Computer-Assisted Learning: Evidence from a Randomized Experiment

Leigh Linden, Abhijit Banerjee, and Esther Duflo

Poverty Action Lab Paper No. 5
October 2003

Poverty Action Lab
TRANSLATING RESEARCH INTO ACTION

Computer-Assisted Learning: Evidence from A Randomized Experiment

Leigh Linden, Abhijit Banerjee, and Esther Duflo

DRAFT
October 2, 2003

Abstract
This note presents the results obtained after the first year of a two-year randomized evaluation of a computer assisted learning (CAL) program in Vadodara, India. The CAL program, implemented by a NGO, took advantage of the donation of four computers to each municipal primary school in Vadodara by the state government. The program provided each child in the fourth standard with two hours of shared computer time in which students played educational games that reinforced mathematics competencies ranging from the standard 1 to the standard 3 level. We find the program to be quite effective. On average, it increased math scores by 0.37 standard deviations. The program effect is slightly higher at the bottom of the distribution but persists throughout the distribution. The program had no apparent spillover on language competencies.

Introduction

Many see Computer Assisted Learning (CAL), as a supplement to regular instruction, as a possible way to improve the quality of education in developing countries. Good educational software can be reproduced at nominal cost, and well-designed educational games can sustain interest and curiosity even in an otherwise dull school environment. The excitement seems to be particularly strong in India, where the high-tech sector is both successful and visible. Many local governments have started providing computers in schools but without offering much guidance about how the schools should use them. The idea of using computers is particularly attractive in urban public schools and in rural areas where the number of qualified teachers is limited and the quality of existing teachers is notoriously poor. Computers have the potential to both directly improve learning and indirectly increase attendance by making school more attractive.

Unfortunately, despite the general excitement, there exists very little rigorous evidence of the impact of computers on educational outcomes and no reliable evidence for India or other developing countries. Furthermore, what evidence that exists is not particularly encouraging. For example, Angrist and Lavy (2002) evaluate a computer assisted learning program in Israeli schools with disappointing results. Among the fourth and eighth grade students evaluated with math and Hebrew exams, the data show no benefits for computer assisted instruction and provide some evidence that children who received such instruction are actually at a disadvantage. It is not clear, however, that these results apply to the use of computers in schools in developing countries since in Israel, the computer-assisted learning replaces time spent in well equipped classrooms with high quality instructors. It is easy to imagine that computers can make a significant improvement in schools in developing countries even if they do not prove to be useful in the developed world.

This paper reports on the results after the first year of a randomized evaluation of a computer assisted learning program in Vadodara, India that attempts to improve school quality and test scores, implemented under the aegis of Pratham, a Bombay-based Non-Governmental Organization. This evaluation was conducted in conjunction with the evaluation of another of Pratham's programs, the balsakhi program (the results of which are presented in a separate paper). Both programs were implemented in the urban slums of Vadodara, where most children of the relevant age do attend school, however desultorily, so that the main impact is on the learning environment of existing students. The computer assisted learning intervention seeks to help children in the fourth standard improve their mathematics skills by playing specially designed educational games on the computer.

We find that in its first year, the program had a large and significant effect on test scores in math. Most of the effect was seen during the second half of the year (from the mid to the post test). The effect is comparable for boys and girls, and significant at all levels of the distribution, although it is somewhat larger for children at the bottom of the skill distribution. The program seems to have had no spillover on language competencies; we find no impact on language test scores. In a future iteration of this document, we will examine the impact on drop out and attendance rates.

2 The Computer Assisted Learning program

Pratham was established in Mumbai in 1994 with support from UNICEF and has since then expanded to several other cities in India. Pratham now reaches over 121,000 children in 20 cities in India and employs about 10,000 individuals. Pratham works closely with the government. Most of its programs are conducted in the municipal schools, and Pratham also provides technical assistance to the government.

