The Persistence of Gains in Learning from Computer Assisted Learning (CAL): Evidence from a Randomized Experiment in Rural Schools in Shaanxi Province in China

Di Mo, Linxiu Zhang, Jiafu Wang, Weiming Huang, Yaojiang Shi, Matthew Boswell, Scott Rozelle

Abstract

Background. Computer assisted learning programs (CAL) have been shown to be effective in improving educational outcomes among students. However, the existing studies on CAL have almost all been conducted over a short period of time. There is very little evidence on how the impact of CAL programs evolves over time. In response, we conducted a clustered randomized experiment involving 2741 boarding students in 72 rural schools in China to evaluate impacts of CAL programs over the long term. Our results indicate that a CAL program that was implemented for one year and a half among third and fifth grade students increased student standardized math scores by 0.25 standard deviations for third graders and 0.26 standard deviations for fifth graders. In addition, we have shown that students gained in math learning in both CAL Phase I (which ran for one semester in spring 2011) and CAL Phase II (which ran for both semesters of the 2011-2012 academic year) programs. By testing for heterogeneous effects, we find that the CAL intervention worked well for both the poorer performing and better performing students in the third and fifth grades. We also find that the third grade girls seem to have improved more than the boys in math in the short term (CAL Phase I).

Working Paper 268

January 2014

reapchina.org/reap.stanford.edu

REAP
Rural Education
Action Program

The Persistence of Gains in Learning from Computer Assisted Learning (CAL): Evidence from a Randomized Experiment in Rural Schools in Shaanxi Province in China

Introduction

In the last decade economists and education experts have studied the impact of computer assisted learning programs (CAL) on the educational performance of students in an attempt to help disadvantaged children in developing countries (e.g. Banerjee, Cole, Duflo & Linden, 2007). These CAL programs utilize modern computing technologies to enhance learning through computerized instruction, drills, and exercises in an environment where teaching and/or tutoring resources are in severe shortage (Kirkpatrick & Cuban, 1998; PCAST, 1997; He, Linden & MacLeod, 2008; Barrow, Markman & Rouse, 2008; Linden, 2008). The software delivers simple teaching functions that might ordinarily be performed by a teacher (Stahl, Koschmann & Suthers, 2006). For example, the software provides explanations about a student's curriculum and gives instructions about solving problems through animated lessons. It can also provide feedback to students by correcting their answers to the exercises and illustrating the different approaches to getting the right answers. There have been several evaluations of CAL programs in developing countries that show positive impacts on student performance (Banerjee, Cole, Duflo & Linden, 2007; Lai et al, 2011).

However, an important limitation shared by nearly all such studies is that they were implemented over fairly short periods of time and therefore have not evaluated whether program impacts persist over time. For example, a study by Lai et al. (2011) only spanned a single semester (four months). Similarly, studies by Banerjee, Cole, Duflo & Linden (2007) and Lai et al. (2012a, 2012b) encompassed less than a year (e.g., eight to nine months). In none of these studies did the research teams evaluate whether the program effect mainly took place in the first semester or whether it accumulated over the academic year. While these studies are helpful in exploring the impacts of a CAL program in the short run, they leave open several interesting questions about the nature of the effect. Is the impact of a CAL program a "one time" effect that diminishes once the novelty of computing wears off? Or, can CAL be considered a way to enhance learning that will continue to benefit students over the longer run?

The evidence on this question—short or long run—is mixed. Some previous studies point out that integrating regular class materials with interactive interfaces and computer-based games may make the learning process more engaging for students (Inal & Cagiltay, 2007; Schaefer & Warren, 2004). Such game-based learning software may increase student motivation and interest in curricula, which may in turn lead to elevated focus and motivation among students. (Chang, 2002; Cotton, 2001; Garcia & Arias, 2000; Garris, Ahlers & Driskell, 2002). However, other studies raise the possibility that the short-term gains in learning derive from the initial excitement of using a novel technology and may not be sustained over the long run (Malone & Lepper, 1987; Lai et al., 2011; Marjanovic, 1999). What is more, after being exposed to a game-based curriculum students may find regular class periods boring and disengage from teacher instruction (Cordova & Lepper, 1993). Under such circumstances, a long-term CAL program may appear to have no impact as reduced learning in regular classes might offset the gains achieved through CAL.

Unfortunately, there is little empirical evidence that can help us understand whether CAL programs can create significant and sustained gains in student learning in developing countries (Banerjee, Cole, Duflo & Linden, 2007). Many of the studies on computer-based learning/teaching programs do not allow for establishing causality due to the absence of a comparable/randomized control group (Blok, Oostdam, Otter & Overmaat, 2002). Many other studies do not have an adequate sample size and, thus, lack statistical power. In the cases in which authors have used valid program evaluation techniques to study the long-run effects of computer-based learning, almost all have been conducted in developed countries and have targeted specific populations. For instance, Günther et al. (2010) found that a computer assisted training program persistently improved cognitive abilities of elderly individuals with age-related memory deficits in United States. Also in the United States, Roesch et al. (2003) found a positive long-term impact of an interactive computer program for medical case studies on dermatology students in medical school. Sustained gains in reading were also found among low-ability readers who participated in a computer-based reading program in France (Ecalle, Magnan & Calmus, 2009).

These studies notwithstanding, important questions about the persistence of the impact of CAL programs on student learning in developing countries remain unanswered. Because the quality of teacher resources is relatively poor in developing countries (Glewwe and Kremer, 2006), and the demands on student learning are potentially greater due to rising student numbers in many parts of the developing world (World Bank, 2013), the absence of evidence on the long-term effects of CAL on students in developing countries warrants our attention. Were CAL programs to be integrated on a larger scale in developing countries, would the effects of CAL be positive and would they persist over time?

This question is particularly relevant to China. In order to narrow the "digital divide" and the educational performance gap between rural and urban schools, China's Ministry of Education has an ambitious plan to invest in the computing infrastructure of rural schools (Yuan, 2012). The recently announced 12th Five-Year Plan for Integrating Information Technology into Education aspires to set up a computer room in every rural school by 2020 (Ministry of Education, 2012). Since the plan requires an enormous investment of fiscal resources, it is important to learn whether these resources can actually be made to promote sustained learning among rural students.

To address this question, the overall goal of this study is to determine the persistence of CAL program effects on the academic outcomes of an underserved student population in a developing country. To achieve this goal, we pursue three specific objectives. First, we estimate whether a one and a half year long math-based CAL program has any impact on academic performance. Second, we compare the program impacts after one and a half years with those of a short-term program that ran for only one semester. Third, we explore the heterogeneous effects of the CAL intervention by investigating whether the treatment effects differ for different subgroups of students in both the short and the longer term.

In order to achieve our objectives, we conducted the largest and longest (in terms of implementation time) field experiment of CAL in China. The field experiment involved 72 rural schools and 2741 rural students. During the year and a half long experiment we conducted a baseline survey (before the launch of the program) and two rounds of evaluation surveys. The first evaluation was implemented one semester after the program started. The second evaluation was implemented three semesters (one and a half years) after the program started. Such a field experiment enables us to examine whether program impacts on student learning are persistent over time.

CAL Program Phase I (Short-term CAL) in Shaanxi Province

In our first attempt to implement the CAL program as a short-term activity, we conducted a clustered Randomized Controlled Trial (RCT)[1] in Shaanxi Province in

[1] Randomized Controlled Trial is a type of experiment that is often used to test the efficacy and effectiveness of an intervention on a target population. In such an experiment, subjects in the target population are randomly allocated to receive treatment or be taken as control (receive no treatment). This method minimizes allocation bias in testing the impact of the treatment on the target population.

China in the spring of 2011. A total of 5943 students in 72 rural schools in Shaanxi Province were involved in the study (36 intervention schools, 36 control). Among the students, 2741 students were boarders and the other 3202 students were non-boarders.[2] During the short-term CAL program, only the boarding students in the 36 treatment schools participated in the CAL classes while the non-boarders served as additional controls.

