The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia

Citation	Barrera-Osorio, Felipe, and Leigh L. Linden. 2009. The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia. World Bank Policy Research Working Paper 4836.
Published Version	http://go.worldbank.org/BZZT7KNLG0
Accessed	February 19, 2015 9:24:17 AM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:8140109
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

WPS4836

Policy Research Working Paper 4836

Impact Evaluation Series No. 29

The Use and Misuse of Computers in Education

Evidence from a Randomized Experiment in Colombia

Felipe Barrera-Osorio
Leigh L. Linden

The World Bank
Human Development Network
Education Team
February 2009

Policy Research Working Paper 4836

Abstract

This paper presents the evaluation of the program Computers for Education. The program aims to integrate computers, donated by the private sector, into the teaching of language in public schools. The authors conduct a two-year randomized evaluation of the program using a sample of 97 schools and 5,201 children. Overall, the program seems to have had little effect on students' test scores and other outcomes. These results are consistent across grade levels, subjects, and gender. The main reason for these results seems to be the failure to incorporate the computers into the educational process. Although the program increased the number of computers in the treatment schools and provided training to the teachers on how to use the computers in their classrooms, surveys of both teachers and students suggest that teachers did not incorporate the computers into their curriculum.

This paper—a product of the Education Team, Human Development Network—is part of a larger effort the network to estimate the impact of Information and Communication Technologies on education. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at fbarrera@worldbank.org.

The use and misuse of computers in education: Evidence from a randomized experiment in Colombia[1]

Felipe Barrera-Osorio
(World Bank)

Leigh L. Linden
(Columbia University)

JEL: C93, I21, I28
Keywords: education, computer programs, randomization

[1] We are grateful to the program Computers for Education, and the Ministry of Communication, for financial support. We thank past and current staff members of the program for their unconditional commitment to the evaluation, particularly María Isabel Mejía Jaramillo, Beatriz Eugenia Córdoba, Francisco Camargo, Martha Patricia Castellanos and Julián Gómez. Fedesarrollo provided personnel and financial help. We are especially grateful to Mauricio Olivera. Camilo Dominguez and Monica Hernandez provided outstanding research assistance. Funding from the Knowledge for Change Program at the World Bank, research grant RF-P101262-TF090460, helped in the final stages of the evaluation. The opinions expressed in this document do not necessarily represent the views of the World Bank.
Barrera-Osorio: fbarrera@worldbank.org; Linden: leigh.linden@columbia.edu.

I. Introduction

The use of information and communication technologies (ICT) in education is becoming a major consideration as developing countries focus on improving the quality of education. Investment in ICT use in education has grown steadily over the past decade in developing countries, even in the some of the most challenging environments in some of the least-developed countries. Several countries are determinedly expanding the supply of computers in their schools in the belief that schools will benefit from the use of the new technologies and that students need to be exposed early. For instance, several countries in Africa, Asia, and Latin America are considering a program to procure $100 laptops for schools. Despite the growing adoption of and demand for ICTs in education, there is very little systematic research and hard data about how ICT is actually used in the classroom and even less about its impact on educational outcomes, social behavior, or employment and worker productivity (InfoDev, 2005).

This article aims to increase the available evidence on the use and the impact of computers in education. We consider the program Computers for Education. The program is an alliance between the public and private sector to refurbish computers donated by private organization, install them in public schools, and run a program that teach teachers to use computer in specific subjects, especially in Spanish. This is an existing large-scale national program in Colombia.

Unfortunately, while ICT programs are one of the most studied interventions in the education literature, robust evaluations of ICT programs are still too scarce to provide general conclusions regarding their effectiveness. The results of the evaluations that do exist are at best mixed.

The evaluation literature of such programs is more abundant in developed countries. Importantly, a large portion of these studies are also correlational analyses for which there are obvious challenges to causal interpretation of the findings. Two studies in the U.S. written by the National Center for Educational Statistics (2001a and 2001b) found a positive relationship between availability of computers in schools and test scores. For the US, Wenglinsky (1998) measured the amount of computers that were used in math classes and scores on math tests and found a positive relationship between use of computers and learning in both 4th and 8th grades. Rouse and Krueger (2004) undertook a randomized study of a popular instructional computer program, known as Fast ForWord, which is designed to improve language and reading skills. They concluded that while the program may have improved some aspects of students' language skills, the gains do not translate into a broader measure of language acquisition or into actual readings skills.

Similar positive relationships have been found in OECD countries between computer use and test scores for mathematics (NCES 2001a, Cox et al., 2003), science (NCES, 2001b, Harrison et. al. 2002) and reading (Harrison et. al. 2002). Kulik (2003) reviews 75 impact evaluations of technology applications in the US, finding the following results, among others: (i) students who used computer tutorials in mathematics, natural science, and social science score significantly higher on tests in these subjects; (ii) the use of computer-based laboratories did not result in higher scores; and (iii) primary school students who used tutorial software tutorial in reading scored significantly higher on reading scores, while very young students who used computers to write their own stories scored significantly higher on measures of reading skills.

In developing countries, six existing studies present generally positive but still mixed conclusions. Linden et. al. (2003) designed an impact evaluation of a computer assisted learning program in Vadodara, India, on cognitive skills using mathematics and language tests. The authors find a positive and significant impact on math scores of 0.375 standard deviations. Similarly, Linden (2008) finds positive effects of a remedial math program when implemented on a supplemental basis and *negative* effects when implemented on a pull-out basis as a substitute for the regular classroom teacher's instruction. Fang, He, and Linden (2008) find strong positive effects on Indian students' English scores of an electronic English-based curriculum.

However, other evaluations do not find such consistent positive results. Angrist and Lavy (2002) find no effect in their evaluation of the 'Tomorrow-98' program, which placed 35,000 computers in schools across Israel between 1994 and 1996. They find no impact on math and Hebrew scores at the fourth or eighth grade level. Finally, the evaluation of the World Links program found positive impact for both students and teachers (Kozma, et. al, 2004, Kozma and McGhee, 1999). This program prepares students and teachers on communication, collaboration and Internet skills in African and Latin American countries. In Uganda, a special designed performance assessment found that World Links schools outperformed the non-World Links schools on measures of communication and reasoning with communication (Quellmalz and Zalles, 2000). However, unlike the other evaluations, these are based on correlational estimates rather than rigorous research designs.

While there is still much to be learned, one general result that seems to emerge from this literature is that positive outcomes of the use of computers in schooling are linked to changes in pedagogy, and introducing technology alone will not change the teaching and learning process. It is not enough to install computers in schools without training (InfoDev 2005). However despite this general result, very little is known about just how (and how often) ICTs are used in developing country classrooms when available. One study has shown, for example, that where ICTs are used for learning, they are chiefly used to present and disseminate information rather than change the way that children are taught (InfoDev, 2005).

Our study builds on this existing literature in several ways. First, because our study is a large scale randomized evaluation of a mature, well developed program, we add to the existing body of rigorous randomized evaluations of ICT programs in education. Second, within this evaluation, we implemented modules that are designed to understand, not just the effect of computers on students' test scores, but also the effect of the computers and associated training on the teachers teaching methods, including their use of computers in the classroom.

