WORKING PAPER Nº IDB-WP-01698

Remote Tutoring in Latin America

Facundo Albornoz
Gonzalo Almeyda Torres
María Lombardi
Victoria Oubiña
Pablo Zoido Lobaton

Inter-American Development Bank
Education Division

April 2025

IDB

Remote Tutoring in Latin America

Facundo Albornoz
Gonzalo Almeyda Torres
María Lombardi
Victoria Oubiña
Pablo Zoido Lobaton

Inter-American Development Bank
Education Division

April 2025

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library
Remote tutoring in Latin America / Facundo Albornoz, Gonzalo Almeyda Torres, María Lombardi, Victoria Oubiña, Pablo Zoido Lobaton.
p. cm. — (IDB Working Paper Series ; 1698)
Includes bibliographical references.
1. Tutors and tutoring 2. Educational innovations- 3. Educational equalization-
I. Albornoz-Crespo, Facundo. II. Almeyda, Gonzalo. III. Lombardi, María. IV. Oubiña, Victoria. V. Zoido, Pablo. VI. Inter-American Development Bank. Education Division. VII. Series.
IDB-WP-1698

JEL Codes: I20, I24, O15
Keywords: tutoring, education, achievement, skills, Latin America

http://www.iadb.org

Remote tutoring in Latin America *

Facundo Albornoz † Gonzalo Almeyda Torres ‡ María Lombardi §
Victoria Oubiña ¶ Pablo Zoido Lobaton ‖

April 2025

Abstract

We study the effect of a randomized one-on-one remote phone tutoring program implemented between 2021 and 2023. The intervention reached more than seven thousand students in seven Latin American countries: Argentina, Brazil, El Salvador, Guatemala, Mexico, Paraguay and Peru. The program targeted students with low initial learning levels and focused on foundational numeracy skills using a differentiated instruction approach. We find that assignment to tutoring increased student test scores by 0.2 SD. Tutoring benefited all students, with no differential effects by gender, age, socioeconomic status, or baseline scores. Students who initially reported having difficulty with concentration or memory experienced larger average effects. Finally, we find that students with lower initial performance exhibited larger improvements in more basic mathematical operations, whereas those with better performance at baseline saw larger gains in more complex operations. This underscores the importance of offering differentiated instruction based on students' initial performance.

JEL Codes: I20, I24, O15
Keywords: tutoring, education, achievement, skills, Latin America

*We are grateful for comments from David Evans.
†University of Nottingham, e-mail: facundo.albornoz@nottingham.ac.uk
‡Inter-American Development Bank, e-mail: gonzaloal@iadb.org
§Universidad Torcuato Di Tella, e-mail: mlombardi@utdt.edu
¶Inter-American Development Bank, e-mail: v.oubina@gmail.com
‖Inter-American Development Bank, e-mail: pabloz@iadb.org

1 Introduction

Education systems worldwide have achieved significant progress in expanding access to schooling, with near-universal enrollment at the primary level (Ritchie et al., 2023). Despite these advances, many students who are currently in school do not have the skills and knowledge required at their grade level. By one measure, more than half of sixth grade students in Latin America do not reach the minimum proficiency level in math (UNESCO, 2023). In response to these challenges, governments are increasingly seeking effective tools to enhance learning outcomes and bridge educational gaps.

Tutoring has proven to be one of the most effective strategies to boost learning achievement. Before the pandemic, substantial experimental evidence demonstrated that in-person tutoring had a significant impact on learning outcomes (Fryer, 2017; Duflo et al., 2024). A meta-analysis by Nickow et al. (2024) found a pooled effect size of 0.29 standard deviations (SD). During the COVID-19 school closures, tutoring became even more popular among researchers and policymakers (Kraft et al., 2024). Remote tutoring programs (via phone calls or platforms) launched during the pandemic were successful at improving student achievement (Carlana and La Ferrara, 2021; Angrist et al., 2022); their effectiveness has also been proven during regular school periods once the pandemic ended (Carlana and La Ferrara, 2024; Gortazar et al., 2024).

While existing research from advanced or low/low-middle income countries suggests that online tutoring has robust impacts across diverse settings, evidence remains scarce for contexts that combine widespread educational coverage with substantial disparities in learning outcomes, such as Latin America. Furthermore, important questions persist about the individual traits linked to greater tutoring benefits. For example, do gender or socioeconomic status influence tutoring effectiveness, or is the intervention equally beneficial to all students facing learning deficits, irrespective of personal background? Are certain characteristics especially useful for identifying student groups likely to gain most from these interventions? This paper aims to expand our understanding of remote tutoring by addressing these critical questions.

Starting in 2021, the Inter-American Development Bank (IDB) launched a remote phone tutoring program to address COVID-19 learning disruptions, beginning in El Salvador (Zoido et al., 2024). Several pilots were later implemented in other Latin American countries (Zoido et al., 2023), offering tutoring services to more than 7 thousand students. The program randomly assigned eight one-on-one math tutoring sessions, each lasting approximately 20 minutes and conducted over the phone. The focus of these tutoring sessions was on foundational numeracy skills, and targeted students who exhibited learning deficiencies at baseline. The program used a differentiated instruction approach (Banerjee et al., 2017; Muralidharan et al., 2019; Angrist et al., 2023; Duflo et al., 2024), offering some degree of content personalization, with initial tutoring content assigned based on each student's performance on the diagnostic test. In most countries, students needed to master a level before progressing to the next one, allowing them to continue practicing a topic until they demonstrated a clear understanding of it. This paper summarizes the results of this multi-country randomized study implemented between 2021 and 2023 in seven Latin American countries: Argentina, Brazil, El Salvador, Guatemala, Mexico, Paraguay and Peru.

We find that assignment to the tutoring program increased student test scores by 0.2 SD, with statistically significant effects across all the numerical operations covered by the intervention. We find that the program is equally effective for students of different subgroups, as the effects on overall performance do not statistically differ based on gender, age, socioeconomic status, or baseline scores. We do find larger effects for students who initially reported having difficulty with concentration or memory. Finally, we find that students with a lower initial performance exhibit larger improvements in more basic operations (addition, subtraction and multiplication), whereas students with relatively better performance at baseline have larger gains in more complex operations (division). These findings underscore the importance of having a differentiated instruction approach to tutoring which tailors support to each student's specific needs, allowing them to focus on the specific skills they need to develop.

Our study contributes to the growing body of research on the effectiveness of remote tutoring. Remote tutoring initiatives have been studied in Botswana (Angrist et al., 2022), Kenya, India, Nepal, the Philippines, and Uganda (Angrist et al., 2023), Italy (Carlana and La Ferrara, 2024), and Spain (Gortazar et al., 2024). Additionally, an online (both at school and at home) tutoring program has been studied in the United States (Kraft et al., 2022). Research consistently finds that these programs are effective at increasing learning achievement. We contribute to this literature by examining the effect of online tutoring initiatives in several upper-middle income Latin American countries. Another important contribution of this paper is that we are able to study whether there is a differential effect for students with attention and concentration challenges, a group which is typically overlooked in the literature. Given the growing importance of learning disabilities and difficulties tied to attention and concentration and their link to technology (and mobile phones in particular) we believe this new evidence points to two key policy implications. Firstly, the potential for one-on-one tutoring led by human tutors (even on the phone) to address the specific learning needs of this population; and secondly, the need for further research in this area. Lastly, the paper also offers support for a differentiated instruction pedagogical approach when working one-on-one. In particular, the finding that low performers follow a different learning pathway when compared with those who have a relatively better performance at baseline highlights the need to adjust learning to where the student is and not where she is supposed to be.

