

TC Document

I. Basic information for TC

• Country/Region:	Ecuador - Andean Group
• TC Name:	Closing Gaps: Teaching and Learning Math in the Early Grades
• TC Number:	EC-T1373
• Team Leader/Members:	Yyannú Cruz Aguayo (team leader); Norbert Schady (co-team leader); Sara Schodt, Vania Pizano, Mingi Kang (SCL/SCL); and Mónica Lugo (LEG/SGO)
• Taxonomy:	Client Support
• Date of TC Abstract:	08 May 2017
• Beneficiary:	Ecuador
• Executing Agency:	Inter-American Development Bank through the Social Sector (SCL/SCL)
• Donors providing funding:	Korea Poverty Reduction Fund
• IDB Funding Requested:	\$1,000,000.00
• Local counterpart funding, if any:	\$0.00
• Disbursement period:	36 months
• Required start date:	June 2017
• Types of consultants:	Firms and individual consultants
• Prepared by Unit:	SCL/SCL
• Unit of Disbursement Responsibility:	SCL/SCL
• Included in Country Strategy:	Yes
• TC included in CPD:	No
• Strategic Alignment:	Social inclusion and equality

II. Objectives and Justification of the TC

- 2.1 **Motivation: Why does early math achievement matter?** While Latin America and the Caribbean have made impressive gains towards universal coverage for preprimary and primary school children, the issue of student achievement - learning - is now at the forefront of the education agenda, as students from the region seem to learn little and routinely perform poorly on international tests compared to other countries with similar income levels. The latest PISA results (2015) show that on average, 63% of the LAC students did not achieve basic proficiency in math, and 46% did not achieve this level in reading. The poor quality of teaching might be a key factor in explaining these results, as teachers are the most important determinant of child development and learning within schools.
- 2.2 Latin American countries do very poorly on international tests of math achievement. The average student in South Korea scored 554 points on the math component of the 2012 PISA. In contrast, the average student in Chile, the highest-scoring country in Latin America, scored 423 points, while the average student in Peru scored only 368 points (OECD, PISA 2012). Put differently, the average student in Peru has a score that is almost *two* standard deviations below the score of a student in South Korea—a very large difference, by any measure¹.
- 2.3 The very low scores of Latin American students on math tests is a pressing concern for many reasons. Mathematical knowledge at early ages has been shown to strongly predict later math achievement in a number of longitudinal studies (Duncan et al. 2007; Siegler 2009). Early math achievement also predicts career choice, employment, and wages in adulthood. Duncan et al. (2011) show that having persistently low mathematical skills at ages six, eight, and ten years has a much larger effect on the probability of college attendance than having persistently low reading skills in the United States. Math skills in elementary school are also better predictors of

¹ For results in Ecuador's math achievement see [Annex V](#).

college attendance than newer “hot topic” areas of study, such as socio-emotional skills and the incidence of various behavior problems.

