
Multiplle Representations	- To interpret and justify solutions	- To set up and solve problems
Acquisition and retention of new knowledge	- Individualized and shared	- Individualized - Learn by doing

The teacher also has access to the “skill-o-meter” both at an individual and full class scale. The teacher knows at any time where individual students are standing and thus can give them, in the traditional classroom, reinforcement on the topics of struggle (Schnaubert, Andres, Narciss, Eichelmann, Goguadze, & Mellis, 2011).

1.4. Purpose of the Present Study

The broad objective of our experimental research is to understand how an educational technology that has been successfully used in developed countries (United States and Canada) can be adapted, contextualized, and integrated in a human and social system of a developing region, specifically the public educational sector of several LA countries. In this case, the teacher, classroom, school, district authorities, and parents are part of the social system. The challenge is to both redesign these social arrangements (e.g. how teachers teach mathematics) and the technology system to fit and improve students’ capability in mathematics. (Focusing solely on the technology would be a gross mistake.)

We want to learn about the different problems that are encountered, not only related to the deployment of the technology, but also in the process of changing authorities’ perceptions, teachers’ strategies, and students’ attitudes. Is it possible to build some experimental cases where, with an innovative technology-supported strategy, we can improve student’s math skills in spite of shortcomings such as lack of resources, teacher time, teaching skills, math expertise, and student motivation?

The key question of this experimental research is the following: can we motivate teachers in LA to change their teaching practices and take full advantage of the potentials of the CT technology? Second, once the teacher is motivated, does he get the resources and support from the school authorities to change his practice? Finally, can motivated teachers induce students to work with the CT technology and improve their math skills?

This paper describes experimentation in public middle schools (5^{th} to 8^{th} grades in a K-12 system) of several LA countries with a teaching strategy based on the CT technology for math learning: we call it the MCT (math cognitive tutor) initiative. A pilot implementation was performed during 2009-2010 in a number of public schools in Chile, Ecuador, Mexico, and El Salvador. Additional extended experimentation continues in Chile (2011-2014) with governmental support. We report results of an assessment study conducted to answer the research questions and evaluate the impact on student’s math performance and motivations. The experimentation shows that improved learning depends directly on the motivation of teachers to change their teaching practices and use CT as a tool to guide and scaffold students’ learning processes.

The overall strategic conception is depicted in Figure 3. In this “strategy-driven” model (as opposed to “technology-driven”), the technology plays an important (but not the main) support role. It empowers the capabilities of the teacher and enriches the learning experience of the student, especially if it is used in the interactive, dynamic environment for which it was designed. In order to advance in this direction, our methodology starts with teacher training and instructional design; technology deployment comes as a final step.

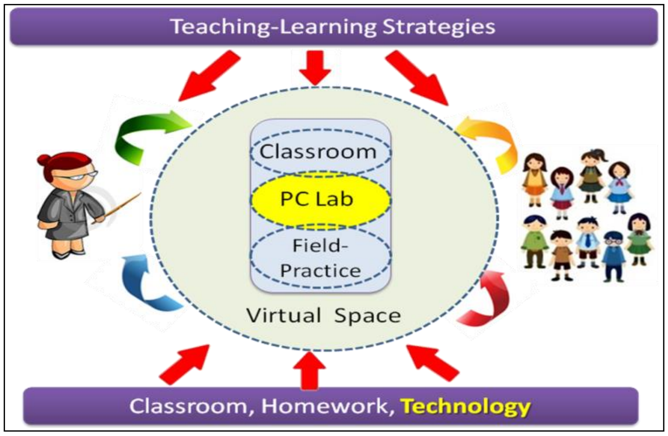

Fig.3. MCT model: technology is one (among several) means for teaching and learning.

2. METHODOLOGY AND METHODS

Building from the experiences in the United States, the MCT initiative seeks an important innovation: the definition and application of new teaching strategies that, based on the CT technology, are adapted to the LA educational idiosyncrasy. This starts with the negotiation of change strategies with the district and school authorities. It follows with the involvement of teachers on training and instructional design courses based on the CT. It culminates with the implementation of the technology-supported strategies in the math classroom.