The paper is organized as follows. In Section 2 we describe the experimental design, and in Section 3 we describe the sample and estimation strategy. Section 4 describes the results, and Section 5 concludes.

2 Experimental design

The interventions took place between October 2021 and December 2023 in seven countries: Argentina, Brazil, El Salvador, Guatemala, Mexico, Paraguay, and Peru. These interventions were developed in collaboration with the respective ministries of education, and were implemented in a group of schools within a subset of administrative units, so the samples are not nationally representative.[1] Appendix Table A.1 summarizes the main characteristics of the implementation in each country.

The program targeted primary and lower secondary students with academic difficulties. In El Salvador, Guatemala and Mexico, school principals identified students in need of tutoring, whereas in Argentina, Brazil, Paraguay and Peru, students were selected based on their performance in the initial diagnostic test.[2] The program covered grades 4 through 9 in most countries, although there was some variation across locations. In Peru, the intervention focused exclusively on secondary school students (7^{th} and 8^{th} grades), while in Paraguay, it targeted primary school students (4^{th} to 6^{th} grades). In Brazil, the intervention was implemented in one grade of primary school (5^{th}) and one grade of secondary school (9^{th}).

At baseline, students completed a survey and took a math test. The student surveys and assessments took place at schools in El Salvador, Mexico, Guatemala, Paraguay, and Peru, while in Argentina, and Brazil they were conducted over the phone. The treatment was randomized at the individual level in El Salvador, Mexico, Argentina, Guatemala, and Brazil; at the school level in Peru; and at the class level in Paraguay. After the intervention was completed, all students retook the math assessment.

[1]However, for simplicity, we refer to each intervention by the name of the country.

[2]Students needed the consent of their caregivers in order to participate in the program. The format of the initial contact with the families, where caregivers were introduced to the program and asked for their consent and active participation, varied by country. In El Salvador, Mexico, Argentina, and Brazil, this was done by phone. In Guatemala, Paraguay, and Peru, in-person meetings with caregivers were held at schools.

The baseline student survey inquired about some basic sociodemographic characteristics, and included some questions on disabilities adapted from the Washington Group/UNICEF Module on Child Functioning (Loeb et al., 2017). In all countries except Argentina, students were evaluated on their math skills using the assessment developed by Angrist et al. (2022).[3] The test consisted of nine open-ended questions, and students had a fixed time to answer each question. The assessment covered two-digit addition, subtraction, multiplication and division, as well as problems on place value, word addition, fractions, and logic (see Appendix B.1). In Argentina, a multiple choice test covering similar topics was used (see Appendix B.2).[4]

The treatment consisted of an 8-session math tutoring program delivered by phone, focusing on foundational numeracy, including place value, addition, subtraction, multiplication, and division. The content was aligned with each country's national curriculum. The tutoring sessions were one-on-one and lasted approximately 20 minutes each. The starting level for each student was determined using their results in the baseline math assessment. The structure of each session was as follows: the first few minutes were spent reviewing the results of math problems sent via SMS before the session. Then, new problems of similar difficulty were introduced, accompanied by step-by-step explanations of the procedures to solve them. In the final minutes, students were presented with a series of real-life challenges to assess their mastery of the operations. The sessions were accessible via regular phones, so owning a smartphone was not required. Almost 60% of the students assigned to treatment attended at least one tutoring session, and most of these students completed the program (Appendix Figure A.1).[5]

The interventions incorporated differentiated instruction, offering some degree of content personalization, and with initial tutoring content assigned according to each student's performance in the baseline math test. Except for Argentina, students were divided into five levels corresponding to the five topics covered in the program (place value, addition, subtraction, multiplication and division). For example, students who did not correctly respond the question on place value began with this operation, whereas those who answered the place value question correctly but could not respond the question on addition started with addition. Students advanced to the next level only after demonstrating mastery of the current level by successfully solving the corresponding exercises. This method allowed students to continue practicing a topic until they demonstrated a

[3]This assessment was adapted by Angrist et al. (2022) from the ASER test, a commonly used evaluation (Banerjee et al., 2007, 2017). The validity of this evaluation was confirmed in our sample (León et al., 2022).

[4]In Peru, the instrument included additional items that are not included in our analysis to maintain consistency.

[5]There was some variation across locations in treatment take-up, as shown in Appendix Figure A.2.

clear understanding of it. In Argentina, a different approach to differentiated instruction was used. Students were assigned to three predefined "circuits" based on their initial performance on the diagnostic test. These circuits varied in their starting point, and had a predetermined content to cover in each tutorial, regardless of whether students demonstrated mastery of the content of previous sessions.[6]

The tutoring sessions were held weekly over a period of eight weeks in most countries. In Guatemala, the tutoring sessions took place twice a week, with the intervention lasting four weeks. In Peru, there was an additional treatment arm in which girls, in addition to receiving math tutoring, had extra sessions on growth mindset.

Tutors could be of any age or gender, though the majority were women. Priority was given to teaching trainees, university students, and certified teachers, although individuals from other backgrounds (e.g., professionals from various fields) also participated. Tutors were primarily teachers in Argentina, El Salvador, and Paraguay; university students in Brazil, Guatemala, and Peru; and teaching trainees in Mexico.

3 Sample and estimation

3.1 Sample and descriptive statistics

Our sample at baseline consists of 13,691 students from seven countries (Appendix Table A.2). At baseline, 40% of students correctly answered the place value question, 64% correctly answered the addition question, 44% correctly answered the question on subtraction, 43% on multiplication, and only 18% on division (Appendix Figure A.3). There is substantial heterogeneity in baseline performance across locations (Appendix Table A.3). For example, the average share of correct responses at baseline in Peru and Brazil was 62% and 60%, and 30% and 36% in El Salvador and Paraguay, respectively. Note however that the age of students varies across samples, and could thus partly account for these differences (Appendix Figure A.4).[7]

Due to randomization, this sample is balanced at baseline across treatment arms, as shown in column 1 of Appendix Table A.4. We achieved an average response rate of 71% at endline, resulting in a final estimation sample of 9,731 students. There are no overall

[6] For example, students in the basic circuit started out with a lesson on how to read, write and compare numbers, whereas those in the intermediate circuit started out on the following topic, and ended with an additional activity on multiplication with rational numbers.

[7] As mentioned previously, the evaluation of Argentina had some differences compared to the one used in the other countries. There, the items were multiple choice questions and the formulation of the questions was different too.

differences between the attrition rates of our treatment and control groups (0.61 percentage point difference with a p-value of 0.390). However, there are statistically significant differences in attrition rates by treatment status in El Salvador, Mexico, Argentina and Guatemala (Appendix Figure A.5).[8] Importantly, our final sample is balanced in terms of age, gender, learning difficulties, socioeconomic status, and baseline scores (column 2 of Appendix Table A.4).[9] Furthermore, in Section 4 below we show that our results are robust to correcting for attrition using bounds (Lee, 2009).