The short-term CAL intervention that was implemented in the 36 treatment schools ran for one semester (or around three to four months). In its most basic form (complete details of the CAL intervention are included in the next section), the CAL program consisted of a remedial, game-based CAL program in math that was held outside of regular school hours among boarding students. To test the effectiveness of the CAL program, students in both treatment and control schools were given standardized math tests before the start of the program and at the end of its implementation.[3]

According to our analysis, which is presented in Figure 1 and Table 1 (and also in Lai et al., 2012a), the short-term CAL program had a positive and significant

---

[2] Boarding students are those students that live in a school-run dormitory between Monday and Friday of each school week. Boarding is optional (occurs at the choice of the parents) but is often necessary because the student's home is so far away from the school that commuting is infeasible. If a student does not board, he/she lives at home with his/her family and is called a non-boarder in this study.

[3] The standardized math test included questions from math exercise books that are available in bookstores. The questions were chosen by education experts (testing specialists) in primary schools in China. The questions were chosen to test the math knowledge and skills that students should learn/master based on the national curriculum. The tests were administered in the same manner to all sample students (in both treatment and control). Different tests were given to students in different grades.

effect on the math test scores of students in the treatment schools. Overall, scores went up by 0.12 standard deviations (significant at the 5 percent level, also see in Table 1, row 1, column 1). Table 1 includes the results of the regressions and shows that when using a full model (see the section below for more details on the exact specifications) the impact on the third grade students was 0.18 standard deviations (significant at the 5 percent level, row 1, column 2) and the effect on the fifth grade students was 0.07 standard deviations (although not significant at the 10 percent level, row 1, column 3).

Despite the positive result in Figure 1 and Table 1, there remains the question of whether the impact of the Phase I short-term program would persist over a longer period of time. If the entire school system were to adopt this program, it would be essential to first learn whether the observed findings represent only a short-term impact or whether the program effect can in fact be sustained. That is why we designed the longer-term second phase of the experiment.

Sampling, Data and Methods for CAL Phase II (The Longer-Term CAL Study)

Sampling and the Process of Randomization

For Phase II of the experiment, we conducted a clustered (at the school level) RCT of the CAL program in Shaanxi rural schools during the entire 2011-2012 academic school year. Each academic year is divided into the spring semester and the fall semester with four months per semester. The second phase of the CAL program was implemented as an extension of the first phase, which ran for the duration of the spring semester of 2011 (as discussed above). As such, CAL Phase II included the same 2741 boarding students in the same 72 rural schools in Shaanxi Province. The sample students were in third and fifth grade when they joined Phase I in the spring of 2011. During Phase II (October 2011 to June 2012) the students were in fourth and sixth grade.

Choosing the sample consisted of several steps. First, to focus our study on poor rural students, we restricted our sample frame to four counties randomly selected out of the ten counties in Ankang Prefecture, an administrative area that covers one of the poorest parts of southern Shaanxi Province. Shaanxi Province is a large (a population of nearly 40 million), rural (more than 60 percent of the population lives in rural areas) and poor province in northwestern China. In 2011, the average per capita income of these four counties was only around 4000 RMB (around $600) per year, compared to rural China's average per capita income of 6977 RMB the same year (CNBS, 2011). After selecting the counties, we obtained a comprehensive list of all *wanxiao* (elementary schools with six full grades, grade one through grade six) in each of the four counties in Ankang Prefecture. We selected all 72 schools that met these criteria to be our sample.

Within the sample of schools, we initially (during Phase I) included both third grade and fifth grade students. We chose students from these grades for several reasons. First, at the time of the launch of the project (spring 2011), we only had remedial tutoring materials for grades 3 to 6 and thus did not choose students from the first or second grade. Second, given the limited number of computers in each school's computer room and the scheduling constraints of boarding students, the CAL program could only accommodate students from two grade levels. We excluded the sixth grade students from consideration because they would have graduated before the Phase II program had begun. Being two grades apart, the third and fifth graders could also offer a sharper comparison of the intervention effects by age group. None of these students had ever participated in a CAL program prior to the spring semester of 2011.

All boarding students in the 72 sampled schools were included in the sample. In the spring semester of 2011, there were a total of 2741 boarding students in the sample, of which 1167 were in the third grade and 1574 were in the fifth grade (Figure 2).

Although there was some sample attrition by the end of Phase II, it is unlikely to have had a large impact on our study. Due to school transfers, illness or injury, 11.5 percent of the students in our sample attritted between the baseline and endline surveys. This attrition rate is low compared with other experiments conducted among primary schools in developing countries (McEwan (2013) reports that the average attrition rate of 76 experiments among primary schools in developing countries is 20%). By the time of the final evaluation of the Phase II program, we were able to follow up with 2426 boarding students in the 72 sample schools (Figure 2, final row).

There were 129 attrited students (11.1 percent) from the third grade and 186 attrited students (11.8 percent) from the fifth grade.

Fortunately for the study's integrity, there was almost no systematic relationship between the treatment and attrition status or student characteristics and attrition status (Table 2). In other words, among the attrited students, there were no characteristics/student-level variables that were correlated (in a statistically significant way) with the treatment/control status of the students (column 1). The treatment students were as likely as the control students to attrit, and attrited students had similar characteristics in both groups.

We randomly chose 36 schools from the 72 schools in our sample to receive the CAL intervention. All of the 1277 boarding students in the third and fifth grades of the 36 treatment schools constituted the treatment group (Figure 2). Among these students, there were 554 third grade students and 723 fifth grade students. The 1464 boarding students in the same grades (613 from the third grade and 851 from the fifth grade) in the other 36 schools served as the control group. Due to attrition, there were 2426 students left in our final analytic sample, among whom 1151 were in the 36 treatment schools, and 1275 were in the control schools (Figure 2).

The balance of the sample across treatment and control groups was also even (Table 3). To show this, we used a set of student characteristics to check the validity of the random assignment. In doing so (and as is standard in the program evaluation literature), we regressed the treatment variable (*whether the student received CAL* *treatment or not*) on the characteristics of the students. According to our data, none of the differences in student characteristics between the treatment and control groups were statistically significant (columns 1 and 2). In addition, almost all the differences between treatment groups are small in magnitude.

Intervention

The main intervention involved computer assisted math remedial tutoring sessions, which were designed to complement the regular in-class math curriculum for the spring semester of 2011 (Phase I) and the entire school year of 2011-2012 (Phase II). During this program, the CAL sessions were given to the students under the monitoring of two teacher-supervisors trained by our research group. The students in the treatment group participated in two 40-minute CAL sessions per week. The sessions were mandatory and attendance was taken by the teacher-supervisors.

The content (instructional videos and games) of each session was designed to help the students reach basic competencies in China’s uniform national math curriculum. The software-based lessons were exactly the same for all students of the same grade in each of the treatment schools.

During each CAL session, students sat at computers and played math games designed to help them review and practice the basic math material that was being taught in their regular school math classes. The CAL teacher-supervisors arranged for the students to sit in pairs, with one pair of students sharing a single computer. The students shared one pair of ear buds so that each could hear the voices, music, and other sounds of the software. Only one student at any given time had control of the mouse, but at regular intervals the students were encouraged to take turns using the mouse.

In a typical session, the students first watched an animated video that reviewed the material on which they were receiving instruction during their regular math class sessions in that same week. The students then played math games with animated characters to practice the skills introduced in the video lecture. If a student had a math-related question, he/she was encouraged to discuss the question with his/her teammate with whom he/she shared the computer. The students were not allowed to discuss with other teams or the teacher-supervisor. Generally, the games involved an animated character engaged in some task (archery, crossing a river, etc). Multiple-choice questions would then appear on the screen one at a time. Successful answers would aid the animated character in their task and incorrect answers would trip them up and/or slow them down. Either way, humorous animations would appear once the students chose their response. Both students in each pair had access to scratch paper at their station to take notes and make calculations. At the end of each game, students were shown how many of the questions had been answered correctly.