Assuming sufficient infrastructure and resources in the school, teacher involvement is the most critical issue in the implementation plan. The goal is to achieve high motivation and strong commitment towards the new strategies. As we will show, the method offers considerable potential to improve learning in mathematics at the primary/secondary levels, long deemed to be among the weakest areas of education throughout the LA region (Hanushek, & Woessmann, 2009).

Together with the definition of new strategies, we take an English version of the software contents and, considering cultural and idiosyncratic differences, transform it into a Spanish version, coming up with a common structure for the entire LA region. Then special care is given to the alignment with the official math curriculum in each participating country.

Even though the underlying theory and structure of the software tool remains the same as in the English version, contents and exercises are localized to the local cultures. Extensive re-programming of the software is required. Special emphasis is given to the analysis of meta-data provided by the software (and other sources) for the definition of remedial actions involving laggard students.

2.1 General Implementation Plan

The specific plan described herein focuses on training (and motivating) teachers in using non-traditional strategies (collaboration in the classroom and personalization in the lab) based on a cognitive tutoring technology that has been adapted and localized for LA schools. The plan was initially designed for interventions in four LA countries, but one of them, El Salvador, could not fully implement it mainly due to lack of resources and support from educational authorities.

The principal components of the plan are:

1) Design of the intervention and experimentation, negotiated with corresponding educational authorities.
2) Adaptation of CT software modules and contents.
3) Specialized training for math teachers and design of instructional units.
4) Local implementation in selected public schools of different LA countries within a two stage pilot program.
5) Creation of local learning communities and a repository of technology enhanced learning materials for mathematics.
6) Assessment data collection and evaluation study.

An initial planning-design stage was accomplished during 2009 and training-implementation activities were performed in 2009-2010 in primary/secondary schools of Chile, Ecuador, and Mexico. Between these three countries, over 730 students and 24 teachers participated in the experimentation.[4] Additionally, a reduced pilot experience was performed in El Salvador (one school, seven teachers, and 77 students). A second experimentation (planning, design, and implementation) is currently underway on 30 public schools in Chile (2011-2014).

2.2 Component 1: Design of the Intervention Plan

Public schools (as opposed to private) were chosen because of the enormous quality gap between the private and public educational system in the LA region. Teachers are better paid in the private system. Private schools have many more resources

[4] Control schools and drop outs (including teachers and students) are not considered in these figures.

and offer little or no access opportunities to the socially vulnerable population. Our plan is guided by the conviction that the MCT system, in the long run, can help reduce the existing gap, provided that governments and private benefactors are willing to invest within normal educational budgetary limits.

In each country, a district/municipality with a vulnerable population was selected which was willing and able to participate in the experimentation. Within that district, experimentation schools were randomly selected among those that did have the resources (computer labs, time for teacher training, and support from their authorities). Teacher motivation was not taken into account at this stage. (Some of the selected schools did drop off during the experimentation due to lack of motivation, amongst other reasons.) Control schools were then selected within the chosen district/municipality (math performance on national tests and socio-economical parameters were similar to the experimentation schools).

Since there were limited resources to build the software contents, a choice had to be made with respect to the grade where the MCT system would be applied. Since the basis for algebra is in 5th to 8th grade, this was the chosen level of intervention. The material would be considered pre-algebra in the United States.

Thereafter, a plan was negotiated with the district and school authorities to define the period of the academic year when the MCT system was to be used and monitored. The experimentation (classroom and lab activities) ran for five months.

2.3 Component 2: Adaptation of Software Modules & Contents

This component focuses on the development of 14 lessons (comprising 54 instruction modules in the MCT software) of pre-algebra fundamentals used over a five month period of experimentation in the 5th – 8th grades. The content of each of these modules was internationally referenced and normalized locally to each country's curriculum.