- 2.4 The low math scores of Latin American students also have implications for gender equality, and for the ability of poor children to break out of the cycle of poverty in adulthood. In many Latin American countries, the math scores of poor children are substantially below those of their better-off counterparts, and girls lag behind boys (Berlinski and Schady 2015). The low level of math achievement of poor children at early ages is one way in which poverty is perpetuated from one generation to the next, while the boy-girl gap in early math test scores will limit the career advancement of women.
- 2.5 Finally, low levels of math achievement have serious implications for economic growth for a country as a whole. Relative to other developing countries, Latin America has a shortage of engineers and other professionals in the hard sciences (De Ferranti et al. 2003). Latin American countries severely lag in Research and Development (R&D), and this, in turn, negatively impacts Total Factor Productivity (TFP) growth. Perhaps not surprisingly, Latin American countries have lost ground relative to the US in TFP since 1960, while countries in East Asia like South Korea have gradually reduced that same gap. Hanushek and Woessman (2012) argue that the difference in growth rates between Latin America and East Asia *can be fully explained* by the difference in test scores between the two regions.²
- 2.6 **Teachers and early math achievement in Ecuador.** Early math achievement is critical, but what determines how much math young children learn? Unlike language, math learning happens almost exclusively in the classroom, and teachers are a critical input into the production of learning. However, estimating the causal effect of teachers on math learning is difficult for a variety of reasons. First, because children are generally not assigned to teachers at random, estimates of the “effects” of teachers on child learning can be confounded by unobserved characteristics of students, teachers, or both. Second, simply knowing that some teachers are better than others has only limited policy relevance, as it is silent about the specific characteristics or behaviors that make some teachers more effective than others.
- 2.7 This TC seeks to extend the work of a groundbreaking project in Ecuador, Closing Gaps, which offers a unique opportunity to understand how young children best learn math. At the request of the Government of Ecuador in 2012, an incoming cohort of approximately 15,000 kindergarten children in 202 schools was randomly assigned to their teachers. These children have been randomly reassigned to teachers year after year in their 1st, 2nd, 3rd, and 4th grades. Random assignment is important because it ensures that estimates of the causal effects of teachers on learning outcomes are not biased by systematic sorting of children to teachers.³ In no other country, developed or developing, has there been an experiment that randomly assigned children to teachers in multiple grades. This means that Closing Gaps offers a truly unique opportunity to answer some of the most pressing questions about how to improve student math learning.⁴

² Hanushek and Peterson (2014) estimate that boosting the test scores of US students on math in the PISA test by 50 points (0.5 standard deviations) would increase economic growth rates by about one percent per year.

³ Random assignment on paper does not translate automatically into random assignment in practice. The most influential study of teacher quality in the United States, the Measuring Effective Teaching (MET) study, randomly assigned children to teachers in single grade. Compliance with the random assignment across the six different study sites in MET ranged from 66 percent (in Dallas) to 27 percent (in Memphis). This stands in sharp contrast with Closing Gaps, where compliance with the random assignment was above 98.5 percent in each grade.

⁴ A recent review by a prominent economist (Costas Meghir, professor at Yale University) of the research carried out under Closing Gaps concludes “In my view the data collected from this experiment is one of the most important recent developments in the Economics of Education capable of answering key questions, such as the effect of teachers and the importance of peers.”

- 2.8 Over multiple years and in every grade, the Closing Gaps study has also collected very rich data on teachers, including filming each teacher teaching in the classroom over at least one complete school day. This has allowed Closing Gaps to establish not only whether some teachers are better than others (and by how much), but also identify what characteristics or behaviors make a given teacher more effective than another one.
- 2.9 What have we learned from the Closing Gaps study to date? First, we have learned that there are large differences in the math test scores of poorer and less poor children, and between boys and girls. On average, at the end of kindergarten, a child whose mother has some university education has math test scores that are 0.45 standard deviations higher than a child whose mother has at most completed elementary school; by the end of 2nd grade, these differences have increased to 0.6 standard deviations. Turning to differences by gender, boys have higher math test scores than girls at the end of kindergarten, but the difference is modest, about 0.07 standard deviations; however, by the end of 2nd grade, this difference has doubled, to 0.15 standard deviations (Carneiro, Cruz-Aguayo, and Schady 2017).
- 2.10 Second, even within the same schools, some teachers are substantially more effective than others in the amount of math learning they produce. A single year exposed to an effective, rather than an ineffective teacher, would allow the child of a mother with only elementary school education to make up between one-third and almost half of the gap in math achievement relative to the child of a mother who has some university education. The magnitude of these teacher effects is comparable in kindergarten, 1st, 2nd, and 3rd grades.
- 2.11 Third, we have learned that some of the effects of having a better teacher fade out over time. Thus, after three years, the magnitude of the effect of having a better kindergarten teacher are roughly one-half of the contemporaneous effects. However, the fade-out of teacher effects in Ecuador appears to be substantially lower than what is observed in the United States (Chetty et al. 2014), although the time horizon over which these effects have been calculated is still short.
- 2.12 Fourth, in math (but not language), the fade-out of teacher effects is substantially lower when a child has more than one good teacher in a row.⁵ Put differently, there are dynamic complementarities associated with teacher quality—a child who is taught by an effective teacher in one grade not only learns more in that grade but is also better prepared to take advantage of better teachers thereafter. In the words of Nobel Prize-winning economist James Heckman, “learning begets learning, and skill begets skill”.
- 2.13 Fifth, the skills that are necessary to be an effective math teacher are likely to be very different in different grades. Thus, for the youngest children, the quality of the interactions of children and teachers are very important. To show this, the Closing Gaps study coded the teacher videos using an instrument known as the Classroom Assessment Scoring System (CLASS, Pianta et al. 2007). The CLASS focuses on the quality of the interactions between teachers and students. Based on their behaviors, teachers are assigned a score that ranges from 1 and 7. A kindergarten student randomly assigned to a teacher with a one point higher CLASS score learns roughly 0.2 standard deviations more math over the course of the year. However, this