There are three main conclusions of this evaluation. First, the program successfully increases the number of computers in the school (by 15 computers) and increases students' use of the computers. Second, despite this success, the program has little impact on students' math and Spanish test scores. The program also has little effect on a host of other academic variables including hours of study, perceptions of school, and relationships with their peers. The reason seems to be that despite the program's focus on using the computers for teaching students in a range of subjects (but especially Spanish), the computers were only used to teach the students computer usage skills. The evidence suggests that students use of the computers for their intended purpose was limited -- only

3 to 4 percent of the students in both treatment and control groups reported to use the computers in the language class for example. Overall the results of this study highlight the importance of program implementation and measuring the impact of an intervention on the actual practice of teachers and the learning experiences of students.

The paper is organized as follows. In the next section we describe the program in more detail. In section three, we discuss the design of the experiment. Section four contains the results that verify the internal validity of the study and section five contains the results of the evaluation. Finally, we conclude in section six.

II. The Computadores para Educar Program

The Computadores para Educar (Computer for Education or CPE) was created in March 2002 by the Minister of Communications, with the objective of refurbishing computers donated by the private sector to install them in public schools. The program trains teachers in the use of computers in the classroom, especially in the areas of language. Since its creation, the program has received 114,541 computers, and refurbished 73,665 that have been installed in 6,386 public schools in 1,018 municipalities. To date, the program includes 83,092 teachers and more than 2 million students.

The program creates a partnership between schools and a local university. For our study, schools were paired with the Universidad de Antioquia. The university then designs a pedagogical strategy for the school and participates in the implementation of a 20 months training component directly in the schools for teachers. The start of the training coincides with the school's receipt of a set of refurbished computers. The initial phase of the training is provided by CPE program directly and lasts for 8 months and covers computers installation and adaptation of classroom techniques, which includes preparation of the rooms where the computers are installed and classroom management strategies. This phase also includes a first step towards the active use of computers in education through teachers' workshops.

The second phase lasts one year and is developed by the partner university to take into account the regional needs of the schools. The objective of this second phase is to train teachers and coordinators on the relationship between technology and learning. Among the objectives are: (i) support the education of the children in basic areas (language, math, natural and social sciences) by integrating the use of ITCs with pedagogic projects and activities, and (ii) encourage collaborative learning, creativity, and improve teachers' and students' confidence in the use of technology by integrating ITCs to their pedagogic processes (CPE, 2008).

The model designed and run by the Universidad de Antioquia focuses on Spanish education. The program is aimed at training teachers in teaching methodologies using computers to strengthen students' reading and writing skills through a theoretical socio-constructivist approach. Its objectives go beyond the impact on standardized tests. In particular, the program tries to integrate technology in learning pursuing the goal of "fomenting a socio-cultural vision over the reading-writing teaching and learning processes" (translated from Henao and Ramirez, no date available, p. 2). Specifically, the model focuses on different aspects of reading and writing, with a special focus on the recent developments of e-mail, Internet, and the Hypermedia formats.

III. The Design of the Experiment

A. Sample

In order to evaluate the program, we randomly assigned 100 interested and eligible schools into a treatment and a control group. Interested schools were selected as follows. First, to minimize program implementation and data collection costs, the schools were chosen to be in close geographic proximity: we chose schools in the north part of Antioquia, Caldas, Choco, Cordoba, Quindio y Risaralda. For all interested schools in this area, we further reduced the sample to schools to those with 80 or more students in order to facilitate the collection of data. From this list of 100 schools a CPE team visited each school to verify the number of students, the existence of a classroom that could be refurbish for the computers, and the type of school (public or private).

Once the final sample was created, we conducted a stratified randomization, stratifying on department and type of school—basic education, basic plus lower secondary, high secondary.[2] In this process half of the schools received the program and half were assigned to a control group which did not receive the program. The lottery was performed at the beginning of August 2006, and the list of schools, with their status, was given to the Ministry of Communication for the implementation of the program.

B. Data Collection

Data was collected in two phases: a baseline survey conducted immediately after the randomization but before the start of the treatment and a follow-up survey conducted two years later. The baseline was conducted between August 14 and September 29,

[2] Type refers to the grades covered by each school. Basic includes grades one through five. Lower secondary includes grades six through eight, and high secondary includes grades nine through eleven.

2006, with the target of reaching 100 schools. The collection of information was done directly at the school in an unannounced visit, and questionnaires were administered to students in grades three through eleven, their math and Spanish teachers, and the school's principal.

The survey company, SEI (Sistemas Especializados de Informacion), was able to gather the information from 97 schools: three schools (two treatment and one control) declined to participate in the data collection activities. Because the follow-up survey was completed two years later, we focus on students who were in grades through nine at baseline – students in grades ten and eleven at baseline would have either dropped out or graduated by the time the follow-up survey was administered. The first three columns of Table 1 contain a tabulation of these schools by research group. In total, the distribution of schools and students is even. The sample contains forty-nine control schools and forty-eight treatment schools, and 3,889 control students and 4,327 treatment students in grades three to nine. Dividing the sample by gender and grade shows a similarly even distribution of students. Across all of the groups, the largest difference is the 182 additional male students in the treatment group.

During May and June of 2008, the same survey company resurveyed the same 97 schools, focusing on the students who were in grades three through nine in 2006. The final sample comprises 5,201 students that were found at follow up; 3,015 students were attritors, yielding an attrition rate of 37 percent. Discussions with local principals suggest that the high rate of attrition is primarily due to high rates of migration in the rural areas chosen for the research project. A high attrition rate alone does not violate the internal validity of a study as long as the types of students who attrit are similar in each research groups (Angrist et al., 2002). While we investigate this in more detail in section five below (and find the attrition rates similar), the distribution of non-attriting students is presented in the last three columns of Table 1. Like the original baseline tabulation, the distribution of students is still evenly distributed between the research groups even at follow-up, despite the high attrition rate.

Both the baseline and follow-up questionnaires followed the same format. Three different questioners were administered respectively to students, math and Spanish teachers and the principal of the school. The students' questionnaire was self-administered by the students with the assistance of the survey team. The variables included in the students questionnaire were socioeconomic characteristics (age, gender, family structure, number of siblings, work status, and allowance, among others,); school outcomes (attendance last week, number of hours of study outside the school, grades in math and language, repetition, and dropout spells in the past, among others); and attitude towards the school and the content of classes, including the use of computers. Finally, the student survey included a shortened version of the national Colombian exam, the *Saber.*[3] For the purpose of analysis, the scores on these tests are normalized relative to the grade-specific control group distribution of scores.

Information was also collected from all math and language teachers present at the school the day of the visit. We collected background information (age, gender, education, experience, etc.), information on their knowledge of computers, and their use of

[3] Four different tests were used, depending on the student's grade, grouping them as follows depending on their grade at the time of the survey: 1) third and fourth, 2) fifth and sixth, 3) seventh and eighth, and 4) ninth, tenth and eleventh. These are national Colombian exams normally administered to students in the odd grades from grade three to grade nine. As a result, we simply administer the exam normally given to a particular grade to that grade and the subsequent grade with the exception of grade eleven which receives the grade nine test (since there is no grade eleven test).

computers in class. Finally, the questionnaire of the school was answered by the principal or the coordinator of the school. It includes general variables at the school level.