Activities at this stage included:

1) Revision of curricula and assessment of teaching strategies used in math in each country. This revision and assessment was done by an MCT team of international consultants. Results from this review and assessment were compared between countries and benchmarked to international standards of learning in mathematics (e.g. to the Program of International Student Assessment, or PISA).
2) Development of CT modules and supporting documentation (teacher's guide, student book, and exercises). The MCT team adjusted existing modules or, where relevant, produced new modules. In some instances, this included the translation and metric conversions; in others, it entailed the development of new modules.
3) Adjustment of cognitive tutoring modules to specific school contexts. The MCT team worked together with local teachers and authorities to make necessary adjustments to the cognitive tutor modules, ensuring their relevance and usability at the school level. This included changes in examples presented or language used, localized to each country.

2.4 Component 3: Training of Teachers and Design of Instructional Units

Training teachers on the new pedagogical strategies and technology was of paramount importance. This included negotiating the integration of CT modules within the existing curricula, defining corresponding learning activities, learning self-paced pedagogies and methodologies, and monitoring and assessing students. There was also community outreach and technical support included. Part of the training endeavor was devoted to working with teachers in the design of instructional units. To provide insight on the new strategies and obtain support for the intervention, we also worked with school planners and administrators and, where relevant, with district educational authorities.

From the initial group of trained teachers, top performers were selected and given training in mentoring, thus creating a pipeline of teacher trainers with the capacity to train others if the initiative were to be extended to other schools. Specialized training in technology management and planning was also provided to school directors. Additional training was given to technical support staff in schools with responsibilities for maintaining ICT infrastructure.

2.5 Component 4: Pilot Implementation in Two Stages

The intervention was planned in two stages. With the purpose of revising the strategies, class/lab methodologies, and contents, a small scale pilot with one or two schools per country and limited MCT lessons was first conducted during 2009 while other MCT modules were being developed. Afterwards, a full scale pilot (with 54 MCT instruction modules) was conducted during 2010. The following activities were performed during the extended pilot (second stage):

- Revision of class and lab infrastructure at each experimentation school, to make sure there was one PC per student, appropriate local connectivity, and Internet access.
- Teachers were provided with the corresponding support documentation (guides, books, and exercises).
- Email, chat, and phone support was provided to teachers during all the experimentation.
- On a weekly basis, the CT reports (with student and course progress) were discussed with the teachers so they could plan in advance the collaborative activities in the classroom.
- Random monitoring of classes and labs was performed periodically.

2.6 Component 5: Creation of Local Learning Communities

This component sought to increase the capacity of math teachers to use new forms of ICT technology to motivate students to pursue scientifically oriented careers. To do so, the project included the creation of local learning communities within and between the schools, thus stimulating the exchange of experiences and best practices.

The MCT specialists worked with groups of teachers both within schools and system-wide to share new ways of teaching math and to create deeper understanding of effective approaches, including project-based activities and the use of different technology enhanced "learning objects" - such as multimedia components to demonstrate concepts, games, and simulations - in the classroom.

Through these activities, a repository of best practices was created and used to promote knowledge exchange among teachers. In the future, the learning community setting will be used to collect, maintain, and analyze basic project data, and to support decision-making at the school/district level.

2.7 Component 6: Data Collection & Assessment Study

This component focuses on the assessment of project-induced changes in schools, teacher practices, and student performance in math. It supports the design of an assessment methodology for capturing and evaluating change through both qualitative and quantitative measures.

Pre-tests and post-tests (pencil-paper based, similar format to national tests) were conducted at the beginning and at the end of the intervention, both at experimental and control schools. Additionally, at the end of the experimentation phase, *s*urveys and interviews were conducted for students, teachers, and school authorities.

In the specific case of student performance, this methodology incorporates a quasi-experimental design. Data collected through this component serves to evaluate changes at the school level as well as the overall effect of the project across countries. The assessment methodology was designed based on field experimentation literature[5] (Goodman, 2005). More about the assessment methodology is presented in the next section.

The MCT team also provided specialized training at the school level on the application of the assessment methodology and tools, maintained responsibility for evaluating the impact in learning across countries, and documented each intervention in a way that allows for eventual replication in other contexts.