3.2 Estimation strategy

We pool the data from the seven experiments, and estimate intent-to-treat effects using the following equation:

$$Score_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Baseline\ Score_i + \gamma_{c(i)} + u_i \tag{1}$$

Our dependent variable, $Score_i$, measures the student's share of correct responses at endline, standardized with respect to the control group from the same country. This variable includes responses to mathematical operations of place value, addition, subtraction, multiplication, and division, as well as an addition word problem.[10] $Treatment_i$ is a dummy for whether student i was assigned to the treatment group. We also control for students' baseline score ($Baseline\ Score_i$), as well as country fixed effects $\gamma_{c(i)}$. In some regressions, we control for age, gender, whether the student reported having difficulty with memory/concentration, and socioeconomic status. Standard errors are clustered at the unit of randomization (class in Paraguay, school in Peru, and individual level elsewhere). We complement our intent-to-treat estimates with treatment-on-the-treated estimates, using the share of sessions attended as the treatment variable.

[8]The attrition rate in the treatment group is 4.07 percentage points lower in El Salvador (p-value=0.005), 13.5 percentage points higher in Mexico (p-value=0.000), 16.2 percentage points lower in Argentina (p-value=0.000), and 5.1 percentage points higher in Guatemala (p-value=0.049).

[9]While there is a small and statistically significant difference in baseline scores (the treatment group has 1.3% of a SD lower scores), we can reject the null hypothesis of a statistically significant difference in all controls (p-value=0.192).

[10]Although students were also asked to solve fraction and logic problems, these were excluded from the outcome variable since these topics were unrelated to the content of the tutoring sessions.

4 Results

Table 1 presents our main results. We find that assignment to treatment increases student test scores by 19.7% of a SD (Panel A, column 1), and this effect is statistically significant at the 1% level. Using Lee bounds to adjust for attrition, the intent-to-treat effects range from 17% to 27% of a SD. Importantly, our results are almost unchanged if we control for age, gender, whether the student reported having difficulty with memory/concentration, and socioeconomic status (column 2), or if we re-weight our sample to give each experiment the same weight (column 3). In Panel B, we present the treatment-on-the-treated estimates, with the share of tutoring sessions attended by the student as the main regressor. We find that attending all the tutoring sessions increases endline test scores by 30.9% of a SD.

We find that treated students exhibit statistically significant learning gains in all numerical operations directly covered by the program, as shown in Appendix Table A.5. In particular, being assigned to treatment increases the likelihood of correctly answering the question on positional value by 9.5 percentage points, addition by 5 percentage points, subtraction by 7.8, multiplication by 5, division by 4.2, and the problem on addition by 3.5 percentage points.

Figure 1 reports the results of separate estimations for each country. We find larger effects in Brazil and Paraguay (54.3% and 57.8% of a SD in the ITT), followed by Mexico and El Salvador (21.6% and 12.3% of a SD in the ITT). While we do not find a statistically significant increase in test scores in Argentina, Guatemala, or Peru, we cannot reject that the intent-to-treat effects are equivalent to those in the other countries, or other studies (see for example Angrist et al., 2023).[11]

We also examine heterogeneous effects and present the results in Table 2.[12] We find that the program is equally effective for students of different subgroups, irrespective of initial disadvantages or demographic differences. In particular, there are no differential effects on overall scores by gender, age, socioeconomic status (SES), or baseline scores. This shows that the remote tutoring interventions under study have a broad-reaching impact, potentially due to their one-on-one nature and differentiated instruction approach, which tailor support to each student's specific needs. We do find suggestive evidence of larger effects for students who reported having some difficulty with concentration or memory at baseline (40% of our sample).[13] For this subgroup, the treatment increases endline

[11] We can reject effects larger than 13.7% of a SD in Argentina, 19% of a SD in Guatemala, and 13.8% of a SD in Peru.

[12] The sample size changes slightly across estimations because of missing values in the characteristics for which we test for heterogeneity, most notably SES status and difficulty with concentration/memory.

[13] We consider that a student has difficulty with concentration/memory if he/she reports that has some, a lot of difficulty, or cannot perform this task.

scores by 24.5% of a standard deviation (vs. 16.8% of a standard deviation for students without difficulties), and the difference in effects is statistically significant at the 10% level (p-value=0.055). Note that in the control group, students who reported having difficulties scored 11.5% of a standard deviation lower than those who did not, so the treatment bridges the gap between these two groups by more than half.

Finally, we analyze whether the impact on learning varies across different types of operations based on students' initial performance. We present our results in Table 3. Although students above and below the median baseline score significantly improve their performance across all types of numerical operations, we find differential effects. Specifically, students with lower overall baseline scores achieve larger gains in addition, subtraction, and multiplication, while those with higher initial performance show greater improvements in division. This finding underscores the importance of having a differentiated instruction approach to tutoring, thus allowing students to focus on the specific skills they need to develop. One potential improvement to the program would be extending its duration for students who start with a lower performance, providing them with additional support to fully catch up by the end.

5 Conclusions

Remote tutoring using low-technology devices results in sizable learning improvements, at least in math and within the age range covered by the interventions studied in this paper. The effect is robust to different contexts, minor variations in implementation, and, importantly, many relevant individual characteristics. Remote tutoring has effects on almost every type of student, with the effect even stronger for students reporting difficulties. Finally, students with lower initial performance exhibit larger learning gains in basic mathematical operations, whereas students who are more capable at baseline improve their skills in more advanced topics.

We also report that the effect of remote tutoring appears to be not significant in some countries when considered independently. The results are noisier in these cases, and we cannot reject the hypothesis that the effects are similar to those observed in other countries. However, it could also be that differences in implementation matter (Angrist and Meager, 2023). For example, the case of Argentina was different because the program had a lower degree of content personalization. Beyond speculative explanations, a definitive understanding of the statistical differences in the effect of remote tutoring across implementations remains an open question that requires further investigation.

A recent concern with remote tutoring, and many other educational interventions, is the challenge of scaling-up these programs effectively (Banerjee et al., 2017; Duflo et al., 2024). The recent expanded meta-analysis by Kraft et al. (2024), which examines 265 randomized controlled trials evaluating tutoring interventions, highlights that while tutoring programs have a meaningful impact on student achievement, the effect sizes tend to diminish as program scale increases. While we do not find that effectiveness diminishes with the size of the intervention, the scalability of the model remains to be rigorously tested.

References

Angrist, Noam, Micheal Ainomugisha, Sai Pramod Bathena et al. (2023) "Building resilient education systems: Evidence from large-scale randomized trials in five countries," NBER Working Paper 31208.

Angrist, Noam, Peter Bergman, and Moitshepi Matsheng (2022) "Experimental evidence on learning using low-tech when school is out," *Nature human behaviour*, 6 (7), 941–950.

Angrist, Noam and Rachael Meager (2023) "Implementation Matters: Generalizing Treatment Effects in Education," EdWorkingPaper No. 23-802.

Banerjee, Abhijit, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton (2017) "From proof of concept to scalable policies: Challenges and solutions, with an application," *Journal of Economic Perspectives*, 31 (4), 73–102.

Banerjee, Abhijit V, Shawn Cole, Esther Duflo, and Leigh Linden (2007) "Remedying education: Evidence from two randomized experiments in India," *Quarterly Journal of Economics*, 122 (3), 1235–1264.

Carlana, Michela and Eliana La Ferrara (2021) "Apart but connected: Online tutoring and student outcomes during the COVID-19 pandemic," CEPR Discussion Paper No. DP15761.

——— (2024) "Apart but Connected: Online Tutoring, Cognitive Outcomes, and Soft Skills," Working Paper 32272, National Bureau of Economic Research.

Duflo, Annie, Jessica Kiessel, and Adrienne M Lucas (2024) "Experimental Evidence on Four Policies to Increase Learning at Scale," *The Economic Journal*, 134 (661), 1985–2008.