Our protocol required that the teachers could only help students with scheduling, computer hardware issues and software operations. This was done to try to control for the possible effect of the CAL supervisor's involvement and thereby to make sure any observed impact would be entirely due to the CAL program itself. In fact, according to our observations (during occasional unannounced visits to randomly selected schools), the software demanded the full attention of students. There was little, if any, interaction between the students and the teacher-supervisors. In addition, while there was a lot of interaction within each of the two-person teams, there was little communication between pair groups.

The intervention team spent considerable time in preparing the necessary hardware, software, curriculum and program implementation protocol in a way that would both facilitate smooth implementation of the CAL program and avoid confounding influences that might bias our results. As the first step, to meet the hardware requirements of the CAL program, we acquired (by way of donation from Dell, Inc.) 640 brand new identical desktop computers and installed the CAL software package on these desktops. We then removed all pre-installed software that would not be used during the CAL intervention (such as Windows built-in games and Microsoft Office) and disabled the Internet and USB functions on all of the computers. In this way, we could not only prevent students and teachers from using the program computers for other purposes that might affect the operation of the regular CAL program, but also avoid the interruptions that might otherwise be caused by accidental deletion of the CAL software or the introduction of viruses. Sealing the computers also ensured the quality of our evaluation of the program effects without capturing any other confounding influences (spillovers) if students had access to knowledge from other sources such as the Internet. It also prevented teachers/students in control groups from copying our CAL software onto other computers.

All teacher-supervisors of the 36 treatment schools also participated in a two-day mandatory training program. The training was designed to prepare the teacher-supervisors for their responsibilities in the CAL classes. The teacher-supervisors' five main responsibilities included: a.) taking attendance; b.) making sure that the CAL curriculum in each session was matched to the curriculum being taught in the students' math class; c.) managing the CAL classrooms so that order was maintained; d.) providing immediate assistance when students experienced difficulty in computer and/or math game software operations (but they were not to instruct the students in math); and e.) taking care of the CAL desktops and keeping close contact with our research group/volunteers regarding technical support or CAL management questions. Because this work was clearly beyond the scope of their normal classroom duties, we compensated the teacher-supervisors with a monthly stipend of 100 yuan (approximately 15 USD). This is an amount roughly equivalent to 15 percent of the wage of a typical rural teacher.

CAL Control Group (the boarding students in the 36 control schools)

The third and fifth grade boarding students in the 36 control schools constituted the CAL control group. Students in the control group did not receive any CAL intervention. To avoid spillover effects from the CAL intervention, the principals, teachers and students (and their parents) of the control schools were not informed of the CAL project. The research team did not visit the control schools except during the baseline and final evaluation surveys. Informed assent was given by the guardian of the students so that the children could participate in the baseline and endline surveys. However, neither students nor teachers in the control group knew that there were students in other schools participating in the CAL program. The students in the control group took their regular math classes at school as usual.

Data Collection

The research group conducted three rounds of surveys in the 72 control and treatment schools. The first-round survey was a baseline survey conducted with all third and fifth grade boarding students in the 72 schools in late February 2011 at the beginning of the spring semester and before any implementation of CAL program had begun. The second-round survey was an evaluation survey conducted in June 2011, a time that coincided with the end of the spring semester of 2011 (and the end of Phase I). The third-round survey was a final evaluation conducted at the end of Phase II in June 2012.

In each round of the survey, the enumeration/survey team (members of which were undergraduate and master's students recruited from a local university) visited all schools (treatment and control schools) and conducted a two-block survey. In the first block of each round of the survey, students were given a math test that gave us our main outcome variable. The test was the same for all students in the same grade across groups and schools in each round of survey. Students were required to finish the test in 25 minutes. Although drawn out of the same pool of questions, the math questions were different for each round of the survey. We also made sure that the math questions in the test did not repeat the exercises included in the computer assisted learning software. Our enumeration team closely monitored the test and strictly enforced the time limits. The math test scores were normalized to create the main outcome variable.[4]

In the second block, enumerators collected data on the characteristics of students and their families. From this part of the survey we created a set of demographic and socioeconomic variables. The dataset includes measures of each student's *gender*, *age* (measured in years), *only child* (if the student is the only child of his or her family), grade repetition (if the student has *ever repeated a grade*), parents' education level (*at least one parent has junior high school or higher degree* and *at least one parent has senior high school or higher degree*), parents' job (*at least one parent has an off-farm job*), *family wealth* (the variable of family wealth equals 1 if the family assets are higher than the median value and 0 otherwise) and computer use (the variable equals 1 if the student had *ever used a computer*).

Statistical Methods

We used both unadjusted and adjusted ordinary least squares (OLS) regression analyses to estimate how the academic performance changed in the treatment group relative to the control group. Our unadjusted analysis regressed changes in the

[4] In order to make test scores from different rounds of survey comparable, scores are normalized relative to the distribution of the baseline test scores of the control group. Specifically, we subtracted the mean of the control group in the baseline and divided by the standard deviation of the control group in the baseline.

outcome variable (i.e., post-program math test score minus baseline math test score) on a dummy variable of the treatment (CAL intervention) status. We used adjusted analyses as well to improve statistical efficiency (we will describe these approaches in detail in the models below). In all regressions, we corrected for the clustered standard errors at the school level.[5]

The unadjusted model is:

$$\Delta y_{is} = \alpha + \beta \cdot treatment_s + \varepsilon_{is} \quad (1)$$

where Δy_{is} is the change in the outcome variable during the program period for child *i* in school *s*, $treatment_s$ is a dummy variable for treatment school students (equal to one for students in the treatment group and zero otherwise) and ε_{is} is a random disturbance term clustered at the school level. By construction, the coefficient of the dummy variable $treatment_s$, β, is equal to the unconditional difference in the change in the outcome (Δy_{is}) between the treatment and control groups over the program period. In other words, β measures how the treatment group changed in the outcome levels during the program period relative to the control group.

In order to improve the efficiency of the estimation, we built on the unadjusted model in equation (1) by including a set of control variables:

$$\Delta y_{is} = \alpha + \beta \cdot treatment_s + \theta \cdot y_{0is} + X_{is}\gamma + \varepsilon_{is} \quad (2)$$

where all the variables and parameters are the same as those in equation (1), except that we added a set of control variables. Specifically, we control for y_{0is}, the baseline math test scores for student *i* in school *s*, and X_{is}, a vector of additional control variables. The variables in X_{is} are student *and* family characteristics (*gender, age,*

[5] The study randomizes treatment at the school level. Therefore, it is possible that the error terms are correlated within schools. To account for intra-cluster correlation, we used standard errors clustered at the school level in all the regressions testing the treatment effect. See Imbens and Wooldridge (2009) for more details.

only child, ever repeated grade, at least one parent has junior high school or higher degrees, at least one parent has senior high school or higher degrees, at least one parent has an off-farm job, family wealth, ever used a computer). By including y_{0is} and X_{is} as control variables, β in equation (2) provides an unbiased, efficient estimate of the CAL treatment effect.

We estimated the treatment effect of the CAL intervention across three time horizons. In order to show the longer-term effects of the CAL program, we first estimated the longer-term treatment effect of three semesters (CAL Phase I and Phase II). We then estimated the treatment effect for Phase II only and compared it with the effect of Phase I. In doing so, we investigate how the CAL program effect evolves from the one-semester program (Phase I) to the two-semester program (Phase II). Both equations (1) and (2) were used in estimating treatment effects across the three time horizons.

Results

The data show that students in the treatment group improved significantly more in their math performance than did students in the control group after taking the CAL classes for three semesters, from the beginning of the spring semester 2011 to the end of the spring semester 2012. The students in the treatment group and the control group in the third grade started at similar levels in pre-test standardized math scores at the start of spring semester of 2011 (Figure 3, Panel A). After three semesters of treatment (Phase I and Phase II), the treatment group improved more in math than did the control group (Panel A). The difference in the change in standardized math test scores between the two groups was 0.21 standard deviations for the third grade students (Panel B).