3. ASSESSMENT STRATEGY

The objective of the MCT evaluation process was to determine the effectiveness of the use of a cognitive tutor system to improve skills in mathematics (specifically in pre-algebra) for 5th to 8th graders in public schools in selected LA countries. We also wanted to understand the motivations (and lack thereof) of teachers, students, and authorities to change their teaching-learning strategies.

3.1 Meaning of Effectiveness

Effectiveness is a multidimensional construct. Rarely is there one ultimate indicator. In most assessments in educational or organizational settings, there are multiple indicators for different constituencies, and these indicators may represent different points in time.

For example, in the context of this project, we could simply look at scores on some standard test. We also would have data on response latency, number of hints used, incremental learning over time, and other measures. All of these metrics would be relevant indicators of effectiveness. But in a very different context, the teachers' attitudes and behavior are other indicators of effectiveness.

However, if the teacher had a successful class with this technology, but would not want to use this technology again, would this be an indicator of ineffectiveness? We think it would. Therefore, our approach to the question of effectiveness does not simply rely on improvements in math, but it is also related to the school and the corresponding district/municipality changing their strategies and adopting the technology for use over time.

3.2 What is the System for the Assessment?

We start at the classrooms where the MCT system is implemented. The relevant players are the students and teachers. For the students, we are interested in changes in individual math performance. But we are equally interested in changes in teacher behavior and attitudes. If the teacher held on to traditional practices, this intervention would not work (as indicated in the research literature (Luckin, et al. 2012; Wellington, 2005; Tamin et al., 2011)). We know from the same literature that the use of technology without a teaching-learning change process provides a minimal added value. Even if student performance increases, lack of peer support will reduce the effectiveness of the intervention. Its long-term viability will be questionable.

In addition to the classroom, the reaction of the teachers' peers will be important. Changes in one part of the system rarely survive if there is not support in other parts of the school. This would be particularly true for the schools' principals, as well as for the technical support people. In previous assessments, these people were critical to the intervention. The understandings of educators and authorities from the government sector will be important for longer-run effectiveness.

[5] Standardized and validated evaluation instruments (pre and post tests, surveys, interviews, and case studies) used in the study are described in detail in specific project reports presented to the Inter-American Development Bank.

3.3 Field Experimentation vs. Lab Experimentation

The proposed MCT assessment takes place in the field. This is substantially different from the traditional laboratory experiment. In that case, one can randomly assign subjects to the experimental and control groups. The advantage of the laboratory experiment is that it maximizes on internal validity. Its weakness is that it does not facilitate external validity. We could have used the MCT system in an experimental lab setting, but the fundamental question of how it would actually operate in a school would remain.

In fact, we have learned that public schools are most willing to participate when the objectives of the project are clearly beneficial for teacher and student performance. This is clearly the case within a "field experiment" such as the MCT initiative.

First, a critical feature of any experiment (lab or field) is to select control groups. In the lab experiment, students sign up, volunteer, or get paid. There the experimenter randomly assigns them to different groups. In doing field experiments in real organizations, we need to get organizations and the people to participate. That initially creates some bias but, within this context, there still can be pseudo-experimental and control groups. We obtained these through random selection of our "willing and able to participate" set of schools.

The major issue with experimental groups is not whether they have been randomly assigned. In our experimentation, the central issues are determining the major variables that might affect math performance, independent of the technology, and figuring out to what extent these are comparable across groups. For example, we would have data on prior math performance of the experimental and control groups, or we could look at the experience, performance, and attitudes of the teachers.

The challenge then is to identify critical predictors of math performance and be sure they are comparable or controlled statistically across experimental and control classes.

Another issue concerns the selection bias inherent in having schools volunteer to participate. The question is whether this is a bias or a necessary "entrance" condition. Given our earlier comments on socio-technical systems, we would argue it is a necessary condition. The technology will not work by itself (if the school and/or the teachers are not motivated). We need voluntary and motivated participation that signals the conditions for a positive or receptive social environment to complement the technology. Otherwise, there will be no changes in math performance.

If we think about the broader question – can this specific strategy-oriented technology improve math performance compared to more traditional methods – then the issue of selection bias gets reframed. We need schools and teachers willing to participate. If we can demonstrate positive results in these types of schools, then more schools would be convinced and will want to participate, with government assistance and incentives. Control groups still would be necessary.