Fryer, Roland G (2017) "The production of human capital in developed countries: Evidence from 196 randomized field experiments," in *Handbook of economic field experiments*, 2, 95–322: Elsevier.

Gortazar, Lucas, Claudia Hupkau, and Antonio Roldán-Monés (2024) "Online tutoring works: Experimental evidence from a program with vulnerable children," *Journal of Public Economics*, 232, 105082.

Kraft, Matthew A., John A. List, Jeffrey A. Livingston, and Sally Sadoff (2022) "Online Tutoring by College Volunteers: Experimental Evidence from a Pilot Program," *AEA Papers and Proceedings*, 112, 614–18.

Kraft, Matthew A, Beth E Schueler, and Grace Falken (2024) "What Impacts Should We Expect from Tutoring at Scale? Exploring Meta-Analytic Generalizability," EdWorking Paper No. 24-1031.

Lee, David S (2009) "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76 (3), 1071–1102.

León, Juan, Felipe Hevia, Samana Vergara-Lope, Tamara Vinacur, and Pablo Zoido (2022) "Validación del instrumento de medición utilizado en el proyecto «Tutorías remotas» para acelerar aprendizajes," Technical Note N° IDB-TN-02612.

Loeb, Mitchell, Claudia Cappa, Roberta Crialesi, and Elena De Palma (2017) "Measuring child functioning: the Unicef/Washington group module," *Salud Pública de México*, 59 (4), 485–487.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J Ganimian (2019) "Disrupting education? Experimental evidence on technology-aided instruction in India," *American Economic Review*, 109 (4), 1426–1460.

Nickow, Andre, Philip Oreopoulos, and Vincent Quan (2024) "The promise of tutoring for PreK–12 learning: A systematic review and meta-analysis of the experimental evidence," *American Educational Research Journal*, 61 (1), 74–107.

Ritchie, Hannah, Veronika Samborska, Natasha Ahuja, Esteban Ortiz-Ospina, and Max Roser (2023) "Global Education," *Our World in Data*, https://ourworldindata.org/global-education.

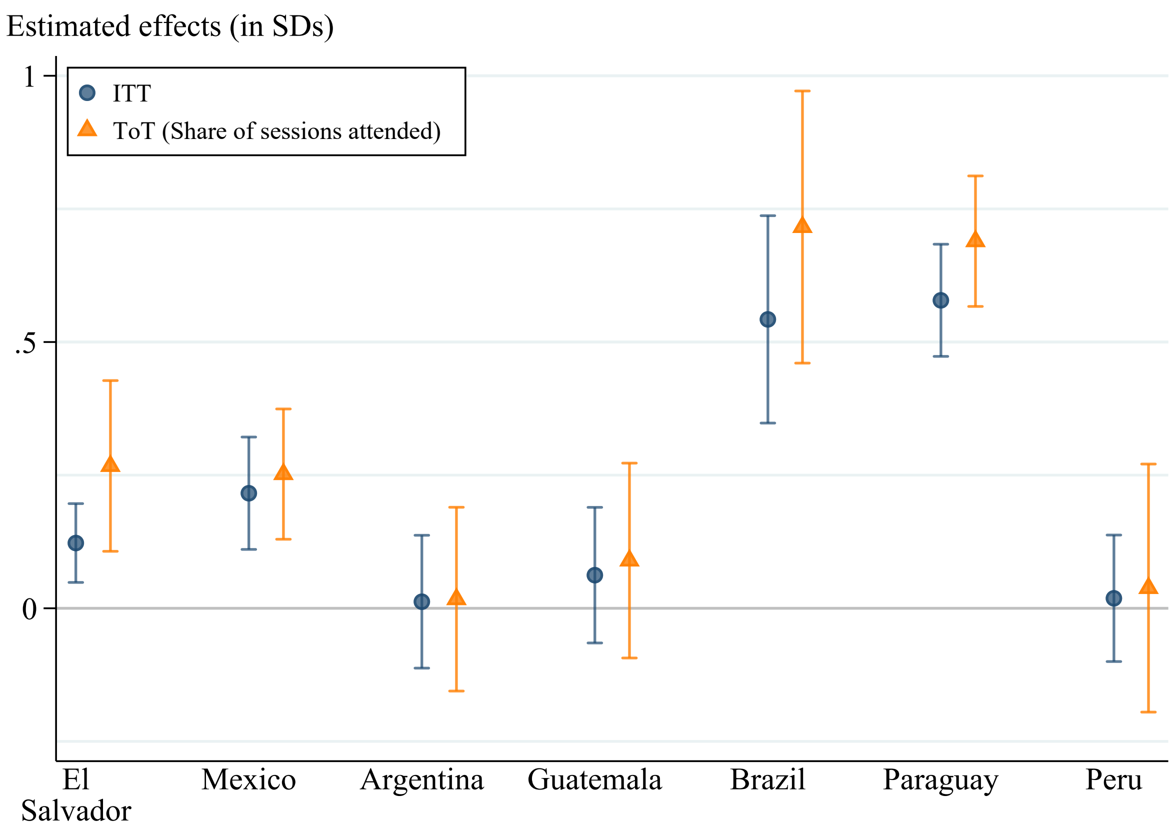
UNESCO (2023) "Fundamental Learnings in Latin America and the Caribbean: Regional Comparative and Explanatory Study (ERCE 2019),"Technical report, United Nations Educational, Scientific and Cultural Organization, https://unesdoc.unesco.org/ark:/48223/pf0000380257_eng/PDF/380257eng.pdf.multi.

Zoido, Pablo, Facundo Albornoz, Andrea Prieto, Gonzalo Almeyda, Jenny Carolina Hernández Cardozo, Victoria Oubiña, and Milton Calderón (2023) "Multiplicar aprendizajes: tutorías a distancia para potenciar la escuela," IADB: Inter-American Development Bank.

Zoido, Pablo, Iván Flores-Ceceña, Miguel Székely, Felipe J. Hevia, and Eleno Castro (2024) "Remote tutoring with low-tech means to accelerate learning: Evidence for El Salvador," *Economics of Education Review*, 98, 102506.

Figures and Tables

Figure 1: Estimated effects by location



Notes: This figure shows the estimates for the intent-to-treat and treatment-on-the-treated effects in each location, along with the 95% confidence intervals.

Table 1: Main estimates

	(1)	(2)	(3)
Panel A: Intent to treat			
Treatment	0.197***	0.197***	0.183***
	(0.024)	(0.024)	(0.026)
Lee bounds (lower)	0.166***		
Lee bounds (upper)	0.274***		
Observations	9,731	9,425	9,731
R^2	0.160	0.159	0.149
Panel B: Treatment on the treated			
Share of tutoring sessions attended	0.309***	0.308***	0.309***
	(0.037)	(0.036)	(0.043)
Observations	9,731	9,425	9,731
R^2	0.171	0.171	0.161
Controls		✓	
Equal weight to each location			✓

Notes: This table shows the results of regressions where the dependent variable is the student's endline score (share of correct responses), standardized with respect to the control group. In Panel A, we present the intent-to-treat estimates, where the main regressor is a dummy for whether the student was assigned to the treatment group. We control for the student's baseline score, and country fixed effects. In column 1, we pool observations, such that each experiment is weighted by the size of the participating population. In column 2, we also control for the student's gender, age, whether he/she reported having difficulty with concentration or memory at baseline, and a dummy for whether the student is below the median SES in their country's sample. In column 3, we re-weight our sample to give each experiment the same weight. Panel B presents the treatment-on-the treated estimates, with the share of tutoring sessions attended by the student as the main regressor. Standard errors clustered at the unit of randomization level in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2: Intent-to-treat estimates – Heterogeneous effects