⁵ To see this, it is useful to compare the learning trajectories of four children. Children 1 and 2 had ineffective kindergarten teachers, while children 3 and 4 had effective kindergarten teachers. As a result, children 3 and 4 learn 0.2 standard deviations more math than children 1 and 2. In 1st grade, children 1 and 3 have ineffective teachers, while children 2 and 4 have effective teachers. Mirroring the results for kindergarten, children 2 and 4 learn 0.2 standard deviations more math than children 1 and 3. Importantly, however, much of the advantage that child 3 had over child 1 at the end of kindergarten is no longer apparent at the end of 1st grade. Why is this the case? Because, in the absence of an effective 1st grade teacher, the effects of a good kindergarten teacher depreciate rapidly. On the other hand, there is almost no depreciation of the kindergarten teacher effects for child 4, who had two effective teachers in a row. By the end of 1st grade, child 4 has test scores that are 0.4 standard deviations higher than child 1.

research also shows that the magnitude of the CLASS effect is smaller in 1st grade than in kindergarten, and smaller still (and not statistically significant) in 2nd grade (Carneiro, Cruz-Aguayo, and Schady 2017). On the other hand, it is likely that the math content knowledge of teachers is largely irrelevant in the earliest grades (when children are being taught basic number recognition and sequencing), but becomes increasingly important in the later grades of elementary school (when children are being taught higher-order math concepts like manipulation of fractions and basic algebra). For example, using data from Peru, Metzler and Woessman (2012) estimate that a one standard deviation increase in teacher math knowledge raises student math achievement in 6th grade by .09 standard deviations. This suggests that the skills that are needed to be an effective math teacher may be very different in, say, kindergarten than in 5th grade.