TABLE II
MCT Implementation in LA region: Experimental & Control Schools per Country

	Treatment (no attrition)				Control		
	Schools	Classrooms	Teachers	Students	Schools	Classrooms	Students
Chile	7	14	8	388	7	14	400
Ecuador	6	6	12	190	6	6	206
Mexico	4	4	4	156	4	12	478
El Salvador	1	2	7	77	1	3	113
TOTAL	**18**	**26**	**31**	**811**	**18**	**35**	**1197**

3.4 The Sample

Table II shows the number of participating (treatment and control) public schools, students, and teachers per country, which provides a sound population for assessment. As mentioned before, districts were selected by their willingness to participate and possession of basic "entry" requirements such as number of PCs per lab, local networks, tech support, Internet connectivity, and authorities' support. Random selection could not be implemented at this first stage of the project, but was used to categorize treatment and control schools.

In general, the selection of the participating districts was a somewhat difficult process. It is obvious that without full support and involvement of the district authorities, the experimentation was impracticable. There were some initially invited districts that were necessary to discard due to their lack of real involvement. Finally, only one district per country participated in the project. All schools within a district were invited to participate, but only a few of them decided to fully experiment with the MCT system. For instance, the selected Chilean district had 25 schools and only 17 of them were willing to participate. Of those 17, ten were assigned to treatment and seven were assigned to control. During the implementation process, three of the experimental schools dropped out for different reasons: problems with infrastructure, lack of involvement in training, reluctance toward teaching changes, and lack of support of school authorities.

Due to the training process - conducted locally in each country following the guidelines of a global strategy - most participating teachers were enthusiastic and willing to adopt the new strategies and technology. A few teachers (about 20% of initial participants) didn't have enough time to complete the training or were simply reluctant to change their methodologies. The later ones constituted drop-outs from the experimentation and in some cases the school as a whole could not participate. Thus, the training activities (performed as university courses) were a good means to filter participants. Attrition schools and teachers were not included in the assessment study. In every case, they dropped out very early in the intervention, most without even using the MCT in their classrooms at all.

3.5 Dimensions & Measures for Assessment

A descriptive transactional methodological design was used for the assessment of the pedagogical practices and learning processes. It integrates qualitative and quantitative methods for the compilation of data: surveys, interviews, classroom observations, and case studies were used.[6] The following dimensions and instruments were considered in the present evaluation study:

1) Math performance of participant (experimental group) and non-participant (control group) students: pre test and post test focused on the contents covered by the experimentation.
2) Performance indicators collected by the MCT software at the student and class level: problems solved per lesson, average time per problem, error/hints percentages, mastered skills, etc.
3) Attitude of students toward the MCT system: surveys taken after experimentation.
4) Attitude of teachers toward the MCT system: surveys taken after training and after experimentation.
5) Beliefs and behaviors of participant and non-participant teachers: surveys taken after experimentation.
6) Attitudes of technical school staff and school principals toward the MCT system: interviews taken after experimentation.
7) Case (ethnographic) studies in selected schools of each participating country: direct observation of classes, interviews of school principals and student groups.

3.6 Quantitative Data: Test Results

A large amount of collected data has been extensively analyzed. We present here observations from the data collected in Chile and Mexico, which both used the Chilean national math exam as an outcome measure. An analysis of quantitative data (pre and post tests) shows the following observations:

1) The students in experimental and control schools have statistically equivalent pre-test scores.
2) However, post-test differences between the experimental and control schools indicate a statistical improvement in math scores for those who used the MCT.
3) There is a variance in math performance among experimental schools.