	Characteristic for heterogeneity				
	Male	Below median age	Below median SES	Below median baseline score	Difficulty with concentration/memory
Treatment	0.207***	0.187***	0.182***	0.167***	0.168***
	(0.030)	(0.027)	(0.031)	(0.032)	(0.028)
Treatment × Characteristic	-0.020	0.029	0.020	0.042	0.077*
	(0.037)	(0.043)	(0.038)	(0.042)	(0.040)
Characteristic	-0.008	-0.097***	-0.070**	-0.641***	-0.115***
	(0.027)	(0.031)	(0.028)	(0.031)	(0.030)
Observations	9,712	9,703	9,567	9,731	9,561
R^2	0.160	0.162	0.158	0.114	0.159
P-value (treatment+interaction=0)	0.000	0.000	0.000	0.000	0.000

Notes: This table shows the results of regressions where the dependent variable is the student's endline score (share of correct responses), standardized with respect to the control group. Our regressors are a dummy for whether the student was assigned to the treatment group, the characteristic specified in the column header, and the interaction between the two. We include country fixed effects in all regressions. We control for the student's baseline score in all regressions except for the one in column 4. In columns 2-4, we split our sample according to age, SES, and baseline score within each country. Standard errors clustered at the unit of randomization level in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Intent-to-treat effects by numerical operation – heterogeneous effects by baseline performance

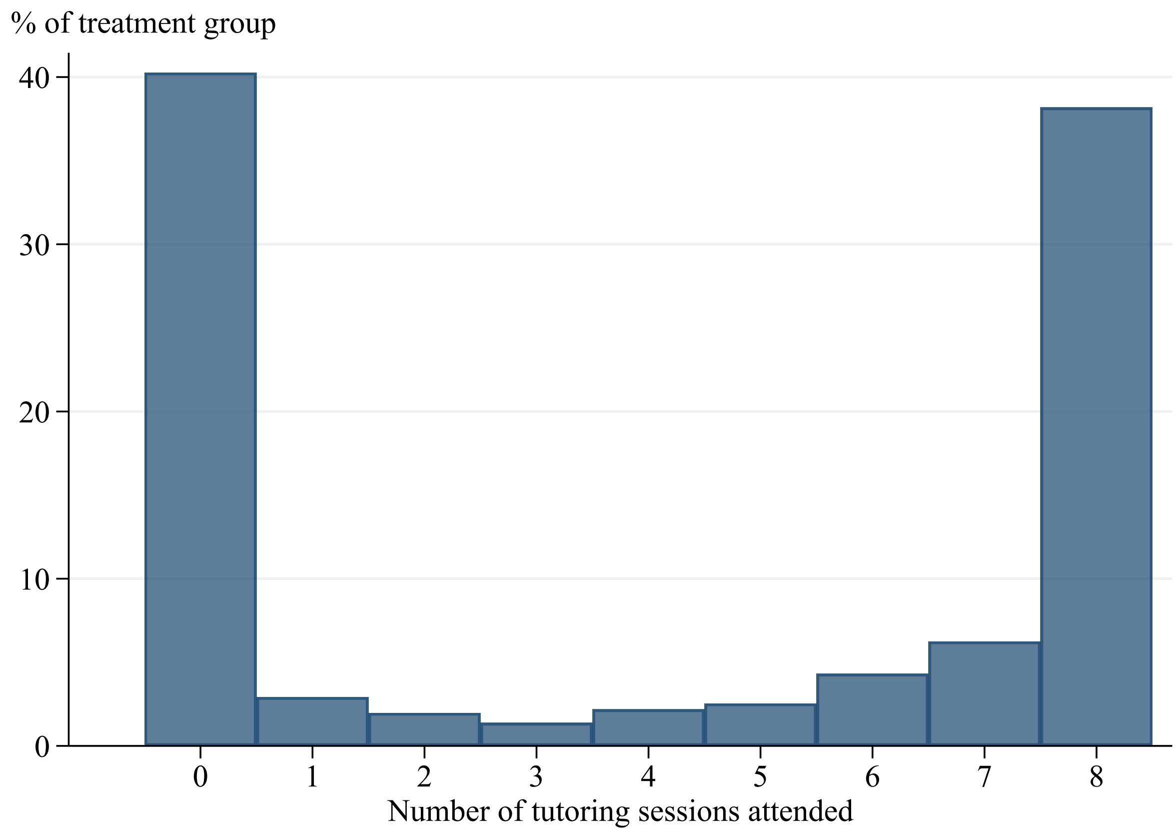
	Positional value	Addition	Subtraction	Multiplication	Division	Addition problem
Treatment	0.085***	0.028***	0.052***	0.018	0.069***	0.032***
	(0.017)	(0.011)	(0.015)	(0.014)	(0.018)	(0.012)
Treatment × below median baseline score	0.017	0.037**	0.045**	0.053***	-0.046**	0.004
	(0.022)	(0.016)	(0.020)	(0.019)	(0.020)	(0.018)
Observations	9,731	9,731	9,731	9,731	9,731	9,731
R^2	0.084	0.084	0.089	0.146	0.106	0.070
Control group mean	0.439	0.782	0.566	0.571	0.293	0.716
P-value (treatment+interaction=0)	0.000	0.000	0.000	0.000	0.050	0.005

Notes: This table shows the results of regressions where the dependent variable is a dummy for whether the student had the correct answer at endline in the question regarding the numerical operation stated in the column header. The main regressors are dummies for whether the student was assigned to the treatment group, and an interaction between that variable and a dummy for whether the student's overall score at baseline was below the median of his/her country. We control for the student's baseline score (overall, and in each numerical operation), and country fixed effects. Standard errors clustered at the unit of randomization level in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