C. Statistical Models

Because the treatment was randomly assigned, we can assess the causal effects of the program by directly comparing the average outcomes of students in the treatment group with students in the control group. To do this, we employ three statistical models estimated through ordinary least squares. First, we use a simple difference model of the following form:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \varepsilon_{ij} \tag{1}$$

The variable Y_{ij} in this specification is the variable whose average value is to be compared between the two groups at the child (i), school (j) level. The variable T_j is a dummy variable for whether or not school j was selected for treatment in the randomization process. The coefficient β_1 then provides the estimated differences in the variable Y_{ij} between the treatment and control group. This specification is primarily used for two purposes: to compare the treatment and control group using information collected in the baseline survey to check for comparability and to conduct simple comparisons of outcome variables (without controlling for baseline characteristics).

Equation (1) can also be used to estimate the treatment effects, but we can improve the precision with which we estimate the treatment effects by controlling for baseline characteristics that are correlated with the outcome of interest. This is done with the following specification which is also estimated through ordinary least squares:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 X_{ij} + \varepsilon_{ij} \qquad (2)$$

The specification is identical to equation (1) with the addition of a vector, X_{ij}, of baseline and school characteristics.

In Table 3, we compare the relative characteristics for students that attrit from the baseline at follow-up and students that do not. To make this comparison we use a difference in differences estimator that compares the difference between attritors and non-attritors in the treatment group to the same difference in the control group. The model is estimated through ordinary least squares using the following specification:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 Attrit_{ij} + \beta_3 Attrit_{ij} * T_j + \varepsilon_{ij} \qquad (3)$$

The variables Y_{ij}, T_j, and ε_{ij} are all defined as in equation (1) and $Attrit_{ij}$ is a dummy variable set to 1 if student i in school j did not appears in the follow-up. The coefficient β_3 in this model is then an estimate of the difference in characteristics between attritors and non-attritors across the treatment and control groups.

Finally, because treatment assignment occurs at the school level, student behavior is likely correlated within schools. Not taking this correlation into account, could lead us to overestimate the precision of the treatment effect estimates and conclude that an effect exists when, in fact, it does not. To account for this in both models, we cluster the standard errors ε_{ij} at the school level (Bertrand, Duflo, and Mullainathan, 2004).

IV. Internal Validity Confirmation

In order to ascribe any observed difference in follow-up characteristics to the implemented program, the two groups created by the randomization must be sufficiently similar in all characteristics that such differences could not be responsible for generating the observed differences in the groups at follow-up. We proceed as follows. First, using all of the information students provided in the baseline, we verify that the randomization process did create initially similar treatment and control groups. Some of those students observed in the baseline, however, then fail to provide information at follow-up, having attritted form the sample. Because these students are not available at follow-up, the attrition process could cause an initially balanced sample to become unbalanced if different types of students attrited from one of the two groups. To check this, we compare the characteristics of attriting students between the two groups and compare non-attriting students using the available baseline characteristics.

A. Baseline Comparison

To start, we compare the respective research groups based on the survey responses at baseline. Table 2 contains the comparison of all students taking a baseline survey in treatment and control schools. Panel A contains the test score characteristics and the rows in Panel B contain basic demographic characteristics. Panel C then contains other relevant academic variables. We present the average of these characteristics for the treatment group in the first column. Column two contains the control group average. And the final column contains the estimated difference using equation (1).

The differences between the treatment and control groups are minor. Of nineteen considered variables three are statistically significant and all of them are practically small. This includes a three percent difference in the probability that students report staying after school, an eight percent difference in the probability that a child talks to a teacher after class, and a 0.11 difference in the number of hours reported studying. The most significant variables – the math and Spanish test scores – show no difference between treatment and control groups. The normalized difference in Spanish scores is 0.07 standard deviations and the difference in math scores is 0.06 standard deviations.[4]

B. Attrition

Even if the treatment and control groups are similar at baseline changes in the composition of the groups between baseline and follow-up could result in large differences between the treatment and control students who complete a follow-up survey. And as long as these attrition patterns are similar across treatment and control groups, the research design would still have internal validity even with extremely high attrition rates (Angrist et al., 2002). We compare the attrition pattern between the two group using baseline characteristics in Table 3 and 4. Table 3 contains a direct comparison of attriting students. Using the same rows as in Table 2, columns one through three contains a simple comparison of attriting students using equation (1). Columns four and five

[4] A more detailed comparison is provided in a baseline report produced after the baseline data collection (Barrera-Osorio et al, 2006). The report compares the samples in much greater detail than is presented here concluding the that the samples are sufficiently similar. This report is available from the authors upon request.

contain the relative difference between attriting and non-attriting students in the treatment and control groups respectively. Column six then contains the estimated difference in these differences between treatment and control groups using equation (3). Thus, the first three columns compare the characteristics of attriting students while the last three columns compare the relative characteristics of attriting and non-attriting students across the two research groups.

The results in Table 3 confirm that the same types of children attrited from the treatment and control groups. Even though on average 35 and 38 percent of baseline students attrited from the treatment and control groups respectively, the types of attriting students are exactly the same. First, the attrition rates are very similar, differing by only 3 percentage points. Second, in almost all characteristics, the differences between the treatment and control groups are minor. The attritors differ in only two characteristics – whether students talk to a teacher outside of class (7.5 percentage points) and the number of hours students study outside of school (0.13 hours) – both of which are small.

Columns four through six then compare the characteristics of attriting and non-attriting students. In both the treatment and control groups, attriting students are significantly different from non-attritors. Attritors are about three-quarters of a year older than non-attritors. They are 11 percentage points more likely to have repeated a previous grade, and 9 to 11 percentage points more likely to have failed a class. They are also more likely to work. The patterns, however, are exactly the same in the treatment and control groups. As shown in column six, the difference in these patterns is only significant for the number of friends.

Table 4 estimates the ultimate outcome of the attrition process by comparing the non-attriting students using the baseline characteristics. The table layout is identical to Table 2. As one would expect from the previous tables, the students remaining in the study after the follow-up survey are similar in their baseline characteristics. In fact, the attrition pattern was so similar, that the results in column three of Table 4 are almost exactly the same as those in column three of Table 2 despite having lost almost a third of the sample. Focusing on the differences in column three, all of the differences are small, and only two differences are statistically significant. This includes an 8.5 percentage point difference in the probability that a child reports seeing a teacher outside of class and a 3.4 percentage point difference in the probability that a child stays after school.

V. Results

Given that the sample of students completing the survey in both research groups is comparable, we can estimate the causal effects of the program by directly comparing the average responses in the treatment and control groups. First, we check to ensure that the treatment was implemented as planned and that the implementation created the expected treatment differential between the two groups – it did. Second, we estimate the differences in outcomes. Overall, the program seems to have had little effect on students' test scores and other outcomes. These results are consistent across grades, subjects, and different types of students. The main reason for these results may be the implementation of the program. Surveys of both teachers and students suggest that the program increases computer use among students and teachers by a surprising small amount, and most of the use of computers by students is for the purposes of learning to use a computer rather than studying language. Additionally, the extra computer use reported by teachers is concentrated in the lower grades with older students' teachers reporting almost no computer use in both groups.