The primary student achievement outcome measure used in this study comes from two comprehensive, grade-level pre-algebra tests given to all the students (treatment and control). Both instruments were measured with the Alfa-Cronbach test, giving a reliability of 0.77 for the pre-test and 0.80 for the post-test.[7]

TABLE III
Statistics of the Difference Scores – Treatment & Control (Chile)

Sample	All Students	Treatment	Control
Observations	788	388	400
Average	0.28	0.95	-0.37
Std Deviation	5.17	4.99	5.27
Std Error	0.18	0.25	0.26
Maximum	19	19	14
Minimum	-22	-15	-22
Median	0	1	0
Mode	2	2	2
25th Percentile	-3	-2	-4
75th Percentile	4	4	3

One test is given near the beginning of the school year. The other is given six months later, after the experimentation with the MCT systems in the classroom and lab. Both exams consist of 44 multiple choice questions. The tests were reviewed and approved by both the Chilean school authorities and the MCT developers prior to their use in the study. The developers agreed that the material in the exams was both grade-level appropriate and covered by the software. The math material focuses on pre-algebra concepts, as does the software used by the treated students.[8]

[6] The same assessment instruments were applied across countries.
[7] A Cronbach-Alfa measure above 0.70 indicates a reliable instrument.
[8] In Chile, there was only one school with both treatment and control students; other schools were solely of one type. In Mexico, randomization occurred within schools at the classroom level. Each Mexican school had four classrooms, and one was assigned to the treatment group. In both countries, control students did not use the MCT system and their math teachers did not participate in MCT training.

Our research team compiled the exam scores and difference scores from student level results on every question from both exams. We know whether a question was answered correctly (1), incorrectly (0), or left unanswered (blank). Adding up the correct responses for each student for each exam yields the total score for each separate exam. There was no added penalty for incorrect answers. The test outcome measure considered is the difference score from the exams. The difference score is measured at the student level by subtracting the initial exam score (pre-test) from the final exam score (post-test). A positive difference score means that a student scored higher on the second exam, while a negative difference score denotes a higher score on the initial exam. Unanswered questions were considered incorrect.

Comprehensive statistics on the difference scores are shown in Tables III and IV when we aggregate across the schools of similar type (treatment vs. control). Across nearly every measure shown in the tables, treatment students have a greater difference score value than control students. In short, a quick glance at Tables III and IV supports the notion of a positive overall treatment effect.

TABLE IV
Statistics of the Difference Scores – Treatment & Control (Mexico)

Sample	All Students	Treatment	Control
Observations	634	156	478
Average	0.96	2.06	0.60
Std Deviation	5.28	5.08	5.30
Std Error	0.21	0.41	0.24
Maximum	23	13	23
Minimum	-25	-11	-25
Median	1	2	1
Mode	0	4	0
25th Percentile	-2	-1	-2
75th Percentile	4	5	4

An ANOVA test reveals significant difference in difference scores[9] (Chile: $F(1,786) = 12.99$, $p = 0.0003$; Mexico: $F(1,632) = 9.11$, $p = 0.0027$) between the treated and control populations in both countries. The 95% confidence interval from a two-sided t-test is (0.60, 2.03) in Chile and (0.51, 2.41) in Mexico. The point estimate of the difference-in-difference scores (1.32 in Chile, 1.46 in Mexico) is approximately 25-30% of the difference score standard deviation across all students within each country, a very substantial improvement. The confidence interval (CI) on the test of whether the sample average from the treated and control groups are the same does not cover zero at the 95% level.

TABLE V
Statistics of the Difference Scores by School, Treatment Only (Chile)

	S1	S2	S3	S4	S5	S6	S7
Average	0.01	1.29	3.29	0.85	0.37	-0.25	2.23
Std Deviation	5.09	5.40	5.52	5.27	4.54	5.08	4.18
Std Error	0.53	0.81	1.20	0.65	0.57	1.04	0.46

TABLE VI
Statistics of the Difference Scores by School, Treatment Only (Mexico)

	S1	S2	S3	S4
Average	5.05	2.93	-1.08	1.10
Std Deviation	4.38	4.84	4.76	4.27
Std Error	0.69	0.75	0.79	0.67

Tables V and VI show the difference scores by school (treatment only). There is significant variation even within treatment schools. Some even show negative difference scores, meaning that, on average, students performed better on the pre-test than the post-test of similar material. We hope to explain these large differences in subsequent implementations by considering MCT data and student, teacher, and school characteristics in the evaluation.