- 2.14 Sixth, Closing Gaps has shown that peers matter. Specifically, having a single very poorly-behaved child in a classroom significantly reduces a teacher's CLASS scores, and results in about 0.1 standard deviations less learning for all other children in the classroom.
- 2.15 **Extending the Closing Gaps study.** This Technical Cooperation, part of the Closing Gaps project, seeks to extend it focusing on math learning and teaching in 5th grade⁶. It is critical to follow children until the end of the elementary school cycle for numerous reasons. First, more years of data are needed to understand how teacher quality interacts across different grades. In the United States, for example, depreciation of teacher effects continues for at least 5 years (Chetty et al. 2014). To date, we have an imperfect understanding of how cumulative effects of teachers evolve in Ecuador or in any other developing country. This is important because, if there are dynamic complementarities in early learning, it would be more efficient to have some children (perhaps those in the poorest or worst-performing schools) have a string of good teachers than to having every child have at least one good teacher.
- 2.16 Second, if the skills that teachers need to be effective are in fact very different in, say, kindergarten than in 5th grade, there might be room to reallocate some teachers from kindergarten to 5th grade (teachers with high content knowledge and low CLASS scores) and others from 5th grade to kindergarten (teachers with high CLASS scores and low content knowledge). By reassigning teachers across grades in a way that optimizes their strengths, it would in principle be possible to raise the math test scores of children in *both* kindergarten and 5th grade. Moreover, unlike virtually any other intervention in education, this could be accomplished at essentially zero cost. This technical cooperation will finance the filming of teachers in 5th grade, and it will also finance the development of a math test that would be applied to all teachers in the Closing Gaps sample of schools.
- 2.17 Third, in Closing Gaps we filmed all kindergarten, 1st, 2nd, 3rd, and 4th grade teachers while they were teaching. As discussed above, these videos were coded using the CLASS. However, the same videos could be coded with a different rubric—for example, to compare how math is taught in different classrooms. Do teachers who spend more time on intentional, focused math teaching produce more math learning? What is the optimal amount of time spent teaching basic math computations as opposed to seeking to bestow children with a deeper understanding of math concepts? These questions have critical implications for pre-service and in-service training, and for curriculum design.
- 2.18 Finally, for Closing Gaps to inform relevant policies applicable to the whole elementary school learning cycle, it is fundamental that we continue following students in the project as they finish elementary school, in order to understand how the causal relationships established between teaching and learning in the early grades play out in the final grades.

⁶ A recent TC was approved to extend Closing Gaps two more years to study language and executive function (ATN/OC-16118-EC).

- 2.19 Closing Gaps has been subject to wide feedback, constructive criticism, and dissemination at different levels. The first academic paper reporting results from the project was recently published in a top peer reviewed journal (Araujo et al. 2016). Meetings have been carried out regularly with the Minister of Education of Ecuador and his team. Based on Closing Gaps findings, the Bank team has also provided concrete policy advice on specific topics around teacher quality to the governments of Uruguay, Peru, Colombia, Mexico, Jamaica, and Trinidad and Tobago. Results from the project have been presented in several conferences and cited in external⁷ and internal academic and policy oriented publications⁸. More recently, the Bank hosted a regional policy dialogue in which Vice Ministers from the LAC region met to discuss the results of the study. Finally, the evidence produced by the study was used to directly design a recently approved loan to the Government of Ecuador (EC-L1155).
- 2.20 This TC is consistent with the Update to the Institutional Strategy (UIS) 2010-2020 (AB-3008) and is aligned with the development challenge of “social inclusion and equality”, as education is a catalyst to break the inter-generational poverty cycle and helps foster social mobility. This TC is also consistent with the Education and Early Childhood Development Sector Framework Document (GN-2708-5), especially in its third dimension –ensuring that all students have access to effective teachers. And its activities are embedded in the social development sector, considered as a priority in the IDB Country Strategy with Ecuador 2012-2017.⁹

III. Description of activities/components and budget

- 3.1 **Component 1: Measurement of math learning outcomes/math knowledge.** Closing Gaps has developed and applied a set of instruments to measure learning outcomes. Each year, these tests have been adapted to take into account the progress that the children in the sample should have made by the end of that school year, based on the curricula and developmentally-appropriate abilities. The activities to be financed under this component include: 1) A review and adaptation of the current set of math tests, under the guidance of measurement experts, and with inputs from the current Korean math testing system. This exercise will include the incorporation of technology to optimize the application of the tests. 2) piloting and appropriate readjustment of the adapted tests to obtain a final version; 3) the application of the end-of-the-school-year math tests to students in their 5th grade; and 4) the application of a math test to 5th grade teachers. The main output of this component is the databases with the test results.
- 3.2 **Component 2: Teacher assessment on classroom practices for teaching Math.** Teachers in South Korea receive an extensive and intensive training before they begin teaching in their own classroom. Specifically, mathematics instruction for teachers in South Korea represents around 43% of their total course load (Schmidt et. al 2011). The prevalence of ineffective classroom practices observed throughout the Latin American region seems likely to explain much of the gap in student mathematics and science competencies (Näslund-Hadley and Bando 2015). As it is fundamental to observe what happens inside the classroom in order to determine which practices generate more (or less) learning outcomes, Closing Gaps will use specialized instruments for class observation such as the CLASS, and will incorporate the use of other instruments to specifically assess the quality of math teaching practices. The main activities to be covered by this component include: 1) the filming of at least one full day of classes for each of the 5th grade teachers according to the CLASS protocols¹⁰ and those of the specialized instrument to assess the quality of math teaching practices; 2) the selection, training