A. Treatment Differential

As described in section four, the two research groups contained similar students at the time of the follow-up exam. The research design then depends on the only difference being the treatment delivered to the schools in the treatment group. Table 5 estimates the differences in the implementation of the treatment at the school level by research group using equation (1). As before, columns one and two contain the average characteristics for the treatment and control groups respectively and we present the estimated difference in the third column. We use three measures of implementation to compare the schools. First, using information from CPE staff regarding whether or not the program was implemented in a specified school, we estimate the average number of schools receiving the treatment in the first row. The estimates document that the treatment assignment was very closely followed by CPE. On average 96 percent of the treatment schools were treated and only 4 percent of the control schools were treated.[5]

We also checked for the implementation of the two main components of the CPE program by CPE and the Universidad de Antioquia – the provision of computers and the training of teachers. Using data from the principals, we estimate in row two the average number of computers in each school. On average, treatment schools have 13.4 computers compared to only 5.1 computers in the control schools. This suggests that the program

[5] Despite this very high compliance rate, we also estimated all of the estimated differences in outcomes using a treatment on the treated model to account for the existing non-compliance. As expected, the estimates are virtually identical to those presented in Tables 6 through 13.

was successful at increasing the computer resources available to students in the treatment schools. Finally, in row three, we report the average responses of the surveyed teachers regarding whether or not they received training from the university. In all, 95 percent of treatment teachers reported receiving training while only 8 percent of the control group reported receiving training. These results document that the program implementation complied with the research design and generated the expected treatment differential between the two research groups.

B. Effects on Test Scores

Given that the groups are on average similar in characteristics except for the receipt of the treatment by the treatment group, we can estimate the effects of the treatment by comparing the average characteristics of the research groups at follow-up. Tables 6 through 9 present results of the impact of the program on test scores.

Table 6 contains the primary outcome measure – the average score on the test administered as part of the follow-up survey. The first four columns of Table 5 presents percentage of correct answers, and the second four estimates the same treatment effects using the test scores normalized by the grade-specific control distribution. Within each group, we first present the average score for the treatment group, the average score for the control group, and then the simple difference estimated using equation (1). Finally, in the fourth column we estimate the differences controlling for baseline characteristics using equation (2).

It is important to note that the estimated respective differences between the estimates from equation (1) in column three and seven (without controls) and the estimates from equation (2) in columns four and eight (with controls) are virtually identical. This underscores the similarity of non-attriting sample. Had there been large differences in characteristics that were correlated with the follow-up test scores, then the estimated differences would have differed significantly. The fact that the estimates are similar is another piece of evidence supporting the comparability of the research groups at follow-up.

The value of the estimated differences shows that, by all measures, the program seems to have had little effect on students' test scores. In Spanish, the subject targeted by the program, students answer on average about 40 percent of the questions correctly, but the treatment group only answers correctly 1.7 percent more than the control group. This difference is too small to be statistically significant at conventional level of significance. Similarly, the program has no ancillary effect on math (neither positive nor negative).

To facilitate comparing the magnitude of these results to those of other studies, the second four columns make the same comparison using normalized scores. The average treatment effect in both subjects is less than 0.1 standard deviations. This is significantly smaller than the estimated effects of other successful programs in the literature. In addition, the estimated standard errors suggest that the experiment was powerful enough to detect treatment effects of at least 0.125 standard deviations (at the ten percent level) which is comparable to the smallest estimated treatment effects for education programs in developing countries (Kremer, Miguel, and Thornton, 2007). In addition, we can reject the hypothesis (at the ten percent level) that the true effect of the program in Spanish is greater than 0.2 standard deviations. Since many viable classroom interventions have increased test scores by 0.3 to 0.47 standard deviations (for example, Muralidharan and Sundararaman, 2008; Banerjee, Cole, Duflo, and Linden, 2007; He, MacLeod, and Linden, 2008; Linden 2008), the test is powerful enough to exclude treatment effects that would place the program in the range of other programs that educators would consider when choosing interventions.

Despite the lack of overall average effects, it is possible that the program may have had significant impacts on individual subjects (as in Glewwe, Kremer, and Moulin, 2007). Table 7 presents the estimated treatment effects by specific components. This includes encyclopedia, identification, paraphrase, pragmatics and grammar in Spanish[6] and of algebra, arithmetic, geometry and statistics in mathematics. As in Table 5, we present the average treatment score, the average control score, the difference estimated with equation (1), and the difference estimated with equation (2). All scores are normalized to the control group distribution. The results are similar to the overall average estimates presented in Table 5. For all competencies, we did not detect any differences across sub-subjects between students in treatment and control schools. The largest estimated difference is 0.13 for letter identification, a difference that is not statistically significant at conventional levels.

Table 8 presents differences on the test by gender. Replicating the format of Table 6, we present the treatment effect measures in normalized test scores. Panel A contains the results for female students and Panel B contains the results for male students. For no subject or group of students are the results different than the overall averages

[6] The encyclopedia component requires blending previous knowledge and information presented in a sample text to answer questions. The identification component asks respondents to repeat information explicitly presented in the text without changing its meaning. The paraphrase questions require recognizing both explicit and implicit information in the text and being able to convey the information back in written form. The pragmatics questions ask students to recognize and understand different types of communicative actions, intentions, aims and purposes of the author and the circumstances under which the text is written. The grammar component asks students to recognize and understand the semantic function of grammatical elements in coherence and cohesion of the text (MEN, 2008).

presented in Table 6. For Spanish, girls show a difference in test scores of 0.069 standard deviations while boys show a difference of 0.087 standard deviations. Neither of these estimates is statistically significant at conventional levels of significance.

Finally, Table 9 divides the sample by grade. Using the reported grade levels from the 2006 baseline survey, we report only the estimated treatment effects using equation (2) for the test scores normalized to the follow-up control distribution for the indicated subject. The results generally confirm the estimates presented earlier. In almost all grades, the estimated differences generally small and statistically insignificant at conventional levels of significance. The exception is students in grade eight and nine. Students in grade eight seem to perform worse due to the program while students in grade nine seem to benefit from the program. However, since these are the only statistically significant results, they are likely due to random variation. The fact that, as we will show below in Table 12, computers were not used in these grades further supports this contention.

Table 10 shows the results for outcome variables other than test scores. The table is organized similarly to Tables 6 and 8 with the differences estimated using equation (1). For most variables, the program seems to have little effect. There is no change in the probability that students like school or like the content of school. Similarly, there is no change in students reported grades or probability of failing a subject. And similarly, there is no change in students' probability of talking to a teacher outside of class. The only statistically significant effects are the probability that students did not attend school in the previous week (reduction of 0.12 percentage points), the probability that a student stayed after school (reduction of 2.8 percentage points), and the reported number of hours worked (reduction of 1.735 hours). However, while significant, they are small.

C. Number and Use of Computers

To better understand the lack of estimated treatment effects on student outcomes, we turn to our measure of computer utilization by teachers and students. A critical objective of the CPE program is to incorporate computers into actual classroom teaching at the school. Indeed, the program aims for the use of computers in Spanish classes. The lack of consistent treatment effects – especially on Spanish test scores – seems to be at least partly the results of teachers' and schools' limited use of the program. Teachers report only a modest increase in computer use – an increase that is concentrated in the lower grades. And students report that their experience of computers due to the treatment – while also modest – occurred in computer science and not in Spanish as originally envisioned by the creators of the program.

Table 11 presents the responses of 426 teachers to our follow-up questionnaire. As before, we show the average response for the treatment group, the average response for the control group, and the estimated difference using equation (1). The responses are disaggregated by the subject that the teacher teaches. The surveys targeted language and math teachers, and if a teacher taught both subjects, then the teacher's responses are included in both the language and math teachers' responses.

The data show that the program created a modest increase in computer use amongst teachers. In response to the question about whether or not they used a computer last week, 42 percent of the treatment language teachers responded affirmatively compared to 17 percent of the control group. This is a difference of 25 percentage points which is statistically significant at the one percent level. The fact that some schools in the control group use computers without the program is hardly surprising, but it does seem surprising that only 44 percent of teachers in the treatment group report using a computer. Interestingly the results were the same for math teachers, suggesting that the computers provided by the program were not used exclusively for the Spanish program, but rather, were used by teachers in general.