[9] "Difference in difference scores" measure the improvements of treated students over control students, using the difference score (post-test minus pre-test) as the outcome variable of interest. A positive difference in difference score means that treated students will improve their test scores more than control group peers.

3.7 Qualitative Data: Selected Survey Results

Some observations in the analysis of student and teachers surveys from Chile are the following:

1) A high percentage (**78%**) of experimental students was satisfied with their improvement of math performance due to the use of the MCT system.
2) After experimentation, a high percentage **(67%**) of experimental students increased their motivation toward learning math.
3) After experimentation, **68%** of students felt more certain about their abilities to solve math problems and decreased their perception of math as a very complex discipline.
4) **81%** of students view the MCT system as a useful tool that substantially helps their learning process.
5) **85%** of students were satisfied with the usability of the MCT software; **78%** of students want to continue using the software tool in their math classes; **88%** of students are motivated by the use of computers in their classes.
6) **87%** of students regarded the teacher as a helpful guide in the learning process using the MCT system.
7) **82%** of students would like to use the cognitive tutoring system methodology in other subject areas.
8) A large percentage (**75%**) of teachers view the MCT system as a **useful and effective** mechanism that empowers their teaching process.
9) **100%** of participating school authorities (principals and area directors) considered the MCT system a useful tool for the teaching and learning of math.
10) **100%** of participating teachers and authorities would like to continue using the MCT system in the future.
11) As a side effect (observed in the case studies), teachers positively viewed the MCT software as a review tool for the refreshment of their own math knowledge.
12) As a side effect (observed in the case studies), the MCT system was an adequate means for advanced students to get involved in helping the learning process of their peers.

Even though the analysis of experimentation results in Chile and Mexico show that math performance did seem to improve and there are very positive attitudes from students and teachers, we are looking at other factors that lead to changes as well. These factors include support from school and district authorities, parental behavior, teachers' background and math knowledge, and school infrastructure and technology support.

4 CONCLUSIONS AND FUTURE WORK

This paper describes an innovative strategy-oriented educational design and experimentation based on cognitive tutoring technology to improve math learning in public primary/secondary schools of three LA countries: Chile, Ecuador, and Mexico.

Assessment results presented herein are encouraging and show that math learning in primary/secondary public schools in LA can be positively impacted with the use of cognitive tutoring technology. A necessary requirement is the involvement of motivated teachers that implement ICT supported innovative strategies and change the traditional passive classroom methodologies. Given positive feedback from teachers, students can get highly motivated to work with the software tool and the math contents. As shown in the survey results, students are encouraged by the personalized feedback and scaffolding provided by the MCT system. They perceive sustained progress in their learning. Teachers feel empowered with a richer teaching environment and use the feedback provided by the software as a guide for remedial actions.

Most of the attrition schools and teachers were anticipated and detected early in the implementation process, by means of the described training activities. According to our surveys and interviews, **100%** of drop-outs were due to lack of resources (preparatory courses demand a significant amount of time and effort from teachers and authorities). Attrition was not included in the assessment results. The fact that **100%** of participating teachers and authorities (not including drop-outs) want to continue using the MCT system in the future is an encouraging result that shows motivation and willingness to change.

The MCT system shows promising results and future actions involving government authorities will be taken to extend its application at a larger scale. Currently, a new experimentation with the MCT system is underway in public schools of Chile, in urban and rural settings. The system is being applied at the 8th grade level and the expectations are high.

REFERENCES

Anderson, J.R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science 26*, 85-112.

Arroyo, I., Woolf, B., Royer, J.M., Tai, M., & English, S. (2010). Improving math learning through intelligent tutoring and basic skills training. In Aleven, Kay and Mostow (Eds.), *Intelligent Tutoring Systems Part I* (pp. 423-432), *Lecture Notes In Computer Science 6094*.

Blessing, S.B., Gilbert, S.G., Oureda, S., & Ritter, S. (2009). Authoring model-tracing cognitive tutors. *International Journal of Artificial Intelligence in Education*, *19*, 189-210.

Chong, A. (2011). *Development connections: Unveiling the impact of new information technologies.* Inter American Development Bank.