Until recently, one of Pratham's core programs was a remedial education program called the balsakhi program. This program, in place in the municipal schools, provided a teacher (usually a young woman, recruited from the local community, who had herself finished secondary school) for children identified as falling behind their peers. While the exact details vary depending on local conditions, the typical instructor met with a group of approximately 15-20 children a day for two hours. Instruction focused on the core competencies the children should have learned in the second and third standards, primarily basic numeracy and literacy. The instructors were provided with a standardized curriculum developed by Pratham. They received two weeks of training at the beginning of the year and ongoing reinforcement while school was in session. The balsakhi program was in place at most of Pratham's sites. It was started in Mumbai in 1994, and then expanded to Vadodara in 1999.

The Computer Assisted Learning (CAL) program in Vadodara takes advantage of both a policy put in place by the government of Gujarat in 2000 as well as the established infrastructure of the balsakhi program. The government delivered four computers to each of 100 municipal primary schools in the city (80% of the public schools). A survey conducted by Pratham in June 2002 suggested that very few of these computers were actually used by children in elementary grade levels. While some schools may have run programs for older students or allowed teachers to use them for administrative tasks, most of the computers remained in their boxes, for want of anyone capable of operating them.

Pratham had previous experience with computer assisted learning, having run a small computer assisted learning program in Mumbai for several years. In particular, they had developed instructional software in the local language, Gujarati. After consultation with the Vadodara Municipal Corporation, they introduced a computer assisted learning program in half of the VMC schools, using the computers already present when possible and replacing or adding computers where necessary.

Pratham hired a team of instructors from the local community and provided them with five days of computer training. These instructors provided children with two hours of shared computer time per week (two children sharing one computer) — one hour during class time and one hour either immediately before or after school. During that time, the children played a variety of educational computer games chosen because they emphasized some of the basic competencies in the VMC mathematics curriculum.

Pratham designed the program to allow the children to learn as independently as possible. The instructors encouraged each child to play games that challenged the student's level of comprehension, and when necessary, they helped individual children understand the tasks required of them by the game. All interaction between the students and instructors was driven by the child's use of the various games, and at no time did any of the instructors provide general instruction in mathematics.

Schools where the CAL program was not implemented were free to continue to use the computer at their convenience, but our observation was that, except for a small number of schools, they did not start to make use of them for instructional purpose.

3 Evaluation Design

3.1 Sample

In 2000, when Pratham decided to expand the balsakhi program to cover the entire city of Vadodara, they decided to take advantage of the expansion to evaluate the effectiveness of the program in the remaining 98 eligible schools in the city. The program was continued in the school years 2001-2002 and 2002-2003. In 2001-2002, half the schools received a balsakhi in standard four and the other half in standard three. In 2002-2003, the 25 remaining primary schools were added to the sample. Those schools that had received a balsakhi in standard three in the previous year now received a balsakhi in standard four. The new schools were randomly assigned a balsakhi in either standard three or four.

In addition, the CAL program was started in approximately half of the municipal primary schools in Vadodara in 2002-2003, focusing exclusively on children in standard four. The sample was stratified according to gender, language of instruction ("medium", in the official terminology) of the school, the average math test scores in the post-test in the previous year as well as treatment or control status for the standard four balsakhi program. Table 1 summarizes the allocation of schools across different groups in the program. Unfortunately, in some schools, computers could not physically be installed,, either because of space constraints or the lack of electricity to run the computers. These schools are excluded from the comparison as well as the treatment group.[1] Thus, in the final sample for the study, 55 schools received the CAL program and 56 serve as the control group.

3.2 Outcomes

The main outcome of interest was whether the interventions resulted in any improvement in cognitive skills. In Vadodara, children were tested at the beginning of the school year (August), in November 2002, and in March 2003. Scores on the pre- and post-tests can be directly compared, as the format of the questions and the competencies tested remained the same. The exam comprised two parts: A math section and a language section. Both parts focused on competencies that the Vadodara Municipal Corporation (VMC) prescribes for children in standards one through four. On the math exam, for example, tasks ranged from basic number recognition, counting, and ordering of single digit numbers to ordering of two digit numbers, addition of single and two digit numbers, and basic word problems.

The first year of the balsakhi program (2000-2001) allowed Pratham to make significant progress in developing a testing instrument (the initial test was too difficult) and effective testing procedures to prevent cheating and exam anxiety. The test was administered with the authorization of the municipal corporation. At least three people are present in the classroom during the test to minimize cheating. To minimize attrition, Pratham returned to the schools multiple times, and children who still failed to appear and who could be tracked down were administered a make-up test outside of school.