The results are similar for the fifth grade students (Figure 4). The data show that students in the fifth grade also improved significantly more in terms of their standardized math test scores than the students in the control group after three semesters of taking the CAL classes. In a statistical sense, the fifth grade baseline standardized test scores of control students are the same as the scores of the treatment students (Figure 4, Panel A). After the CAL intervention of three semesters, the students in the treatment group improved by 0.29 standard deviations more than did the students in the control group (Panel B).

The multivariate regression analyses (adjusted and unadjusted) are consistent with our graphical descriptive analysis (Table 4). Using only the third grade students or only the fifth grade students, the estimated CAL treatment effects on math test scores using results from the unadjusted model are 0.21 standard deviations for the third grade students (row 1, column 1) and 0.30 standard deviations for the fifth grade students (row 1, column 2). The estimated treatment effects for both grades are statistically significant (significant at the 10 percent level in the case of the third grade cohort; and significant at the 1 percent level in the case of the fifth grade cohort).

When we add the additional control variables, using the adjusted model, the results from the more efficient estimator demonstrate that the treatment effect is still large and statistically significant (Table 4). In the case of the third grade students, the estimated treatment effect is 0.25 standard deviations (row 1, column 3). In the case of the fifth grade students, the estimated treatment effect is 0.26 standard deviations. Both of the estimates are significant at the 1 percent level (row 1, column 3 and 4). An increase of one-fourth of a standard deviation can amount to a considerable gain in performance. Such an increase in performance is estimated by some to be equivalent to 0.6 years of schooling (Glewwe, Park & Zhao, 2011). A similar effect size was found in other prominent education experiments. For example, the Tennessee Star Program sought to measure the effect of reducing class size by one third (from a classroom of 22 to 25 students to a classroom of 13 to 17 students—Mosteller, 1995). The program was considered successful in that test scores were raised by 0.25 standard deviations. Our measured effect size is similar in magnitude.

Comparing the Effects of the CAL Intervention between Phase I and Phase II

Using the results from the regression model based on the specification in equation (2), the results show that the effect of the CAL treatment appears to persist in the longer run (Table 5). Our point estimate shows that during CAL Phase I (March 2011 to June 2011), the estimated treatment effect for the third grade boarding students is equal to 0.18 standard deviations and is significant at the 5 percent level (row 1, column 1). During CAL Phase II (September 2011 to June 2012), the point estimate of the treatment effect for the third grade students is still positive. The magnitude is 0.07 standard deviations, though statistically insignificant (row 1, column 3). Hence, this result (0.25 standard deviation shift over the two study phases as shown in Table 4) suggests that the impact persisted in the longer term.

Consistent with the third grade, the fifth grade boarding students also improved in the longer-term CAL program. The estimated treatment effect of Phase I is equal to 0.11 standard deviations, though this is statistically insignificant (Table 5, row 1, column 2). However, during Phase II the estimated treatment effect becomes significant at the 10 percent level and the point estimate is equal to 0.15 standard deviations (row 1, column 4). Like the third graders, fifth graders improved in both phases (in the short term and the longer term) to achieve an overall learning improvement of 0.26 standard deviations after three semesters of CAL classes.

The high interest level in the CAL software among the treatment students supports these results. The ratings of student interest in the software (0-100 points) at the end of CAL Phase II suggest that the students were highly interested in the software regardless of their previous computer experience or academic performance (Appendix 1). The mean rating of student interest was 88 points for the third grade students and 83 points for the fifth grade students. Third grade students that had used computers before the CAL program had an interest rating as high as 88 points (row 1, column 2). The third grade students without any computer experience before CAL had a rating of 89 points (row 1, column 3). However, the difference is not significant (row 1, columns 4-5). The difference between the fifth grade students with and without previous computer experience is also small and not significant (row 2, columns 2-5). Moreover, both the better and worse performing third grade and fifth grade students showed high interest in the software (none of the differences were significant; rows 1 and 2, columns 6-9). These findings suggest that all students, regardless of time exposed to computer technology and level of academic performance, maintained a high level of interest in CAL.

Heterogeneous effects of the CAL intervention

In order to test for heterogeneous program effects, we also included in the regression model specified in equation (2) interaction terms between the treatment dummy variable and two key covariates. For example, we tested whether the change in math test scores differed for students who were better performing in math at the time of the baseline relative to students who were poorer performing. This was done by including in the regression an interaction term between the treatment dummy variable and the variable of *baseline math test score*. We also tested if boys benefited differently from the program than girls by including the interaction term between the treatment dummy variable and the variable of *gender*.

The estimated results using Equation (2)— which includes the interaction term between the treatment dummy variable and the baseline math test score—demonstrate that the CAL intervention worked similarly well for the better performing and poorer performing students in the third and fifth grades (Table 6). We find no significant evidence of heterogeneous program effects of CAL on standardized math test scores (row 2). Students in the third and fifth grade who scored relatively high and relatively low on the baseline math test did equally well after the entire CAL treatment of three semesters (Phase I and Phase II, row 2, columns 1 and 4). When estimating with Phase I only and Phase II only, we do not find that the poorer and better performing students had significant differences in their math improvements (row 2, columns 2, 3, 5 and 6).

However, there does seem to be an interesting heterogeneous effect among girls and boys in the third grade (Table 7). The third grade girls improved more than the boys by 0.21 standard deviations in math in Phase I (the difference is significant at the 5 percent level, row 2, column 2). More specifically, after the CAL intervention in Phase I girls in the treatment group improved by 0.29 standard deviations in math relative to the girls in the control group (row 1, column 2), while the boys in the treatment group improved by 0.08 (0.29-0.21) standard deviations in math relative to the boys in the control group (rows 1 and 2, column 2). The difference in math improvement between girls and boys is ultimately reduced as boys seem to catch up by improving more than the girls in Phase II (the coefficient on the interaction term between the treatment and being a boy is positive for Phase II, row 2, column 3). As a result, the difference in the treatment effect between the girls and boys during the entire treatment of three semesters is 0.09 standard deviations and insignificant (row 2, column 1). In contrast, boys and girls were affected by the treatments similarly in the fifth grade (none of the differences across the three time horizons are significant and all the scales are small, row 2, columns 4-6).

Conclusion

In this paper we present the results from a randomized field experiment of a Computer Assisted Learning (CAL) program in 72 rural public schools in Ankang Prefecture, Shaanxi Province. The study involves 2741 third grade and fifth grade boarding students. To evaluate the effectiveness of the program we randomly chose 36 schools from the entire sample as treatment schools and had third and fifth grade students undergo the CAL intervention. Phase I of the program was held for an intervention period of one semester (half a year in the spring semester of 2011). Phase II of the program was implemented for an intervention period of one academic year (the 2011-2012 school year). The remaining 36 schools served as control schools. This paper contributes to the limited understanding of whether a CAL program has a persistent (longer-run) impact on student learning.

Our results indicate that the CAL program that was implemented for one year and a half had significant beneficial effects on student academic outcomes. Two 40-minute CAL math sessions per week for one and a half years increased student standardized math scores by 0.25 standard deviations for the third grade boarding students and 0.26 standard deviations for the fifth grade boarding students. In addition, we have shown that the program effect on student gains in math learning is persistent over one and a half years. In other words, the students continued to improve in math when using the CAL program even after they had become accustomed to using both computers and the software.

By testing for heterogeneous effects, we found that the CAL intervention worked similarly well for the poorer performing and better performing students in the third and fifth grades during the entire treatment of three semester (Phase I and Phase II, combined), during Phase I only (short term) and during Phase II only (longer term). We also found that the third grade girls improved more than the boys in math in the short term (CAL Phase I).