The next set of questions attempt to identify what the teachers used the computers to do. For each topic, teachers were asked how many times they used a computer for the identified task in the previous week. Again, teachers responded that they used the computers in the treatment group (for both language and math) about a half a day more a week than the control group, differences that are both statistically significant at the five percent level. However, when asked how many times the computers were used for in class activities, the teachers responses indicate that the computers were not used more often in the treatment group for these purposes. Teachers, for example, may have been using the computer for preparatory activities rather than for in class activities.

Using the information on which teachers did and did not use computers, we can in turn ask for each school which grades had a teacher that reported using a computer in the previous week. These results are presented in Table 12. The format is the same as for Table 10, except that the analysis is conducted at the school level, difference are again estimated using equation (1). The results suggest that for both language and math, the program generated changes in schools' use of computers only in the early grades. In Spanish, for example, the program seems to increase computer in only 20 and 16 percent of the schools respectively in grades three and four, both differences are statistically significant at the five percent level. However, in higher grades, the differences are much smaller. There is a difference of 10 percentage points in the fifth grade that is not statistically significant at conventional levels, and then the difference for all other grades is statistically insignificant and less than 10 percent in magnitude.

We also included questions for students regarding computer use, and their results corroborate the data from the teachers. This data are presented in Table 13. In the first three columns we present the average responses for the treatment group, the control group, and the difference (estimated using equation (1)) for all 5,201 students responding to both the baseline and follow-up surveys. In the second three columns, we present the same three statistics, but only for students who report having used a computer in the last week. Thus the first three columns show the overall average rates of use for each individual purpose across all students, and the second column identifies differences in computer use between the treatment and control students for students using a computer.

Similar to the responses for the teachers, about 25 percent more students report using a computer in the last week, an increase to 66 percent for the treatment group versus 41 percent for the control group. Focusing on where students used a computer, the increase in the treatment group seems to be largely due to increased computer use at school (a 30 percent increase) which is largely consistent with the idea that the use is generated by the provision of computers through the CPE program. There is no difference in this respect between students in general and only students using computers.

Given that students use the computers more often at school, which subjects do they use them for? Panel B presents estimated differences in the subjects in which students' report using computers. It seems that the largest (and only statistically or practically significant change) is in the use of computers for computer science class. The magnitude of this result, 29 percent, is consistent with the overall reported use of computers for both groups of students. This is also consistent with the data from teachers suggesting that they did not report an increase in the use of computers in class. Both confirm that despite the intent of the program to increase computer use in the teaching of Spanish, the program did not have that effect.

VI. Conclusion

The results of this evaluation provide a sobering example of the potential limits of ICT interventions aimed at improving the methods that teachers use in the classroom. Our estimates suggest that this widely implemented national program has no effect on students' academic performance. The primary reason seems to be a failure of the implementation of the program. Despite receiving computers, training, and technical assistance, the teachers in the program simply failed to incorporate the new technology into their classroom teaching.

This example provides an important lesson both for researchers and for policy makers. For policy makers it emphasizes the importance of program implementation and monitoring. In this case, the program simply assumed that once equipped and trained, teachers would voluntarily incorporate the provided technology into their classrooms. Mere training and equipment does not seem to be sufficient.

For researchers, it suggests two important research questions. First, the obvious question is once trained and equipped, how can we get teachers to use the resources that they are given? There are examples of teacher using new techniques and technologies (Duflo, Hanna, Ryan, 2007; He, MacLeod, and Linden, 2008), but why these efforts were successful at changing teacher behavior and other were not remains an open question. Second, while all too often process evaluations are incorrectly interpreted as measures of impact (Duflo, 2004), process evaluations remain an important component of impact evaluations. Particularly in a literature with results that are as mixed as those that estimate the effects of ICT on education, it is important to be able to distinguish between programs that have little effect because of implementation issues and those that fail because of poor pedagogical design.

References

Angrist, J. and Lavy, V. (2002) "New Evidence on Classroom Computers and Pupil Learning". *Economic Journal*, 112, pp. 735-765.

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), pp. 1235-1264.

Barrera-Osorio, F., C. Dominguez, L. Linden, and M. Olivera. 2006. "Computers for Education: Evidence from a Randomized Experiment. Baseline report", processed, Fedesarrollo, Bogota, Colombia.

Bertrand, M., E. Duflow, and S. Mullainathan (2004) "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 pp 249-275.

Computadores para Educar (2008) www.computadoresparaeducar.gov.co. Accessed December 9, 2008.

Cox, M., Abbott, C., Webb, M., Blakely, B., Beauchamp, T. and Rhodes, V. (2003) ICT and Attainment—A Review of the Research Literature, Coventry: Becta (ICT in Schools Research and Evaluation Series)

Duflo, E. (2004). "Scaling Up and Evaluation," *Annual World Bank Conference on Development Economics 2004*. pp 341-369.

Duflo, E., Hanna, R. and Ryan, S. (2007) "Monitoring Works: Getting Teachers to Come to School," NBER Working Paper No. 11880, 2005; BREAD Working Paper No. 103, November 2007.

Glewwe, P., Kremer, M. and Moulin, S. (2007) "Many Children Left Behind? Textbooks and Test Scores in Kenya," NBER Working Papers 13300, National Bureau of Economic Research, Inc.

Harrison, C., C. Comber, T. Fisher, K. Haw, C. Lewin, E. Lunzer, A. McFarlane, D. Mavers, P. Scrimshaw, B. Somekh, and R. Watling, (2002) “ImpCT2: The Impact of Information and Communication Technilogies on Pupil Learning and Attainment,” *mimeo* Department for Education and Skills (DfES)/Becta, London.

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," mimeo Columbia University.

Henao, O. and Ramirez, D. (date not available) “Impacto de una propuesta de capacitación en didáctica de la lecto-escritura con un enfoque socio-constructivista y el uso de TICs, en las concepciones teóricas y metodológicas de los docentes,” mimeo Universidad de Antioquia, Medellin, Colombia.

InfoDev (2005), “Knowledge Maps: ICTs in Education”, November, www.infoDev.org.

Kozma, R. and McGhee, R. (1999). *World Links for Development: Accomplishments and challenges. Monitoring and evaluation annual report: 1998-1999*. Menlo Park, CA: SRI International.

Kozma, R., McGhee, R., Quellmalz, E. and Zalles, D. (2004), “Closing the digital divide: Evaluation of the World Links program”. *International Journal of Educational Development*, 24(4), pp. 361-381.

Kremer, M., Miguel, E. and Thornton, R. (2007). “Incentives to Learn,” mimeo MIT University.

Kulik, J. (2003), "Effects of Using Instructional Technology in Elementary and Secondary Schools: What Controlled Evaluation Studies Say", SRI Inernational, Arlington, VA, May.

Linden, L., Banerjee, A. and Duflo, E. (2003), "Computer-Assisted Learning: Evidence from a Randomized Experiment", Poverty Action Lab Paper No. 5, October.

Linden, L. (2008) "Complement or Substitute? The Effect of Technology on Student Achievement in India," mimeo Columbia University.

Ministerio de Educación Nacional de Colombia (MEN) (2008). http://menweb.mineducacion.gov.co/saber/componentes.php. Accessed December 10, 2008.