ONLINE APPENDIX

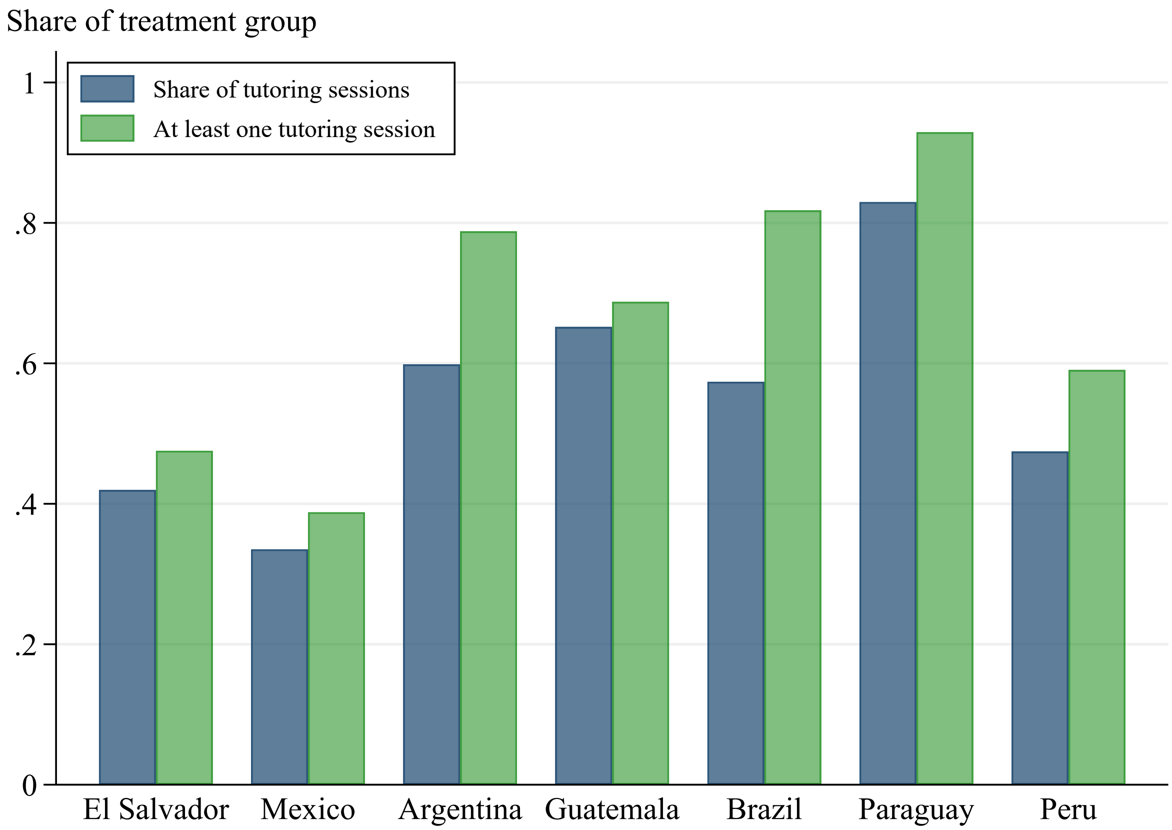
Appendix A Appendix Figures and Tables

Figure A.1: Number of tutoring sessions attended by students in the treatment group



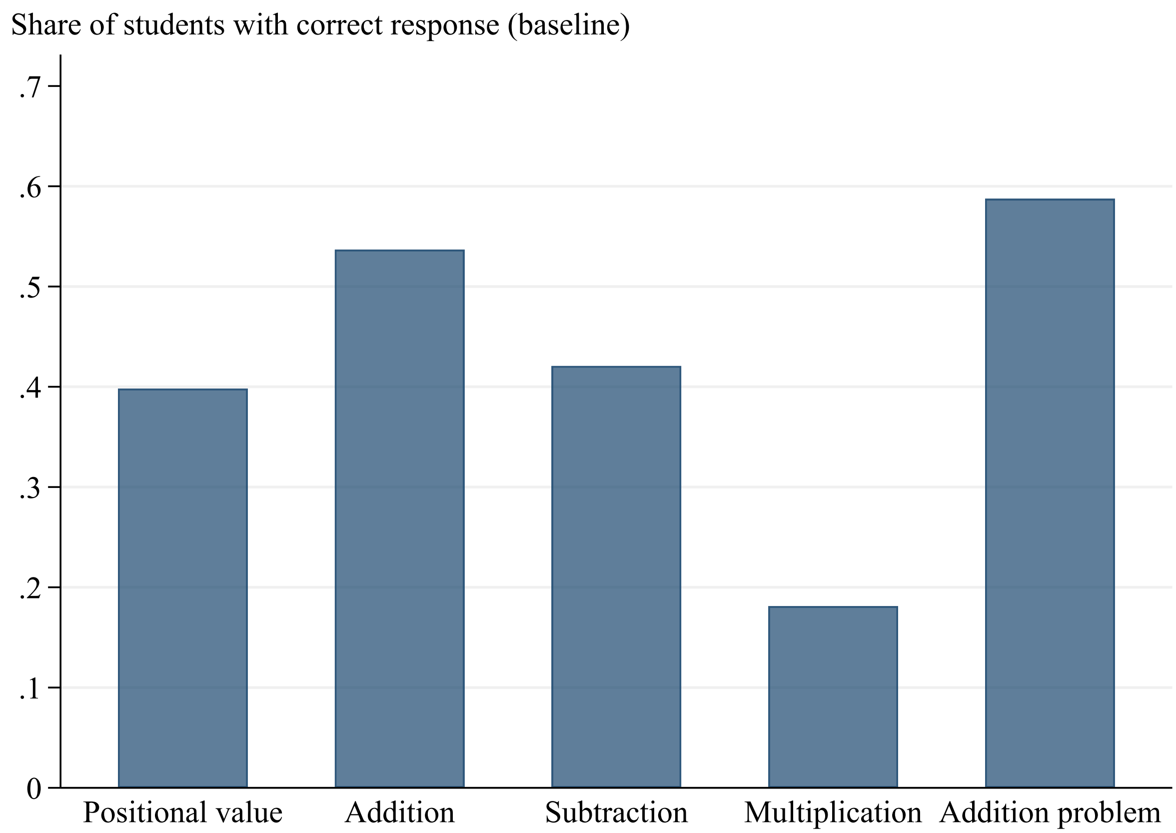
Notes: This figure shows the distribution of the number of tutoring sessions attended by students assigned to the treatment group.

Figure A.2: Treatment take-up by location



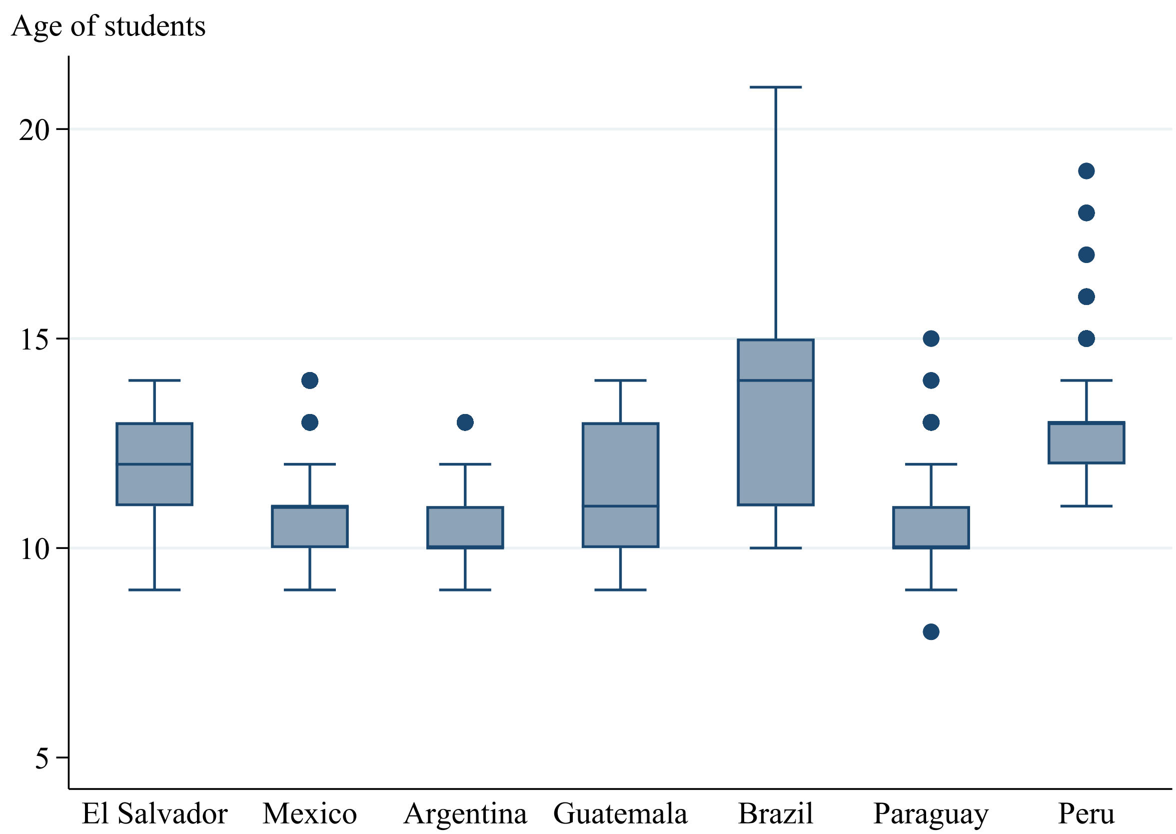
Notes: This figure shows the average take-up of tutoring sessions in the treatment group in each location. The blue bars show the share of tutoring sessions taken by students in the treatment group, and the green bars show the share of students from the treatment group that took at least one tutoring session.