Another outcome of interest is attendance and school dropout rates, which are collected weekly by Pratham employees (to avoid using the official rolls, which are often manipulated). In addition, we collected several intermediate outcomes for the CAL program. In particular, at every session, we collected data on the games played by the children, and at what level they were played. This data will be analyzed in the next draft of this paper.

3.3 Statistical Framework

Given the randomized allocation of both programs, we expected the results of the 2002 pretest to be similar between the CAL treatment and control groups. Noting y_{ijk} the test score of child i in school j on test k (k is either “pre”, “mid” or “post”), we start by comparing test scores in the pre-test and either the mid-test or the post-test, and we run the following regression to assess the significance of any differences:

$$y_{ijPRE} = \alpha + \beta C_j + \varepsilon_{ijPRE}\ , \qquad (1)$$

where C_j is a dummy indicating whether school *j* receives the CAL program. This regression is run separately for the math exam, the verbal exam, and the total score on

[1] There is no ambiguity as to whether a school could or could not have received the treatment if it had been in the control group, since the treatment and control schools were switched for the second year of the program. Schools that had to be excluded in 2003-2004 are also excluded from the treatment group in 2002-2003. An alternative strategy is to keep all schools in the sample, irrespective of whether or not the program could have been run there, and to scale up the difference between the treatment and control school by dividing it by the probability of being treated. The results are essentially the same.

the exam. The standard errors are clustered at the school level.[2] We then run the same regression in the mid-period (k = MID) and the post-period (k = POST):

$$y_{ijMID} = \alpha + \beta C_j + \varepsilon_{ijMID}, \tag{2}$$

$$y_{ijPOST} = \alpha + \beta C_j + \varepsilon_{ijPOST}, \tag{3}$$

This provides a first estimate of the program effect. Because tests scores are very strongly auto-correlated, the precision of the estimate is increased by relying on differences-in-differences estimates. This estimate also controls for any pre-existing differences between the treatment and the control group.

$$y_{ijk} = \lambda + \delta C_j + \theta POST + \gamma\left(C_j * POST\right) + \varepsilon_{ijk}. \tag{4}$$

Again, this regression is run separately for the math, verbal, and the total score on the exam, and the standard errors are clustered by school. We run the same regression for the differences from pre to mid-test and from mid to post-test. In the absence of large differences between the treatment and the control groups, the coefficients β in equation 3 and γ in equation 4 should be similar.

Since the Balsakhi (B_j) and the CAL (C_j) programs were run in the same schools, we can estimate the effects of each program together in a single regression:

$$y_{ijk} = \lambda + \delta_1 B_j + \delta_2 C_j + \theta POST + \gamma_1\left(B_j * POST\right) + \gamma_2\left(C_j * POST\right) + \varepsilon_{ijk}, \tag{5}$$

This equation provides estimates of the average effect of each program, controlling for receiving the other one. Because the probability of receiving the CAL program conditional on receiving the balsakhi program is, by construction, the same as the probability of receiving the CAL program conditional on NOT receiving the balsakhi program, the estimate of the effect of the CAL program should not be affected. However, the estimates of each program will be more precise, since the regression controls for the "noise" introduced by the other program. Finally, we can examine whether the CAL program and the Balsakhi program have interaction effects, using the following specification:

$$y_{ijk} = \lambda + \delta_1 B_j + \delta_2 C_j + \theta POST + \gamma_3\left(B_j * POST\right) + \gamma_4\left(C_j * POST\right) + \gamma_5\left(B_j C_j * POST\right) + \varepsilon_{ijk} \tag{6}$$

In this regression, γ_3 is an estimate of the effect of the balsakhi program in schools where there is no CAL program, γ_4 is an estimate of the effect of the CAL program in schools where there is no balsakhi program, and γ_5 is an estimate of the difference between the effect of the CAL program in schools that have a balsakhi and in schools that do not have a balsakhi. Finally, we estimate the effect of the CAL program in the two sub-samples with and without balsakhi and the effect of the Balsakhi program in the sample with and without the CAL program.