Muralidharan, K. and Sundararaman, V. (2008) "Teacher Performance Pay: Experiment Evidence from India," mimeo University of California at San Diego.

National Center for Educational Statistics (NCES) (2001a) "The Nation's Report Card: Mathematics 2000", Washington DC.

National Center for Educational Statistics (NCES) (2001b) "The Nation's Report Card: Science 2000", Washington DC.

Quellmalz, E. and D. Zalles (2000) *World Links for Development: Student assessment Uganda field test*. Menlo Park, CA: SRI International.

Rouse and Krueger (2004) "Putting computerized instruction to the test: a randomized evaluation of a 'scientifically based' reading program" *Journal Economics of Education Review*, 23(4), pp. 323-338.

Wenglinsky, H. (1998) "Does it compute? The relationship between educational technology and student achievement in mathematics", Princeton, NJ: ETS.

Table 1: Distrubtion of Surveyed Schools

	Baseline Survey			Follow-Up Survey		
	Control	Treatment	Total	Control	Treatment	Total
Schools	49	48	97	49	48	97
Students	3889	4327	8216	2403	2798	5201
Students by Gender						
Male	1851	2133	3984	1146	1410	2556
Female	2038	2194	4232	1257	1388	2645
Students by Grade						
Third	1230	1209	2439	764	804	1568
Fourth	375	365	740	242	251	493
Fifth	370	441	811	235	279	514
Sixth	767	833	1600	455	536	991
Seventh	576	689	1265	345	427	772
Eighth	321	469	790	198	276	474
Ninth	250	321	571	164	225	389

Note: This table contains a tabulation of the schools and students that were surveyed in the baseline survey and then both the baseline and follow-up surveys. One hundred schools were originally randomized, but three schools (2 treatment and 1 control) refused to participate in data collection.

Table 2: Average Characteristics of Students Completing Baseline Survey

Characteristics	Treatment Average	Control Average	Difference
Panel A: Test Scores			
Language score	0.06	-0.01	0.07
	(0.09)	(0.08)	(0.12)
Math score	0.04	-0.02	0.06
	(0.08)	(0.08)	(0.11)
Total score	0.07	-0.02	0.08
	(0.10)	(0.09)	(0.13)
Panel B: Demographic Characteristics			
Gender	0.51	0.52	-0.02
	(0.01)	(0.02)	(0.02)
Age	12.05	11.85	0.20
	(0.26)	(0.35)	(0.43)
N. parents in the household	1.55	1.59	-0.04
	(0.02)	(0.02)	(0.03)
N. siblings	3.77	4.03	-0.27
	(0.20)	(0.18)	(0.27)
Receives allowance	0.76	0.72	0.04
	(0.02)	(0.03)	(0.03)
N. friends	17.88	15.52	2.36
	(1.79)	(1.16)	(2.12)
Hours of work	6.50	7.58	-1.08
	(0.37)	(0.76)	(0.84)
Panel C: Academic Variables			
Transport time to school	2.43	2.41	0.02
	(0.05)	(0.07)	(0.08)
Attended school last year	0.97	0.97	0.00
	(0.00)	(0.01)	(0.01)
Repeated a grade in the past	0.35	0.36	-0.01
	(0.02)	(0.02)	(0.02)
N. days absent this year	2.07	1.89	0.18
	(0.10)	(0.13)	(0.16)
Failing subject	0.37	0.35	0.03
	(0.02)	(0.02)	(0.03)
Stays at school after classes	0.09	0.12	-0.032*
	(0.01)	(0.02)	(0.02)
Talks to teacher outside class	0.62	0.70	-0.082**
	(0.02)	(0.03)	(0.03)
N. hours study outside school	1.45	1.33	0.112*
	(0.05)	(0.04)	(0.06)

Note: This table contains the average characteristics of all students completing the baselind survey in grades three through nine. This includes 8,216 students -- 4,327 in the treatment group and 3,889 in the control group. Differences are generated by estimating equation (1) using ordinary least squares with standard errors clustered by school. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 3: Comparison of Attriting Students by Baseline Characteristics

	Attritors			Attritors less Non-Attritors		
Characteristics	**Treatment Average**	**Control Average**	**Difference**	**Treatment Average**	**Control Average**	**Difference**
Percentage of Baseline Students	0.353	0.382	-0.029			
			(0.031)			
Panel A: Test Scores						
Language normalized score	0.012	-0.018	0.03	-0.071	-0.018	-0.053
	(0.064)	(0.082)	(0.104)	(0.057)	(0.044)	(0.071)
Math normalized score	0.049	-0.047	0.095	0.011	-0.047	0.058
	(0.095)	(0.075)	(0.120)	(0.050)	(0.042)	(0.065)
Total normalized score	0.039	-0.039	0.078	-0.042	-0.039	-0.002
	(0.089)	(0.091)	(0.126)	(0.065)	(0.043)	(0.077)
Panel B: Demographic Characteristics						
Gender	0.527	0.526	0.002	0.031	0.002	0.029
	(0.016)	(0.021)	(0.026)	(0.023)	(0.021)	(0.031)
Age	12.527	12.328	0.199	0.738***	0.777***	-0.04
	(0.260)	(0.357)	(0.439)	(0.123)	(0.131)	(0.179)
N. parents in the household	1.458	1.516	-0.058	-0.137***	-0.112***	-0.025
	(0.028)	(0.030)	(0.041)	(0.024)	(0.027)	(0.036)
N. siblings	3.88	4.101	-0.222	0.174*	0.11	0.064
	(0.175)	(0.181)	(0.251)	(0.092)	(0.121)	(0.151)
Receives allowance	0.762	0.702	0.060*	-0.001	-0.027	0.026
	(0.015)	(0.030)	(0.033)	(0.017)	(0.018)	(0.025)
N. friends	17.823	14.533	3.29	-0.083	-1.582**	1.499*
	(1.629)	(1.215)	(2.022)	(0.650)	(0.623)	(0.895)
Hours of work	7.343	8.286	-0.944	1.391**	1.199**	0.192
	(0.544)	(0.912)	(1.056)	(0.572)	(0.594)	(0.820)
Panel C: Academic Variables						
Transport time to school	2.384	2.379	0.004	-0.075**	-0.052*	-0.023
	(0.055)	(0.078)	(0.095)	(0.030)	(0.031)	(0.043)
Attended school last year	0.957	0.954	0.003	-0.020***	-0.025***	0.005
	(0.007)	(0.006)	(0.009)	(0.006)	(0.007)	(0.009)
Repeated a grade in the past	0.422	0.43	-0.008	0.111***	0.107***	0.003
	(0.014)	(0.018)	(0.022)	(0.014)	(0.017)	(0.022)
N. days absent this year	2.196	1.906	0.29	0.226	0.032	0.194
	(0.162)	(0.123)	(0.202)	(0.155)	(0.145)	(0.211)
Failing subject	0.444	0.403	0.041	0.113***	0.094***	0.019
	(0.022)	(0.033)	(0.039)	(0.018)	(0.030)	(0.034)
Stays at school after classes	0.103	0.131	-0.028	0.016	0.011	0.005
	(0.011)	(0.019)	(0.022)	(0.011)	(0.014)	(0.018)
Talks to teacher outside class	0.627	0.702	-0.075**	0.018	0.007	0.011
	(0.022)	(0.024)	(0.032)	(0.014)	(0.018)	(0.023)
N. hours study outside school	1.433	1.302	0.131**	-0.019	-0.05	0.032
	(0.035)	(0.046)	(0.057)	(0.045)	(0.045)	(0.063)