Figure A.3: Learning at baseline by numerical operation



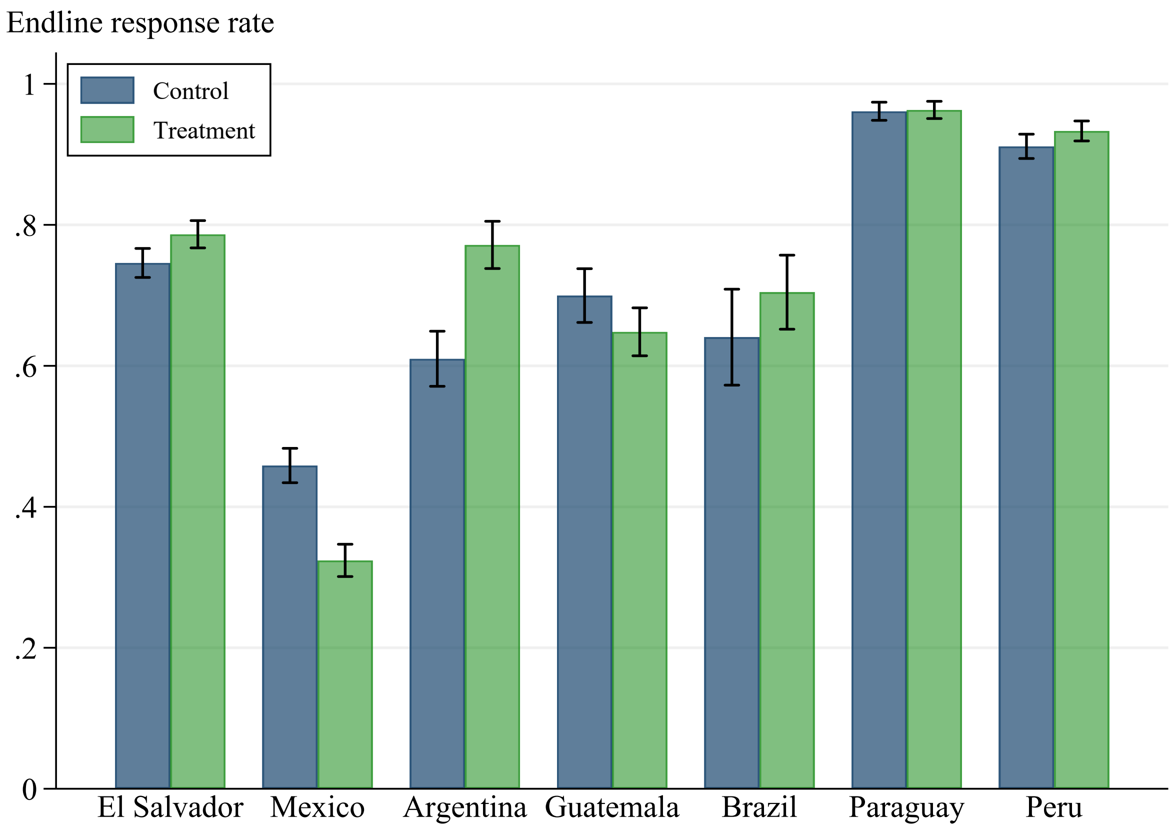
Notes: This figure shows the share of students that had the correct response in the baseline evaluation for each of the numerical skills that were evaluated.

Figure A.4: Age distribution by location



Notes: This figure shows the distribution of student age in each country's sample.

Figure A.5: Endline response rate by location



Notes: This figure shows the average response rates to the endline survey for each group (treatment and control) in each location, with the corresponding 95% confidence interval.

Table A.1: Description of each intervention

	El Salvador	Mexico	Argentina	Guatemala	Brazil	Paraguay	Peru
Number of students	3,440	3,210	1,204	1,315	483	1,793	2,246
Number of schools	68	88	NA	31	39	58	60
Student grades	4th-9th (primary and upper secondary)	4th-9th (primary and lower secondary)	5th-8th (primary and lower secondary)	4th-9th (primary and lower secondary)	5th (primary) and 9th (secondary)	4th-6th (primary)	7th-8th (secondary)
Unit of randomization	Individual	Individual	Individual	Individual	Individual	Class	School
Sensitization format	Phone	Phone	Phone	At school	Phone	At school	At school
Teaching at the right level?	✓	✓	×	✓	✓	✓	✓
Number of tutoring sessions per week	1	1	1	2	1	1	1
Survey format	At school	At school	Phone	At school	Phone	At school	At school
Dates	Oct 2021-Jun 2022	Jan-Jun 2022	2022	Apr 2022-Feb 2023	Aug-Dec 2023	Jun-Dec 2023	May-Dec 2023
Administrative units	3 departments	4 municipalities	1 city	4 departments	1 state	3 departments and the capital city	1 city
Location	Cabanas, Chalatenango, and Morazán	Celaya, Irapuato, León and Silao.	Buenos Aires city	Jalapa, Jutiapa, Santa Rosa and Escuintla	Pará	Asunción, Paraguarí, Central and Cordillera	Ayacucho

Notes: This table summarizes the key features of the intervention across locations.

Table A.2: Sample size at baseline and endline in each location

	Baseline	Endline
El Salvador	3,440	2,636
Mexico	3,210	1,256
Argentina	1,204	832
Guatemala	1,315	881
Brazil	483	328
Paraguay	1,793	1,725
Peru	2,246	2,073
Total	13,691	9,731

Notes: This table shows the number of observations in each location. Column 1 displays the total number of students (at baseline), and column 2 shows the number of students that responded the endline survey.

Table A.3: Learning at baseline by numerical operation in each location

	El Salvador	Mexico	Argentina	Guatemala	Brazil	Paraguay	Peru
Share correct responses	0.301	0.492	0.430	0.399	0.604	0.360	0.622
Positional value	0.311	0.403	0.601	0.250	0.563	0.341	0.513
Addition	0.460	0.706	0.554	0.660	0.760	0.539	0.902
Subtraction	0.273	0.466	0.525	0.445	0.619	0.315	0.642
Multiplication	0.208	0.504	0.331	0.319	0.629	0.335	0.760
Division	0.104	0.236	0.498	0.084	0.369	0.114	0.122
Addition problem	0.450	0.635	0.525	0.637	0.685	0.516	0.771
Observations	3,440	3,210	1,204	1,315	483	1,793	2,245

Notes: This table shows for each location the average share of correct responses and the share of students that had the correct response in the baseline evaluation for each of the numerical skills that were evaluated. Except for Argentina, all evaluations have six items (one for each of the numerical operations). The evaluation in Argentina has nine items, one in addition, subtraction, positional value and addition problem, two on multiplication, and three on division.