Interesting questions remain about whether the CAL program can be made more efficient in improving student learning in developing countries. For example, future research can be conducted to explore whether the interaction between the two students who share one computer during the CAL sessions is beneficial or harmful to student learning. Switching control of the computer may reduce the learning time of a single student (Rogers & Lindley, 2004). However, the interaction and discussion between students may improve the efficiency of learning (Stahl, Koschmann & Suthers, 2006). Future studies should be conducted to explore whether interaction or what kind of interaction between students during the CAL classes can help the students gain more in academic performance in rural China. Moreover, studies should also be conducted to explore the impacts of CAL on other key subjects in the national primary school curriculum, such as Chinese and English. Chinese language skills are particularly important because they have been found to affect off-farm work opportunities and wages (Li, Sato & Sicular, 2011). English test scores have also been found to be one of the indicators that best predicts students' chance of college admission and level of post-college income (Li, Meng, Shi & Wu, 2012). If CAL can be made to effectively improve Chinese and English language skills in addition to math skills among rural students, there will be important policy implications. This will be especially significant if the central government in China fulfills its stated goal of placing computer rooms in every rural school.

References

Banerjee, A., Cole, S., Duflo, E., and Linden, L. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics*. 122 (3): 1235-1264.

Barrow, L., Markman, L., and Rouse, C. 2008. "Technology's Edge: The Educational Benefits of Computer-Aided Instruction." *National Bureau of Economic Research working paper No. 14240*.

Blok, H., Oostdam, r., Otter, M. E., and Overmaat, M. 2002. "Computer-Assisted Instruction in Support of Beginning Reading Instruction: A Review." *Review of Educational Research*, 72(1): 101-130.

Chang, C. Y. (2002). Does computer-assisted instruction + problem solving = improved science outcomes? A pioneer study. *Journal of Education Research, 95(*3), 143-151.

CNBS [China National Bureau of Statistics]. 2011. *China National Statistical Yearbook, 2011.* China State Statistical Press: Beijing, China.

Cordova, D. I., and Lepper, M. R. 1993. "Intrinsic motivation and the process of learning: beneficial effects of contextualization, personalization, and choice." *Journal of Educational Psychology*, 88 (4): 715-730.

Cotton, K. 2001. "Computer-assisted Instruction." *School Improvement Research Series,* 10.

Garcia, M. R., and Arias, F. V. 2000. "A Comparative Study in Motivation and Learning through Print-oriented and Computer-oriented Tests." *Computer Assisted Language Learning*, 13(4-5): 457-465.

Garris, R., Ahlers, R., and Driskell, J. E. 2002. "Games, motivation, and learning: a research and practice model." *Simulation & Gaming*, 33(4): 441-467.

Glewwe, P., and Kremer, M. 2006. "Schools, Teachers, and Education Outcomes in Developing Countries." *Handbook on the Economics of Education.* (New York, NY: Elsevier), 2: 945-1017.

Günther, V. K., Schäfer, P., Holzner, B. J., and Kemmler G. W. 2010. "Long term improvements in cognitive performance through computer-assisted cognitive training: A pilot study in a residential home for older people." *Aging & Mental Health*, 7(3): 200-206.

He, F., Linden, L., and MacLeod, M. 2008. "How to Teach English in India: Testing the Relative Productivity of Instruction Methods within the Pratham English Language Education Program." Working Paper. Columbia University Department of Economics.

Inal, Y., and Cagiltay, K. 2007. "Flow experiences of children in an interactive social game environment." British Journal of Educational Technology, 83(3): 455–464.

Kirkpatrick, H., and Cuban, L. 1998. "Computers Make Kids Smarter – Right?" *Technos Quarterly for Education and Technology*, 7 (2): 26-31.

Lai, F., Luo, R., Zhang, L., Huang, X. and Rozelle, S. 2011. "Does Computer-Assisted Learning Improve Learning Outcomes? Evidence from a Randomized Experiment in Migrant Schools in Beijing." REAP working paper.

Lai, F., Zhang, L., Qu, Q., Hu, X., Shi, Y., Boswell, M., and Rozelle S. 2012a. "Computer Assisted Learning as Extracurricular Tutor? Evidence from a Randomized Experiment in Rural Boarding Schools in Shaanxi." REAP working paper.

Lai, F., Zhang, L., Qu, Q., Hu, X., Shi, Y., Boswell, M., and Rozelle S. 2012b. "Does Computer-Assisted Learning Improve Learning Outcomes? Evidence from a Randomized Experiment in Public Schools in Rural Minority Areas in Qinghai, China." REAP working paper.

Lepper, M. R., and Malone, T. W. 1987. "Intrinsic motivation and instructional effectiveness in computer-based education." In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction* (3): 255-286. Hillsdale, NJ: Lawrence Erlbaum.

Li, H., Meng, L., Shi, X., and Wu, B. 2012. "Does Attending Elite Colleges Pay in China?" *Journal of Comparative Economics*, 40: 78-88.

Li, S., Sato, H. and Sicular T., eds. 2011. "*Rising Inequality in China: Challenge to a Harmonious Society*." Cambridge University Press, UK.

Linden, L. 2008. “Computer or Substitute? The Effect of Technology on Student Achievement in India.” *Working Paper*. Columbia University Department of Economics.

Malone, T., and Lepper, M. 1987. “Making Learning Fun: A Taxonomy of Intrinsic Motivations of Learning.” In R. E. Snow and M. J. Farr (Ed.), *Aptitude, learning, and instruction, vol. 3. Conative and affective process and analyses*, Hillsdale, NJ: 223-253.

Marjanovic, O. 1999. “Learning and Teaching in a Synchronous Collaborative Environment.” *Journal of Computer Assisted Learning*, 15: 129-138.

McEwan, P. J. 2013. “Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments”. Working paper.

Ministry of Education. 2012. *12th Five-Year Plan for Integrating Information Technology into Education.*

PCAST [President’s Committee of Advisors on Science and Technology]. 1997. “Panel on Educational Technology.” *Report to the President on the Use of Technology to Strengthen K-12 Education in the United States*, Washington DC: Office of the President.

Roesch, A., Gruber, H., Hawelka, B., Hamm, H., Arnold, N., Popal, H., and Stolz, W.2003. Computer assisted learning in medicine: a long term evaluation of the 'Practical Training Programme Dermatology 2000'. Medical Informatics and the Internet in Medicine, 28(3):147-159.

Rogers, Y., and Lindley, S. 2004. "Collaborating around Large Interactive Displays: Which Way is Best to Meet?" *Interacting with Computers*, 16(6): 1133-1152.

Schaefer, S., and Warren, J. 2004. "Teaching computer game design and construction." Computer-Aided Design, 36: 1501–1510.

Stahl, G., Koschmann, t., and Suthers, D. 2006. "Computer-supported Collaborative Learning: A Historical Perspective." In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences*, Cambridge, UK: 409-426.