Note: This table describes the characteristics of students attriting at followup. The first three columns contain the averge characteristics of attriting students. Differences are estimated using equation (1). The last three columns compare the relative characteristics of attriting and non-attriting students using equation (3). Standard errors in all models are clustered by school. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 4: Comparison of Non-Attriting Students Using Baseline Characteristics

Characteristics	Treatment Average	Control Average	Difference
Panel A: Test Scores			
Language normalized score	0.083	0	0.083
	(0.104)	(0.083)	(0.133)
Math normalized score	0.037	0	0.037
	(0.081)	(0.081)	(0.114)
Total normalized score	0.08	0	0.08
	(0.111)	(0.097)	(0.147)
Panel B: Demographic Characteristics			
Gender	0.496	0.523	-0.027
	(0.016)	(0.024)	(0.029)
Age	11.79	11.551	0.239
	(0.268)	(0.361)	(0.447)
N. parents in the household	1.595	1.628	-0.033
	(0.019)	(0.023)	(0.030)
N. siblings	3.705	3.991	-0.286
	(0.222)	(0.201)	(0.298)
Receives allowance	0.763	0.729	0.034
	(0.019)	(0.029)	(0.034)
N. friends	17.906	16.115	1.791
	(1.914)	(1.154)	(2.223)
Hours of work	5.951	7.087	-1.136
	(0.398)	(0.729)	(0.826)
Panel C: Academic Variables			
Transport means to school	1.607	1.444	0.163
	(0.150)	(0.066)	(0.163)
Attended school last year	0.976	0.979	-0.003
	(0.004)	(0.005)	(0.007)
Repeated a grade in the past	0.311	0.322	-0.011
	(0.020)	(0.018)	(0.027)
N. days absent this year	1.969	1.873	0.096
	(0.096)	(0.167)	(0.191)
Failing subject	0.331	0.309	0.022
	(0.016)	(0.020)	(0.026)
Stays at school after classes	0.087	0.12	-0.034*
	(0.010)	(0.015)	(0.018)
Talks to teacher outside class	0.61	0.695	-0.085**
	(0.021)	(0.031)	(0.037)
N. hours study outside school	1.452	1.353	0.099
	(0.057)	(0.043)	(0.071)

Note: This table contains a comparison of non-attriting students by baseline characteristics using 5,201 non-attriting students (2,798 tretment and 2,403 control). Difference are generated by estimating equation (1) using ordinary least squares with standard errors clustered by school. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 5: First Stage, Distribution of Treatment by Research Group

Variable	Treatment Average	Control Average	Difference
CPE Reported School Treated	0.958	0.041	0.918***
	(0.029)	(0.029)	(0.041)
Number of Computers at School	13.383	5.102	8.281***
	(1.279)	(0.753)	(1.485)
Percentage of Teachers Trained	0.947	0.082	0.865***
	(0.031)	(0.04)	(0.05)

Note: This table compares characteristics of the 97 schools in our sample. All regressions are estimated using equation (1) at the school level. Significance at the one, five, and ten percent levels are indicated by ***, **, and * respectively.

Table 6: Follow-Up Test Scores

	Percentage Correct				Normalized Score			
Test Sections	Treatment Average	Control Average	Difference	Difference w/ Cntrls	Treatment Average	Control Average	Difference	Difference w/ Cntrls
Spanish Section	0.42	0.402	0.017	0.015	0.099	0	0.099	0.077
	(0.014)	(0.013)	(0.019)	(0.015)	(0.071)	(0.059)	(0.092)	(0.076)
Math Section	0.238	0.23	0.008	0.014	0.014	0.07	0.07	0.088
	(0.018)	(0.011)	(0.021)	(0.019)	(0.019)	(0.098)	(0.110)	(0.109)
Total Score	0.334	0.321	0.013	0.015	0.015	0.111	0.111	0.109
	(0.014)	(0.011)	(0.018)	(0.015)	(0.015)	(0.096)	(0.116)	(0.104)

Note: This table contains a comparison of the treatment and control groups using the tests administered in the follow-up survey. Results are presented first for the percentage of correct answers and then using normalized scores. The first column contains the average scores in the treatment group. The second column contains the average scores in the control group. Column three contains the simple difference using equation (1), and column four contains the difference estimate controling for baseline characteristics using equation (2). All standard errors are clustered by school. Sample includes 5,201 students who completed follow-up survey. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 7: Treatment Estimates by Subject

Subject	Treatment Average	Control Average	Difference	Difference w/ Controls
Panel A: Spanish Subjects				
Encyclopedia	0.08	0.00	0.08	0.08
	(0.04)	(0.04)	(0.05)	(0.05)
Identification	0.11	0.00	0.11	0.13
	(0.06)	(0.04)	(0.07)	(0.10)
Paraphrase	0.07	0.00	0.07	0.07
	(0.06)	(0.05)	(0.08)	(0.06)
Pragmatics	-0.04	0.00	-0.04	-0.06
	(0.05)	(0.04)	(0.06)	(0.06)
Panel B: Math Subjects				
Algebra	-0.04	0.00	-0.04	0.09
	(0.08)	(0.05)	(0.09)	(0.14)
Aritmetics	-0.01	0.00	-0.01	-0.01
	(0.06)	(0.03)	(0.07)	(0.07)
Geometry	0.07	0.00	0.07	0.10
	(0.09)	(0.05)	(0.10)	(0.10)
Statistics	0.11	0.00	0.11	0.12
	(0.08)	(0.04)	(0.09)	(0.09)

Note: This table contains a comparison of the treatment and control groups using the tests administered in the follow-up survey. Scores are normalized relative to the control group follow-up survey distribution. The first column for each subject area contains the average scores in the treatment group. The second column contains the average scores in the control group. Column three contains a simple difference estimates using equation (1) and column four contains the difference estimate controling for baseline characteristics using equation (2). All standard errors are clustered by school. Sample includes 5,201 students who completed follow-up survey. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 8: Difference in Test Scores by Gender at Follow-Up

Test by sections	Treatment Average	Control Average	Difference	Difference w/ Controls
Panel A: Female Students				
Language score	0.146	0.051	0.095	0.069
	(0.077)	(0.066)	(0.101)	(0.087)
Math score	0.088	-0.039	0.127	0.143
	(0.117)	(0.047)	(0.125)	(0.122)
Total score	0.155	0.017	0.138	0.136
	(0.110)	(0.065)	(0.127)	(0.115)
Panel B: Male Students				
Language score	0.055	-0.047	0.102	0.087
	(0.072)	(0.065)	(0.096)	(0.079)
Math score	0.051	0.035	0.016	0.027
	(0.087)	(0.063)	(0.107)	(0.107)
Total score	0.068	-0.015	0.083	0.079
	(0.091)	(0.076)	(0.118)	(0.102)

Note: This table contains the difference between treatment and control groups disaggregated by gender. Panel A presents the results for female students while Panel B presents the results for male students. The first column presents the average score on the specified subject for the treatment group. The second column contains the average score for the control group. The third column presents the simple difference in averages estimated using equation (1) and the final column presents the average difference controling for baseline characteristics using equation (2). All standard errors are clustered by school. Sample includes the 5,201 students who completed the follow-up survey. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 9: Difference in Test Scores by Grade at Follow-Up