4 Results

[2] If instead we used a random effect model, with a nested random effect at the school and division level, the point estimates are very similar, and the estimated standard errors are smaller, making the results more significant.

4.1 Descriptive Statistics: Level of Competencies and Pre-intervention Differences

Tables 2 presents the descriptive statistics of all the test scores for the treatment and control samples used in this analysis. The scores are presented both as raw scores (out of 50 points for each of the math and verbal tests) and normalized relative to the distribution of the pre-test score in the control group.[3]

The randomization appears to have been successful; there are no large or systematic differences between the pre-test scores of the treatment and control group. Children in the control group perform slightly better than the treatment group on the math, verbal and total scores, but none of the differences between the groups prior to the implementation of the program are significant.

The raw scores give an idea of how little these children actually know. The average math score on the pre-test is about 15 (or 30%), both in the control and treatment groups. Since one math question is multiple-choice, on average a student who knows nothing will score 0.9 points. If a student can consistently order two numbers and add two single digit numbers, she can earn 7 additional points and get half way to the average fourth standard performance. Children in the bottom of the distribution of pre-tests score on average between 11 and 12 points on the two tests combined, indicating that they do not know how to count and have trouble even copying sentences, a task that requires no comprehension of the actual words.

4.2 Attrition

Table 3 presents the attrition that occurred between the pre-test and the mid-test or post-test, broken down by treatment status. Attrition was slightly but insignificantly higher in the comparison group than in the treatment group: attrition was 4.2% in the CAL treatment group and 6.1% in the control group at the post test. The fact that there was no differential attrition rate in the treatment and control groups suggests that the estimate of the treatment effects should not be biased, unless different types of people drop out from the sample in the treatment and the control groups (Angrist, 1995; and Powell, 199X). This does not seem to be the case here: the second row in each panel presents the difference between the score at the pre-test of children who were not present at the post-test, by treatment status. The third column presents the differences-in-differences in the treatment and comparison groups. Children who will eventually leave the sample tend to be at the bottom of the distribution of the pre-test scores in the treatment group, but not necessarily for the control group.

4.3 Effects of the CAL Program

Table 2 presents the first estimates of the effect of the CAL program as simple differences between the treatment and control groups' scores on the mid-test and post-test. On the math test, the difference between treatment and control groups is positive for both the mid-test and the post-test, although the difference is much larger for the post-test. The difference in math test scores is an insignificant 0.05 standard deviations on the mid-test, but jumps to a highly significant 0.33 standard deviations on the post-test.

---

[3] We subtract the mean of the control group in the pre-test, and divide by the standard deviation.

On the verbal test, the difference is insignificant for both the mid-test and the post-test. For the total score, the difference between the treatment and control groups is still significant (0.20 standard deviations). Overall, this is a substantial difference, by the standard of most interventions in the education literature. Among children who were at the bottom of the distribution of the pre-test scores, the differences between treatment and control groups on the total score is only 0.10 standard deviations on the mid-test but jumps to 0.31 on the post-test. The middle third of the distribution demonstrates a smaller, but still significant effect, while the difference for the top third is positive, but insignificant.

Because test scores have a strong persistent component, the precision of these estimates can be improved significantly, however, by turning to a differences-in-differences specification (equation 4). Since the randomization appeared to be successful, and almost all the children who took the post-test also took the pre-test, the point estimates should be similar in the simple differences and the differences-in-differences specification. The confidence intervals should, however, be tighter. Table 4 presents differences-in-differences estimates of the effects of the CAL program, for various sub-groups.

Overall, the CAL program always increased math test scores significantly, and most of the gains seem to have occurred during the second part of the year. The second year of the program will help us determinate whether this is due to the fact that the children need longer exposure to the computers to really benefit from it, or whether it is due to the fact that the program was just starting and improved during the year. In addition, the program did not increase verbal scores significantly in any group for any interval of time. Over the entire year, the program increased average math scores by 0.37 standard deviations for all students. Math scores among the bottom, middle and top thirds of the pre-test distribution rose by 0.43, 0.35, and 0.33 respectively. Boys and girls seemed to benefit equally (0.38 versus 0.37). Total scores demonstrate an effect of 0.22 standard deviations for all students. By thirds in the pre-test distribution, the effect of the program is 0.29, 0.18 and 0.17. Again, boys and girls benefit equally (0.23 versus 0.22).