Table A.4: Balance tests

	Full sample	Responded endline
Male	0.010	0.007
	(0.009)	(0.011)
Age	-0.003	-0.003
	(0.004)	(0.005)
Difficulty with concentration/memory	-0.004	-0.003
	(0.010)	(0.012)
Low SES	-0.008	-0.016
	(0.011)	(0.014)
Baseline score	-0.004	-0.013**
	(0.005)	(0.007)
Observations	13,312	9,425
R^2	0.003	0.010
P-value (Joint significance)	0.651	0.192

Notes: This table shows the results of regressions where the dependent variable is a dummy for whether the student was assigned to treatment, and the main regressors indicate the student's gender, age, whether he/she reported having difficulty with concentration or memory at baseline, a dummy for whether the student is below the median SES in their country's sample, and the student's baseline score (standardized with respect to the control group). All regressions control for location fixed effects. In column 1, the sample is composed of all students that took the baseline survey, whereas column 2 restricts the sample to students that also responded to the endline survey. Standard errors clustered at the unit of randomization level in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.5: Intent-to-treat estimates by numerical operation

	Positional value	Addition	Subtraction	Multiplication	Division	Addition problem
Treatment	0.095***	0.050***	0.078***	0.050***	0.042***	0.035***
	(0.012)	(0.009)	(0.011)	(0.010)	(0.011)	(0.009)
Observations	9,731	9,731	9,731	9,731	9,731	9,731
R^2	0.084	0.083	0.088	0.145	0.105	0.070
Control group mean	0.439	0.782	0.566	0.571	0.293	0.716

Notes: This table shows the results of regressions where the dependent variable is a dummy for whether the student had the correct answer at endline in the question regarding the numerical operation stated in the column header. The main regressor is a dummy for whether the student was assigned to the treatment group. We control for the student's baseline score (overall, and in each numerical operation), and country fixed effects. Standard errors clustered at the unit of randomization level in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix B Survey tools

Appendix B.1 Assessment used in Brazil, El Salvador, Guatemala, Mexico, Paraguay and Peru

Instructions for the Survey Administrator

- Ask the students to solve the math problems independently on a draft paper, including their answers.
- Explain to the students that this is not a test, and it is perfectly fine if they do not get the answers correct.
- Students should take no longer than **30 seconds** for place value questions or **2 minutes** for regular arithmetic operations.
- If it seems like someone is helping, kindly request them to refrain. If assistance continues and/or the student takes more than the allotted time (**30 seconds** for place value or **2 minutes** for arithmetic operations), mark the response as incorrect.
- Upon completion, ask the student and/or caregiver to read aloud each of their answers.
- Record all correct answers for each student and note the highest level of operation they can successfully perform.

Procedure

- All students begin by answering the **PLACE VALUE** question.
- Regardless of whether they answer correctly, students proceed to the **BASIC LEARNING MODULE**.

Assessment

1) Juan has 47 apples and organizes them by place value. How many **TENS** does he have?

  - **Correct answer**: Responds **4 tens** in less than 30 seconds.

- **Incorrect answer**: Does not respond, gives a different number, or takes more than 30 seconds.

2) What is 52 plus 39?

- **Correct answer**: Responds **91** in less than 2 minutes.
- **Incorrect answer**: Does not respond, gives a different number, or takes more than 2 minutes.

3) What is 42 minus 29?

- **Correct answer**: Responds **13** in less than 2 minutes.
- **Incorrect answer**: Does not respond, gives a different number, or takes more than 2 minutes.

4) What is 28 multiplied by 3?

- **Correct answer**: Responds **84** in less than 2 minutes.
- **Incorrect answer**: Does not respond, gives a different number, or takes more than 2 minutes.

5) What is 65 divided by 8?

- **Correct answer**: Responds **8.1** or **8.12** in less than 2 minutes.
- **Incorrect answer**: Does not respond, gives a different number, or takes more than 2 minutes.

6) A man drives 24 km. Then he drives 17 km. How many kilometers has he traveled in total?

- **Correct answer**: Responds **41** in less than 2 minutes.
- **Incorrect answer**: Does not respond, gives a different number, or takes more than 2 minutes.

7) The day before the day after tomorrow is Saturday. What day is today?

- **Correct answer**: Responds **"Friday"** in less than 2 minutes.
- **Incorrect answer**: Does not respond, gives another day of the week, or takes more than 2 minutes.

The interviewer says: “Do you want to try answering another logic question?”

- **If yes, proceed to Question 9**
- **If no, proceed to Question 10**

8) The day after yesterday is Tuesday. What day is today?

- **Correct answer**: Responds **"Tuesday"** in less than 2 minutes.
- **Incorrect answer**: Does not respond, gives another day of the week, or takes more than 2 minutes.

9) What is $\frac{3}{8} + \frac{4}{8}$?

- **Correct answer**: Responds $\frac{7}{8}$ in less than 2 minutes.
- **Incorrect answer**: Does not respond, gives another number, or takes more than 2 minutes.

Appendix B.2 Assessment used in Argentina

1) Which of the following options allows me to form $1,230 using bills from our country?

(i) 1 of $1,000, 2 of $100, and 30 of $10 ✓
(ii) 1 of $1,000 and 23 of $10
(iii) 1 of $1,000, 200 of $100, and 3 of $10
(iv) 1 of $1,000 and 23 of $100

2) What is the range of the result for: $3,550 + 6,740$?

(i) Less than 9,000
(ii) Between 9,000 and 10,000
(iii) Between 10,000 and 11,000 ✓
(iv) More than 11,000

3) Knowing that $2,800 + 2,300 = 5,100$, what is $5,100 - 2,799$?

(i) 2,299

(ii) 2,301 ✓

(iii) 2,401

(iv) 2,501

4) On National Route No. 7, the distance between the city of Mendoza and Buenos Aires is approximately 1,050 km. If Pedro has already traveled 490 km, how many kilometers does he still need to travel?

(i) 460 km

(ii) 560 km ✓

(iii) 640 km

(iv) 660 km

5) Solve 7×99. Who is doing it correctly?

(i) Daniela, who calculates 7×100 and subtracts 7. ✓

(ii) Lucrecia, who calculates 7×100 and subtracts 1.

(iii) Vito, who calculates 7×90 and adds 9.

(iv) Augusto, who calculates 7×90 and adds 7.

6) If $28 \times 130 = 3,640$, how can you calculate 28×13? Which of the following options is correct?

(i) Subtract 10 from the result of 28×130.

(ii) Subtract 100 from the result of 28×130.

(iii) Divide the result of 28×130 by 10. ✓

(iv) Multiply the result of 28×130 by 10.

7) In a game, I need to equally distribute a deck of 40 cards among 7 children. How many cards does each one get?

(i) Less than 5 cards each, with no leftover cards.

(ii) 5 cards each, with 2 leftover cards.

(iii) 5 cards each, with 5 leftover cards. ✓

(iv) 6 cards each, with no leftover cards.

8) Ana has \$300 to pay for her commutes to work. She spends \$27 per day on transportation to and from work. How many days can she afford to commute with that amount?

(i) 5

(ii) 6

(iii) 11 ✓

(iv) 12

9) Joaquín wants to divide 11 chocolates equally among 5 friends so that everyone receives the same amount and none is left over. How much chocolate will he give to each friend?

(i) 2 chocolates

(ii) $2\frac{1}{5}$ chocolates ✓

(iii) 3 chocolates

(iv) $5\frac{1}{2}$ chocolates