Yuan, X. 2012. "Strategy for Improving Public Service: Push Forward the Integration of ICT in Rural Areas in China (in Chinese)." Retrieved from: http://www.stdaily.com/stdaily/content/2012-02/20/content_429399.htm

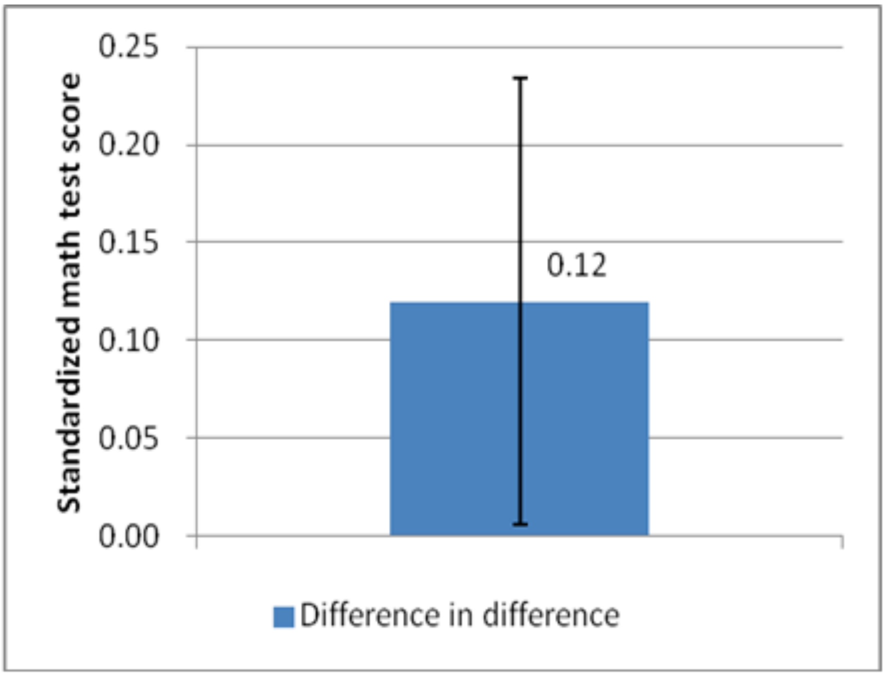


Figure 1. Difference in difference in the standardized math test scores before and after the CAL Phase I Program (March 2011 and June 2011) between the treatment and control groups in both the third and the fifth grades

Cited from Lai et al. (2012a).

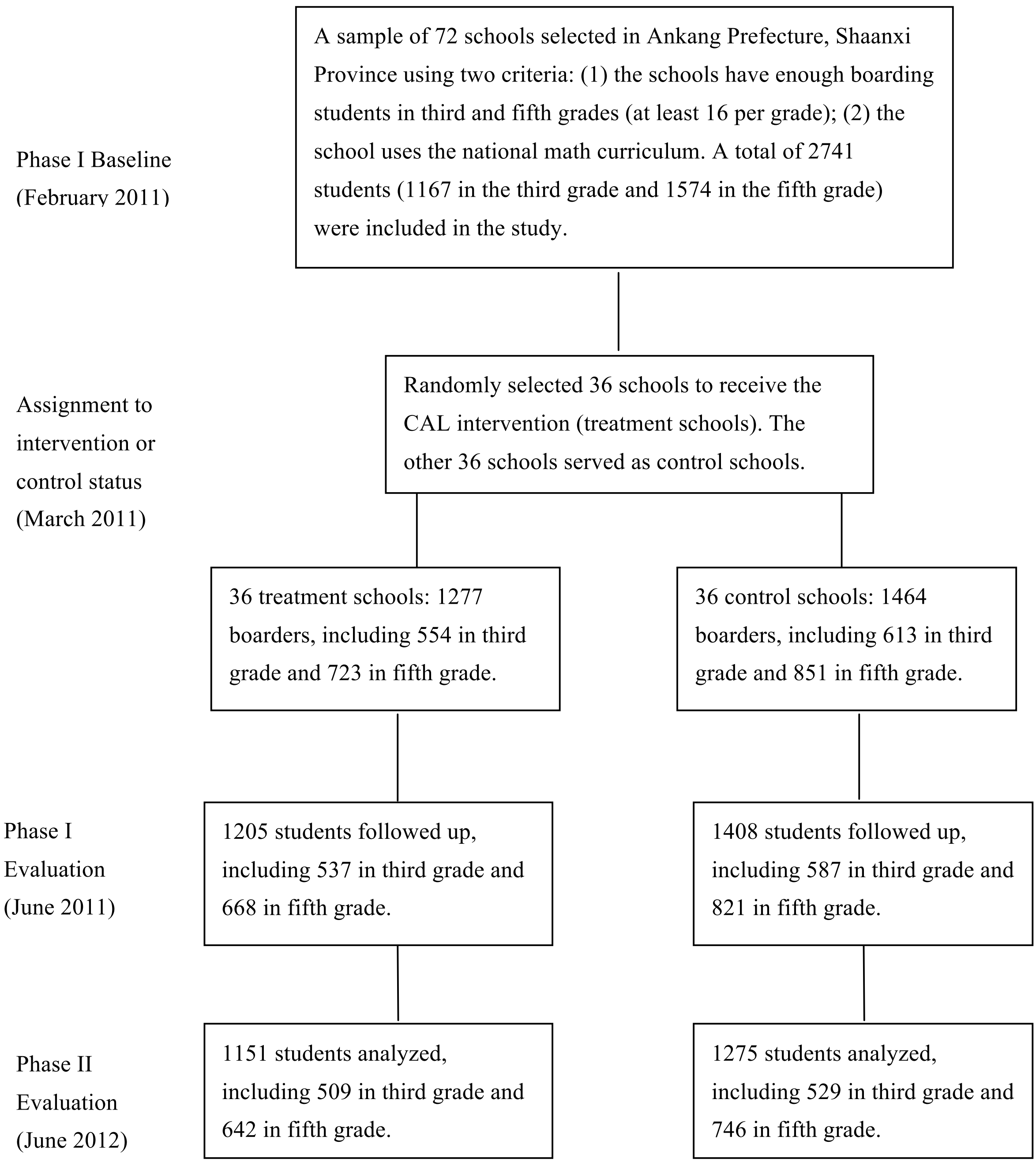
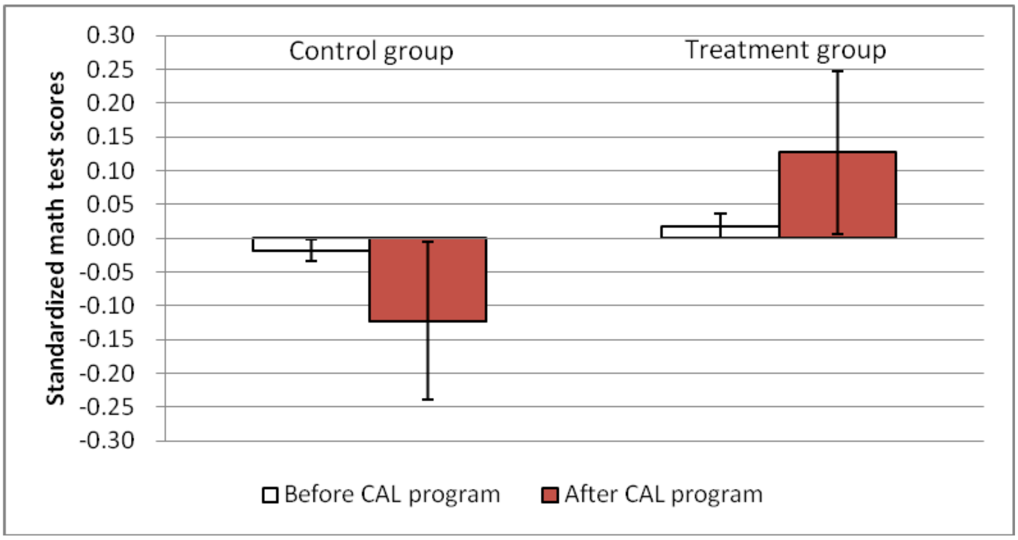
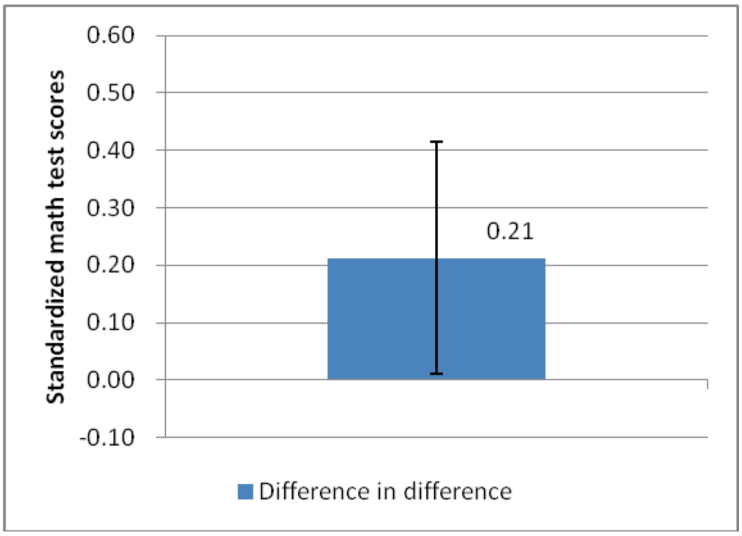