4.4 Comparison with the balsakhi program

Table 7 presents the estimates from the specifications comparing the balsakhi and CAL programs. The effect of the CAL program hardly changes at all when we control for the balsakhi program. The effect of the CAL program, γ_2, is again 0.37 standard deviations for the math test, again insignificant for the language test and 0.21 standard deviations for the total score. Similarly, these estimates do not change when we include the interaction between the balsakhi and CAL programs. Compared to the balsakhi program, the effect on the math test is similar, but the CAL program had no effect on language competencies while the balsakhi program had a significant effect. As the interaction term is insignificant on either test, we conclude that the balsakhi program does not make the CAL program more or less effective and vice versa.

We can also compare the cost of one year of the balsakhi program with one year of the CAL program. The cost per student per year of the balsakhi program is 107 rupees, or approximately 2.25 dollars. The recurring expenditures of the CAL program are 367 rupees, but the cost of the CAL program including the start-up costs of the computers and software (assuming they are depreciated over five years) is 722 rupees. Thus, using the estimates from equation 5, we can calculate the relative cost effectiveness of each program. CAL increases the math score by 0.37 standard deviations and the overall test score by 0.21 whereas the Balsakhi program increases the math score by 0.25 and the total score by 0.24. Since CAL costs 6.7 times as much as the Balsakhi program per student, the Balsakhi program is 4.6 times more cost effective for math and 7.7 times more cost effective for the total score.

5 Conclusion

This paper reports the preliminary results of a computer assisted learning program. While the first semester of the intervention showed very insignificant effects, the overall effect over the entire year showed substantial improvements over the first four months. Average scores on a 50-point math test rose from 14.9 to 29.0 in the treatment group but only from 15.5 to 25.0 in the control group. This intervention had a somewhat bigger effect on the bottom third of the students, which might have been due to the selection of competencies reinforced by the software. The use of computer assisted learning, however, has the capacity to allow each child to move through the competencies at their own pace. The second year of the CAL program is experimenting with allowing children to move forward immediately after mastering a particular competency.

Since the computer games that the children played focused exclusively on math competencies, it is understandable that the language test scores showed no effect from the intervention. Unfortunately, this indicates that the more interactive, computer-based approach to learning might not have created a greater enthusiasm for learning overall. After analyzing the attendance data, we will be able to track this secondary effect more clearly.

Table 1: Sample Design

	Intended Number of Schools	Actual Number of Schools	Number of Divisions	Number of Children
	(1)	(2)	(3)	(4)
CAL	61	55	87	2823
No CAL	61	56	90	3096

Table 2: Summary Statistics: Written Tests

	PRE TEST			MID TEST			POST TEST		
	Treatment	Control	Difference	Treatment	Control	Difference	Treatment	Control	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Number of Tests	2823	3096	-273	2584	2879	-295	2704	2907	-203
Average Score (Points)									
Math	14.944	15.549	-0.605	24.555	23.937	0.617	28.997	25.090	3.907
			(0.730)			(0.850)			(0.898)
Verbal	16.987	17.068	-0.082	26.066	26.074	-0.007	24.563	24.353	0.211
			(0.703)			(0.821)			(0.834)
Total	31.931	32.618	-0.687	50.621	50.011	0.610	53.560	49.443	4.117
			(1.348)			(1.609)			(1.673)
Average Score (Normalized)									
Math	-0.049	0.003	-0.052	0.770	0.717	0.053	1.148	0.816	0.333
			(0.062)			(0.072)			(0.076)
Verbal	-0.011	-0.003	-0.008	0.876	0.877	-0.001	0.729	0.709	0.021
			(0.069)			(0.080)			(0.081)
Total	-0.033	0.000	-0.033	0.874	0.845	0.030	1.017	0.817	0.200
			(0.065)			(0.078)			(0.081)