⁷ Murnane and Ganimian 2014.

⁸ Forthcoming book on Math (Education Division) and in the Development Effectiveness Overview 2013 and 2014.

⁹ The TC objectives are also aligned with the OC-SPD for Social Development objectives of “Enhance relevance, quality, and volume of IDB lending in support of social sector priorities” and “strengthen public institutions efforts to become more effective an efficient in social programming, group targeting, and project execution”.

¹⁰ Previously validated in other countries and with the careful guidance from experts of both Harvard University and the University of Virginia

and evaluation of a team of highly qualified analysts, who will undergo a rigorous certification process to code the videos; 3) the edition and coding of the classroom videos following the standardized protocols; and 4) the recoding of existing videos from previous years of the study using the math rubric. The main output of this component is the databases with the coded videos.

- 3.3 **Component 3: Data analysis and dissemination.** A number of high quality policy briefs, academic papers, reports and presentations will be produced, including a comparative analysis of math teaching practices in Korea and in Ecuador. On this basis, we will produce a set of recommendations drawn from Korea's best practices. This component will be implemented jointly with scholars from Korean Universities¹¹ and/or Korean research institutions, and at least one research fellow from Korea, who is part of the SCL team. The main output of this component is a series of working papers and policy notes with concrete advice and recommendations to be used by Ecuador and other countries in the region as input for policy reforms.
- 3.4 In accordance with the Fund's requirements, a yearly report of findings and recommendations and advances in the results matrix will be produced. Once the execution of the TC is completed a final report will be delivered.

Indicative Budget (US\$)

Activity/Component	IDB/Fund Funding	Counterpart Funding	Total Funding
Component 1. Measurement of math learning outcomes/math knowledge	\$ 605,000.00	-	\$ 605,000.00
Activity 1. Preparation of math assessment instruments for 5th grade students	\$ 25,000.00	-	\$ 25,000.00
Activity 2. Pilot for adjusted math assessment instruments for 5th grade students	\$ 80,000.00	-	\$ 80,000.00
Activity 3. Application of end-of-school-year tests for 5th grade students	\$ 350,000.00	-	\$ 350,000.00
Activity 4. Application of math tests for 5th grade teachers	\$ 150,000.00	-	\$ 150,000.00
Component 2. Teacher assessment on classroom practices for teaching Math	\$ 260,000.00	-	\$ 260,000.00
Activity 1. Classroom teaching observation-film 5th grade	\$ 100,000.00	-	\$ 100,000.00
Activity 2. Selection and certification of coders for 5th grade	\$ 25,000.00	-	\$ 25,000.00
Activity 3. Edition and coding of 5th grade classroom videos	\$ 50,000.00	-	\$ 50,000.00
Activity 4. Recoding of existing classroom videos	\$ 85,000.00	-	\$ 85,000.00
Component 3. Data analysis and dissemination	\$ 135,000.00	-	\$ 135,000.00
Activity 1. Data processing and analysis for policy briefs, papers and reports	\$ 100,000.00	-	\$ 100,000.00
Activity 2. Presentations and dissemination activities	\$ 35,000.00	-	\$ 35,000.00
Total	\$1,000,000.00	-	\$1,000,000.00

- 3.5 The activities to be financed through this TC comprise consultancy services (individuals and firm) and the acquisition of goods is not foreseen for any of the components.