Figure 2: Experiment Profile

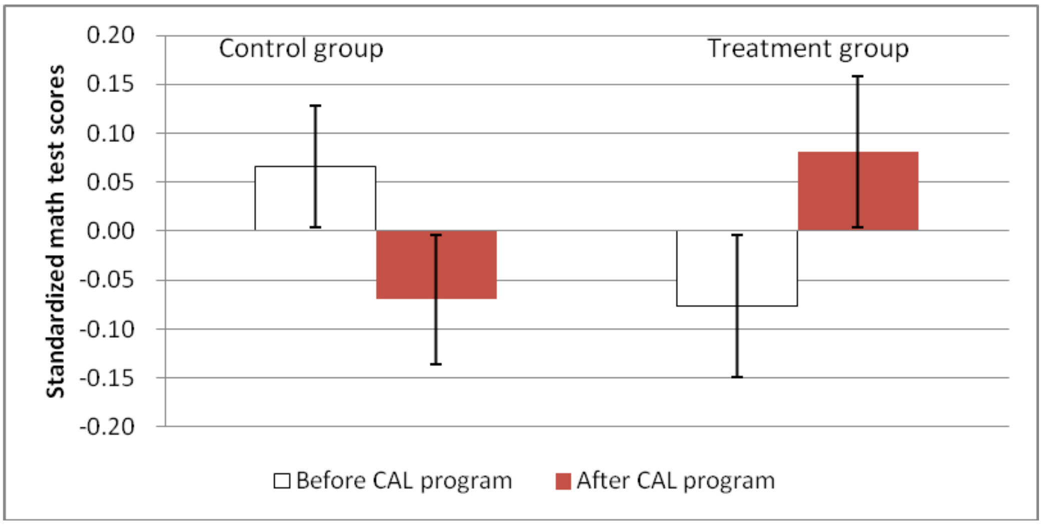


Panel A. Standardized math test scores before and after the entire CAL intervention (Phase I and Phase II): treatment and control groups in third grade.

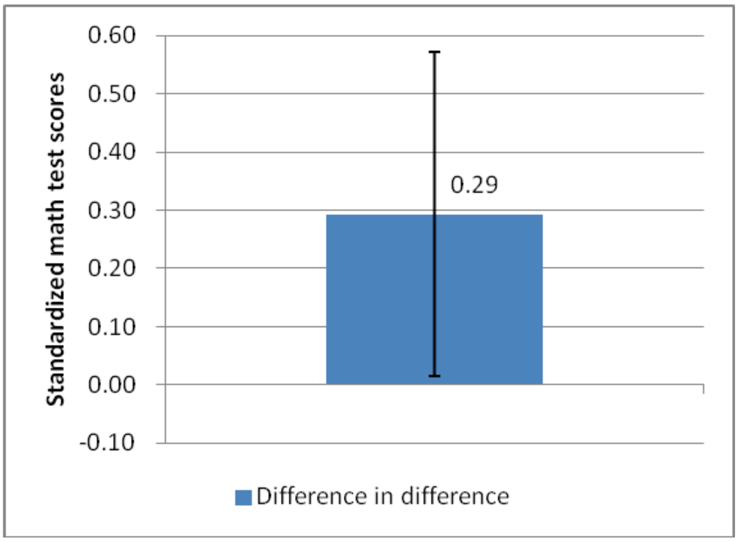


Panel B. Difference in difference in the standardized math test scores before and after the entire CAL intervention (Phase I and Phase II) between treatment and control groups in third grade

Figure 3. Change in the standardized math test scores of third grade students before and after the entire CAL intervention (Phase I and Phase II)



Panel A. Standardized math test scores before and after the entire CAL intervention (Phase I and Phase II): treatment and control groups in fifth grade.



Panel B. Difference in difference in the standardized math test scores before and after the entire CAL intervention (Phase I and Phase II) between treatment and control groups in fifth grade

Figure 4. Change in the standardized math test scores of fifth grade students before and after the entire CAL intervention (Phase I and Phase II)

Table 1. Ordinary Least Square analysis of CAL Phase I (March through June, 2011) on student standardized math test scores for third and fifth grade students.

		All (Third grade and Fifth grade)	Third grade	Fifth grade
		(1)	(2)	(3)
[1]	Treatment (1=treatment group; 0=control group)	0.12**	0.18**	0.07
		[0.05]	[0.08]	[0.07]
[3]	Control variables	Yes	Yes	Yes
[4]	Observations	2613	1124	1489
[5]	R-squared	0.26	0.29	0.25

Cited from Lai et al. (2012a).

* significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors in brackets clustered at school level.

Table 2. Comparison of student characteristics between attrited and non-attrited students in the sample and between treatment and control students during the entire CAL intervention (Phase I and Phase II).

		Difference between attrited and non-attrited students[c]		Difference between the treatment and control groups within attrited students
		Third grade	Fifth grade	All attrited students[d]
		(1)	(2)	(3)
[1]	Baseline math score (units of standard deviation)[a]	-0.01	-0.03	0.02
		(0.01)	(0.02)	(0.04)
[2]	Baseline Chinese score (units of standard deviation)[b]	-0.02**	-0.01	-0.05
		(0.01)	(0.01)	(0.04)
[3]	Gender (1=boy; 0=girl)	-0.01	0.03*	0.03
		(0.02)	(0.02)	(0.05)
[4]	Age(years)	0.01	-0.00	0.01
		(0.01)	(0.01)	(0.03)
[5]	Only child (1=yes; 0=no)	0.01	0.03**	0.02
		(0.02)	(0.02)	(0.05)
[6]	Ever repeated grade (1=yes; 0=no)	0.02	0.02	-0.07
		(0.02)	(0.02)	(0.06)
[7]	At least one parent has junior high school or higher degrees (1=yes; 0=no)	0.02	0.01	0.06
		(0.02)	(0.02)	(0.06)
[8]	At least one parent has senior high school or higher degrees (1=yes; 0=no)	0.03	-0.01	-0.02
		(0.03)	(0.03)	(0.09)
[9]	At least one parent has an off-farm job (1=yes; 0=no)	0.02	-0.04*	0.03
		(0.02)	(0.02)	(0.08)
[10]	Family wealth (1=higher than the median; 0=otherwise)	0.02	0.02	-0.08
		(0.02)	(0.02)	(0.06)
[11]	Ever used a computer (1=yes; 0=no)	-0.05	-0.05	0.15
		(0.04)	(0.08)	(0.12)
[12]	Observations	1,167	1,574	305

* significant at 10%; ** significant at 5%; * significant at 1%. Robust standard errors in brackets clustered at school level.

[ab] The baseline math score is the score on the standardized math test that is given to all sample students before the CAL Program.

[c] The sample includes both the sample observations (non-attrition) and the attrition observations.

[d] The sample is limited to the attrited observations.

Table 3. Difference in characteristics between students in the treatment group and the control group during the entire CAL intervention (Phase I and Phase II).