	Grade in 2006						
Test by Sections	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Spanish Section	0.139	0.115	0.047	0.092	-0.020	-0.257***	0.302**
	(0.157)	(0.140)	(0.160)	(0.115)	(0.158)	(0.098)	(0.134)
Math Section	0.170	0.225	0.125	-0.042	0.029	-0.186	0.355**
	(0.226)	(0.213)	(0.237)	(0.158)	(0.164)	(0.151)	(0.160)
Total Score	0.189	0.215	0.133	0.032	-0.007	-0.298***	0.426***
	(0.225)	(0.168)	(0.241)	(0.145)	(0.175)	(0.111)	(0.138)

Note: This table reports the estimated difference in normalized test scores between the treatment and control groups using equation (2). Estimates are disaggregated by grade and subject as indicated. Standard errors are clustered by school. The sample size for each grade is given in Table 1, but all 5,201 students who completed the follow-up survey are included in this table. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 10: Difference in Other Academic Outcomes at Follow-Up

Characteristics	Treatment Average	Control Average	Difference
Attended school last year	0.96	0.94	0.02
	(0.007)	(0.014)	(0.016)
Did not attend school last week	0.246	0.37	-0.124*
	(0.024)	(0.067)	(0.070)
N. days she did not attended	1.793	2.064	-0.271
	(0.114)	(0.212)	(0.239)
Likes school	0.977	0.973	0.004
	(0.004)	(0.003)	(0.005)
Likes contents at school	0.987	0.986	0.001
	(0.002)	(0.004)	(0.004)
Grade report	4.023	4.095	-0.072
	(0.047)	(0.037)	(0.060)
Failing subject	0.396	0.365	0.03
	(0.028)	(0.032)	(0.042)
Stays at school after classes	0.064	0.093	-0.028*
	(0.010)	(0.011)	(0.015)
Talks to teacher outside class	0.647	0.657	-0.009
	(0.021)	(0.029)	(0.035)
Hours of work	8.185	9.92	-1.735**
	(0.493)	(0.613)	(0.782)

Note: This table contains the average responses of students in the respective groups. Differences are estimated using equation (1). Standard errors are clustered at the school level. Sample includes all 5,201 students who completed a follow-up survey. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 11: Teachers' Reported Utilizaton of Computers at Follow-Up

	Language teachers			Math teachers		
Variables	**Treatment Average**	**Control Average**	**Difference**	**Treatment Average**	**Control Average**	**Difference**
Did you use computer last week?	0.416	0.170	0.246***	0.438	0.182	0.256***
	(0.075)	(0.041)	(0.085)	(0.074)	(0.043)	(0.085)
How many times last week did you?						
Use a computer in relationship to your classes	0.876	0.388	0.488**	0.932	0.436	0.496**
	(0.188)	(0.110)	(0.217)	(0.173)	(0.116)	(0.207)
Use a computer for group work in class	1.478	1.444	0.033	1.486	1.621	-0.135
	(0.176)	(0.257)	(0.309)	(0.197)	(0.404)	(0.445)
Use a computer for lectures	0.866	0.679	0.187	0.789	0.903	-0.114
	(0.275)	(0.137)	(0.306)	(0.263)	(0.182)	(0.318)
Use a computer for class excercises	1.552	1.036	0.517	1.414	1.226	0.188
	(0.234)	(0.235)	(0.329)	(0.206)	(0.275)	(0.341)
Use a computer in class for Internet research	0.746	0.321	0.425	0.761	0.419	0.341
	(0.326)	(0.124)	(0.347)	(0.305)	(0.198)	(0.362)
Allowed class free use of the computeres	1.299	0.893	0.406	1.324	1.387	-0.063
	(0.320)	(0.245)	(0.400)	(0.303)	(0.408)	(0.505)
Requires use of a computer for homework	2.313	3.000	-0.687	2.569	3.133	-0.564
	(0.302)	(0.507)	(0.584)	(0.323)	(0.504)	(0.594)

Note: This table presents the results of teachers reported use of computers for the indicated activities. Statistics are provided for teachers teaching Spanish and teacher teaching math. Teachers who teach both subject are included in both sets of statistics. Differences were estimated using equation (1) with standard errors clustered at the school level. Sample includes 426 teachers. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 12: Fraction of Schools Reporting Computer Use at Follow-Up by Grade

Grade	Language teachers			Math teachers		
	Treatment Average	Control Average	Difference	Treatment Average	Control Average	Difference
Third	0.45	0.25	0.202**	0.45	0.22	0.222**
	(0.07)	(0.06)	(0.10)	(0.07)	(0.06)	(0.10)
Fourth	0.43	0.27	0.160*	0.45	0.25	0.202**
	(0.07)	(0.06)	(0.10)	(0.07)	(0.06)	(0.10)
Fifth	0.36	0.27	0.10	0.38	0.27	0.12
	(0.07)	(0.06)	(0.10)	(0.07)	(0.06)	(0.10)
Sixth	0.09	0.02	0.07	0.13	0.06	0.07
	(0.04)	(0.02)	(0.05)	(0.05)	(0.04)	(0.06)
Sevent	0.06	0.06	0.00	0.13	0.10	0.03
	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)	(0.07)
Eight	0.04	0.06	-0.02	0.11	0.12	-0.02
	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	(0.07)
Ninth	0.04	0.04	0.00	0.09	0.08	0.00
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.06)
Tenth	0.02	0.02	0.00	0.06	0.08	-0.02
	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.05)
Eleventh	0.02	0.02	0.00	0.06	0.06	0.00
	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.05)

Note: This table contains the fraction of schools, in each research group, that have at least one teacher teaching in the indicated grade that reports using a computer. Differences were estimated using equation (1) at the school level. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.

Table 13: Student Reported Computer Utilization in Previous Week at Follow-Up

	All Students			Students using Computers		
	Treatment Average	Control Average	Difference	Treatment Average	Control Average	Difference
Did you use a computer last week?	0.66	0.41	0.252***			
	(0.04)	(0.05)	(0.06)			
Where did you use a computer last week?						
at school	0.57	0.27	0.300***	0.85	0.64	0.214*
	(0.04)	(0.06)	(0.07)	(0.04)	(0.10)	(0.11)
at home	0.04	0.05	-0.01	0.06	0.11	-0.05
	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)
at an internet café	0.08	0.07	0.00	0.12	0.17	-0.06
	(0.04)	(0.03)	(0.05)	(0.05)	(0.08)	(0.09)
at a friends home	0.02	0.03	-0.01	0.04	0.07	-0.038**
	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
In which subjects did you use a computer?						
Mathematics	0.03	0.02	0.01	0.03	0.02	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Spanish	0.04	0.03	0.02	0.05	0.03	0.01
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
Natural Sciences	0.03	0.03	0.00	0.03	0.04	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Social Sciences	0.02	0.02	0.00	0.02	0.02	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Computer Science	0.51	0.21	0.294***	0.52	0.22	0.296***
	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.07)
Art	0.02	0.02	-0.01	0.02	0.02	-0.01
	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)

Note: This table contains estimates of differences in students' reported use of computers. The first three columns contain estimates for all 5,201 children completing a follow-up survey. The last three columns contain estimates for just those students who report having used a computer in the last week. All estimated differences are generated by estimating equation (1) with standard errors clustered at the school level. Significance at the one, five, and ten percent levels is indicated by ***, **, and * respectively.