Table 3: Attrition Patterns

	CAL	NoCAL	Difference
	(1)	(2)	(3)
Pre-Test to Post-Test			
Percent attrition	0.042	0.061	-0.019
			(0.020)
Difference in normalized score at pretest: attriters-stayers	-0.207	0.059	-0.266
			(0.125)
Pre-Test to Mid-Test			
Percent attrition	0.085	0.070	0.015
			(0.017)
Difference in normalized score at pretest: attriters-stayers	-0.077	-0.187	0.110
			(0.112)
Mid-Test to Post-Test			
Percent attrition	0.018	0.041	-0.023
			(0.018)
Difference in normalized score at pretest: attriters-stayers	-0.321	0.072	-0.394
			(0.151)

Table 4: Difference-in-difference estimate of the impact of the CAL program (OLS and IV)

	Math (1)	Verbal (2)	TotalScore (3)
OLS			
Pre-Test to Post-Test	0.375	0.014	0.221
	(0.085)	(0.073)	(0.080)
Pre-Test to Mid-Test	0.114	0.009	0.070
	(0.069)	(0.068)	(0.068)
Mid-Test to Post-Test	0.275	0.010	0.162
	(0.085)	(0.077)	(0.083)
IV			
Pre-Test to Post-Test	0.374	-0.002	0.212
	(0.090)	(0.080)	(0.086)
Pre-Test to Mid-Test	0.114	-0.009	0.061
	(0.075)	(0.075)	(0.075)
Mid-Test to Post-Test	0.271	0.008	0.158
	(0.088)	(0.081)	(0.086)

Table 5: Difference-in-difference estimate of the impact of the CAL program

	Math (1)	Verbal (2)	TotalScore (3)
Pre-Test to Post-Test			
Bottom Third	0.428	0.088	0.287
	(0.109)	(0.090)	(0.103)
Middle Third	0.352	-0.029	0.186
	(0.091)	(0.082)	(0.083)
Top Third	0.329	-0.025	0.175
	(0.088)	(0.090)	(0.088)
Pre-Test to Mid-Test			
Bottom Third	0.116	0.024	0.078
	(0.089)	(0.106)	(0.098)
Middle Third	0.088	-0.059	0.021
	(0.084)	(0.080)	(0.077)
Top Third	0.106	0.046	0.083
	(0.074)	(0.075)	(0.072)
Mid-Test to Post-Test			
Bottom Third	0.302	0.046	0.195
	(0.127)	(0.115)	(0.124)
Middle Third	0.257	0.013	0.153
	(0.093)	(0.086)	(0.089)
Top Third	0.247	-0.049	0.117
	(0.078)	(0.075)	(0.076)
Pre-Test to Post-Test			
Boys	0.381	0.023	0.229
	(0.098)	(0.082)	(0.091)
Girls	0.373	0.010	0.217
	(0.110)	(0.091)	(0.101)
Pre-Test to Mid-Test			
Boys	0.125	0.034	0.088
	(0.092)	(0.092)	(0.093)
Girls	0.101	-0.016	0.050
	(0.077)	(0.079)	(0.074)
Mid-Test to Post-Test			
Boys	0.265	0.005	0.153
	(0.109)	(0.103)	(0.109)
Girls	0.290	0.021	0.175
	(0.103)	(0.087)	(0.097)

Table 6: Comparing the results of the CAL program and the Balsakhi program

	Math (1)	Verbal (2)	TotalScore (3)
Equation 4 (and the equivalent for Balsakhi)			
CAL	0.375	0.014	0.221
	(0.085)	(0.073)	(0.080)
Balsakhi	0.265	0.184	0.242
	(0.092)	(0.072)	(0.082)
Equation 5			
CAL	0.367	0.008	0.213
	(0.083)	(0.071)	(0.078)
Balsakhi	0.253	0.184	0.236
	(0.084)	(0.073)	(0.079)
Equation 6			
CAL	0.457	0.012	0.267
	(0.102)	(0.111)	(0.108)
Balsakhi	0.343	0.188	0.289
	(0.128)	(0.122)	(0.129)
Interaction	-0.186	-0.008	-0.110
	(0.160)	(0.179)	(0.174)