De Ferranti, D., Perry, G., Gill, I., Guasch, L., Malloney, W., Sánchez, C., & Schady, N. (2003). *Closing the gap in education and technology.* The World Bank.

Feenstra, L., Aleven, V., Rummel, N., & Taatgen, N. (2010). Multiple interactive representations for fraction learning. *Proceedings 10th Int. Conference on Intelligent Tutoring Systems (ITS), (*pp. 221-223).

Fritz, K. P., & Carrie L. (2009), "Applying Organizational Research to Public School Reform: The Effects of Teachers Human and Social Capital on Student Performance", Academy of Management Journal, Vol. 52, N° 6, 1101-1124, 2009.

Goodman, P.S. (2002) *Technology enhanced learning: Opportunities for change*. New Jersey, London: Lawrence Erlbaum Associates, Publishers.

Goodman, P.S. (2005). *The Rushton quality of work experiment.* New York: John Wiley & Sons.

Goodman, P.S., & Haran U. (2009). Self–Managing Teams. In Levine, J.M., & M.A.Hogg, M.A. *Encyclopedia of Group Processes and Intergroup Relations*. Thousand Oaks, CA: Sage Publications.

Hanushek, E.A. & Woessmann, L. (2009). *Schooling, cognitive skills and the Latin American growth puzzle.* Inter American Development Bank and CESifo.

Koedinger, K., Anderson, J.R., Hadley, W., & Mark, M.A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8,* 30-43.

Koedinger, K., & Corbett A. T. (2006). Cognitive tutors: Technology bringing learning science to the classroom. In *The Cambridge Handbook of the Learning Sciences* (pp. 61-78). Cambridge University Press.

Koedinger, K., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review 19,* 239-264.

Luckin, R. (2008). The learner centric ecology of resources; a framework for using technology to scaffold learning. *Computers & Education 50,* 449-462.

Luckin, R., Blight, B., Manches, A., Ainsworth, S., Crook, C., & Noss, R. (2012). *Decoding Learning: The proof, promise and potential of digital education*. Nesta Operatig Company, U.K. Available: http://www.nesta.org.uk.

Mitrovic, A., Ohlsson, S., Barrow, D.K. (2013). The effect of positive feedback in a constraint-based intelligent tutoring system. *Computers & Education 60,* 264-272.

Rau, M., Aleven, V., & Rummel, N. (2009). Intelligent tutoring systems with multiple representations and self-explanation prompts support learning of fractions. *Proceedings 14th Int. Conference on Artificial Intelligence in Education (AIED),* (pp. 441-448).

Ritter, S., Anderson, J.R., Koedinger, K., & Corbett, A. T. (2007). The cognitive tutor: applied research in mathematics education. *Psychonomics Bulletin & Review 14(2),* 249-255.

Ritter, S. (2011). *The research behind the Carnegie Learning math series* [Online]. Available: http://www.carnegielearning.com.

Scheurmann F. & Pedró F. (2009). *Assessing the effects of ICT in education. Indicators, criteria and benchmarks for international comparisons.* European Commission.

Schnaubert, L., Andres, E., Narciss, S., Eichelmann, A., Goguadze, G., & Mellis, E. (2011). Student behavior in error-correction-tasks and its relation to perception of competence. In *Towards Ubiquitous Learning Proceedings of the 6th European Conference on Technology Enhanced Learning* (pp 370-383), Berlin: Springer.

Tamin, R., Bernard, R.M., Borokhovski, E., Abrami, P.C., & Schmid, R.F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research, 81(1),* 4-28.

Warschauer, M. (2003). *Technology and social inclusion: Rethinking the digital divide*. Cambridge, MA: MIT Press.

Wegerif, R., MacLaren, B.M., Chamrada, M., Scheuer, O., Mansour, N., Miksatko, J. & Williams, M. (2010). Exploring creative thinking in graphically mediated synchronous dialogues. *Computers & Education 54(3),* 613-621.

Wellington, J. (2005). Has ICT come of age? Recurring debates on the role of ICT in education. *Research in Science and Technological Education, 23(1),* 25-39.