IV. Executing agency and execution structure

- 4.1 As requested by the beneficiary, this TC will be Bank executed through the Social Sector Department (SCL/SCL). The activities programmed for this TC need to be carried out in strict adherence with the existing methodology employed throughout the past five years of the study, in which the Bank has also been the executing agency using the same executing structure¹². The Bank has thus developed a competitive advantage.

¹¹ We have approached and started discussions with Korea University.

¹² ATN/OC-13003-EC, ATN/OC-13782-EC, ATN/OC-14124-EC, ATN/OC-14450-EC, ATN/JF-13365-EC are 100% disbursed; ATN/OC-15468-EC 100% committed and 99.17% disbursed; ATN/OC-15266-EC 100% committed and

- 4.2 The project will primarily collect information during the 2017–2018 school year. However, there is also information that needs to be collected at the beginning of the following school year, such as administrative information and teacher surveys. To allow time for the execution of these activities, the TC will disburse in 36 months and execute in 30 months.
- 4.3 The Bank will contract individual consultants, consulting firms and non-consulting services in accordance with the Bank's procurement policies and procedures. The Bank will procure the consulting services according to the applicable Bank's policies and procedures. For Components 1 and 2, single source procurement is recommended, to be executed by the service provider "Habitus MillwardBrown S. A." (Habitus) for data collection, this in accordance with paragraph 2.7 and 2.9 of the Operational Guidelines for the Selection and Contracting of Consulting Firms in Bank-executed operational work. The justification for single source is based on the necessity of a firm with a proven record of expertise in technical fieldwork logistics and on collecting this type of information in several specific key areas of the country. Habitus was contracted by PRC (Corporate Procurement) initially for this project and has been granted approval by the Corporate Procurement Committee (PRC) of the Bank of a major amendment adding additional scope of services to the existing contract.¹³ Habitus has acquired unique experience in performing surveys and tests, providing accurate data and coded videos, following very strict and specific research protocols. Furthermore, Habitus has been working for the past five years in close coordination with MinEduc, so it has an unparalleled experience to work with the Ministry and school authorities for this project. This institutional knowledge is necessary for an appropriate follow-up. Hence, Habitus provides unique services and is the best candidate to continue carrying out the necessary tasks for the project.

V. Project risks and potential issues

- 5.1 Specific implementation risks are considered low/medium for the components. A few school principals/teachers/parents might not comply with the planned activities, including testing of students and filming. In those cases, the Ministry of Education has committed to mediating and correcting any deviations from the plan. In the four years of the implementation of the project there has never been an extreme case where a stakeholder refused to participate.
- 5.2 This TC is part of a longitudinal project that started in 2012 and as such it has overcome many changes in government counterparts, so it has proved its sustainability over time. Currently, the Ministries of Finance and Education have confirmed that Closing Gaps is a highly relevant project and that the country will continue to support this effort.

VI. Exceptions to Bank policy

- 6.1 None

VII. Environmental and Social Strategy

- 7.1 The ESG classification for this TC is "C" according to the Environment and Safeguards Compliance Policy (OP-703). There are no potential negative environmental and/or social impacts associated to this TC anticipated ([See filters](#)).

Required Annexes:

- [Annex I - Government request](#)
- [Annex II - Results Matrix](#)
- [Annex III - Terms of Reference](#)
- [Annex IV - Procurement Plan](#)

93.24% disbursed; ATN/OC-15690-EC 100% committed and 90.72% disbursed; ATN/OC-16118-EC 4.8% disbursed. Each of these TCs has produced at least two datasets, that are currently being used in the production of academic papers and policy notes. Technical cooperation resources approved for this project from 2011 up to date amount \$11,596,729.00.

¹³ See IDB Docs document <http://www.iadb.org/projectDocument.cfm?id=40414595>