Dependent variable: whether the student received CAL treatment (1=yes; 0=no)		Third grade [c] (1)	Fifth grade [c] (2)
[1]	Baseline math score (units of standard deviation)[a]	0.02	-0.04
		(0.03)	(0.04)
[2]	Baseline Chinese score (units of standard deviation)[b]	-0.04	0.01
		(0.03)	(0.03)
[3]	Gender (1=boy; 0=girl)	-0.02	-0.03
		(0.03)	(0.03)
[4]	Age(years)	0.00	0.04
		(0.03)	(0.02)
[5]	Only child (1=yes; 0=no)	0.02	-0.05
		(0.05)	(0.03)
[6]	Ever repeated grade (1=yes; 0=no)	0.00	-0.00
		(0.04)	(0.04)
[7]	At least one parent has junior high school or higher degrees (1=yes; 0=no)	0.01	-0.03
		(0.03)	(0.03)
[8]	At least one parent has senior high school or higher degrees (1=yes; 0=no)	0.06	0.04
		(0.05)	(0.05)
[9]	At least one parent has an off-farm job (1=yes; 0=no)	-0.05	-0.02
		(0.05)	(0.04)
[10]	Family wealth (1=higher than the median; 0=otherwise)	-0.01	-0.00
		(0.04)	(0.04)
[11]	Ever used a computer (1=yes; 0=no)	0.05	0.08
		(0.11)	(0.12)
[12]	Observations	1,038	1,388
[13]	R-squared	0.111	0.120

* significant at 10%; ** significant at 5%; * significant at 1%. Robust standard errors in brackets clustered at school level.

[ab] The baseline math score is the score on the standardized math test that is given to all sample students before the CAL Program.

[c] The sample includes the remaining sample (non-attrition).

Table 4. Ordinary Least Squares estimators of the impacts of the entire CAL intervention (Phase I and Phase II) on student math test scores.

Dependent variable: standardized post-CAL math test score - standardized baseline math test score (standard deviations)					
		Third grade	Fifth grade	Third grade	Fifth grade
		(1)	(2)	(3)	(4)
[1]	Treatment (1=treatment group; 0=control group)	0.21*	0.30***	0.25***	0.26***
		(0.11)	(0.10)	(0.08)	(0.08)
[2]	Baseline math score (units of standard deviation)[a]			-0.64***	-0.59***
				(0.04)	(0.03)
[3]	Baseline Chinese score (units of standard deviation)[b]			0.18***	0.15***
				(0.04)	(0.03)
[4]	Gender (1=boy; 0=girl)			0.13**	0.02
				(0.05)	(0.04)
[5]	Age(years)			-0.11**	-0.10***
				(0.04)	(0.03)
[6]	Only child (1=yes; 0=no)			0.06	0.02
				(0.08)	(0.05)
[7]	Ever repeated grade (1=yes; 0=no)			-0.03	0.04
				(0.06)	(0.05)
[8]	At least one parent has junior high school or higher degrees (1=yes; 0=no)			0.00	0.04
				(0.07)	(0.04)
[9]	At least one parent has senior high school or higher degrees (1=yes; 0=no)			-0.07	-0.05
				(0.08)	(0.08)
[10]	At least one parent has an off-farm job (1=yes; 0=no)			-0.06	0.04
				(0.07)	(0.07)
[11]	Family wealth (1=higher than the median; 0=otherwise)			-0.07	0.10**
				(0.05)	(0.04)
[12]	Ever used a computer (1=yes; 0=no)			0.07	0.25*
				(0.10)	(0.14)
[13]	Observations	1,038	1,388	1,038	1,388
[14]	R-squared	0.011	0.024	0.322	0.293

* significant at 10%; ** significant at 5%; * significant at 1%. Robust standard errors in brackets clustered at school level.

[ab] The baseline math score is the score on the standardized math test that is given to all sample students before the CAL Program.

Table 5. Ordinary Least Squares estimators of the impacts of the CAL Program during Phase I and Phase II on student math test scores for third and fifth graders.

Dependent variable: standardized post-CAL math test score - standardized baseline math test score

		CAL Phase I		CAL Phase II	
		Third grade	Fifth grade	Third grade	Fifth grade
		(1)	(2)	(3)	(4)
[1]	Treatment (1=treatment group;	0.18**[b]	0.11 [c]	0.07 [b]	0.15* [c]
	0=control group)	(0.08)	(0.07)	(0.10)	(0.08)
[2]	Control variables [a]	Yes	Yes	Yes	Yes
[3]	Observations	1,038	1,388	1,038	1,388
[4]	R-squared	0.301	0.261	0.048	0.038

* significant at 10%; ** significant at 5%; * significant at 1%. Robust standard errors in brackets clustered at school level.

[a] Control variables include all the variables in Table 2 and the township dummies.

[b] Wald test shows that the treatment effect on the third grade students is not significantly different between Phase I and Phase II (p-value=0.48).

[c] Wald test shows that the treatment effect on the fifth grade students is not significantly different between Phase I and Phase II (p-value=0.73).

Table 6. Ordinary Least Squares estimators of the heterogeneous effects of CAL intervention on standardized math test scores by baseline math performance.

Dependent variable: standardized post-CAL math test score-standardized baseline math test score							
		Third grade			Fifth grade		
		Phase I and Phase II	Phase I	Phase II	Phase I and Phase II	Phase I	Phase II
		(1)	(2)	(3)	(4)	(5)	(6)
[1]	Treatment (1=treatment group; 0=control group)	0.25***	0.19**	0.07	0.26***	0.10	0.16**
		(0.08)	(0.08)	(0.09)	(0.08)	(0.07)	(0.08)
[2]	Treatment * Standardized baseline math test score	0.01	0.04	-0.03	-0.03	-0.07	0.05
		(0.06)	(0.06)	(0.09)	(0.06)	(0.06)	(0.07)
[3]	Control variables [a]	Yes	Yes	Yes	Yes	Yes	Yes
[4]	Observations	1,038	1,038	1,038	1,388	1,388	1,388
[5]	R-squared	0.322	0.301	0.048	0.293	0.262	0.038

* significant at 10%; ** significant at 5%; * significant at 1%. Robust standard errors in brackets clustered at school level.

[a] Control variables include all the variables in Table 2 and the township dummies.

Table 7. Ordinary Least Squares estimators of the heterogeneous effects of the CAL intervention on standardized math test scores by gender.

Dependent variable: standardized post-CAL math test score-standardized baseline math test score		Third grade			Fifth grade		
		Phase I and Phase II	Phase I	Phase II	Phase I and Phase II	Phase I	Phase II
		(1)	(2)	(3)	(4)	(5)	(6)
[1]	Treatment (1=treatment group; 0=control group)	0.30***	0.29***	0.01	0.29***	0.10	0.19**
		(0.09)	(0.09)	(0.11)	(0.09)	(0.08)	(0.08)
[2]	Treatment * Gender (1=boy; 0=girl)	-0.09	-0.21**	0.11	-0.05	0.02	-0.07
		(0.10)	(0.09)	(0.10)	(0.08)	(0.09)	(0.10)
[3]	Control variables [a]	Yes	Yes	Yes	Yes	Yes	Yes
[4]	Observations	1,038	1,038	1,038	1,388	1,388	1,388
[5]	R-squared	0.322	0.304	0.049	0.293	0.261	0.038

* significant at 10%; ** significant at 5%; * significant at 1%. Robust standard errors in brackets clustered at school level.

[a] Control variables include all the variables in Table 2 and the township dummies.

Appendix 1. Student interest rating in the CAL software by the end of Phase II.

			Interest rating by computer experience				Interest rating by baseline math score			
		Mean interest rating of the total sample	Students that had used computer before the program	Students that had never used computer before the program	Difference = (2)-(3)	P-value	Students with baseline math score higher than the median	Students with baseline math score lower than or equal to the median	Difference = (6)-(7)	P-value
		(1)	(2) Mean (Std Dev)	(3) Mean (Std Dev)	(4)	(5)	(6) Mean (Std Dev)	(7) Mean (Std Dev)	(8)	(9)
[1]	Third grade	88.28	88.09 (19.34)	89.33 (18.64)	-1.24	0.60	87.29 (21.63)	88.91 (17.53)	-1.63	0.35
[2]	Fifth grade	83.39	83.48 (18.28)	82.34 (18.82)	1.14	0.68	82.39 (20.43)	84.27 (16.20)	-1.88	0.20