



USAID
FROM THE AMERICAN PEOPLE

Evaluation of *Amazonía Lee* Reading Intervention in Peru

First report



October 2016 (revised January 2021)

This report was produced for review by the United States Agency for International Development under task order AID-OAA-M-12-00020. It was prepared by Larissa Campuzano, Steve Glazerman, Nancy Murray, Sarah Liuzzi, and Camila Fernández of Mathematica Policy Research. Photo credit: Ivonne Padilla, Mathematica

Evaluation of *Amazonía Lee* Reading Intervention in Peru

First report

DISCLAIMER

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

ABSTRACT

This report presents baseline results of the impact evaluation of *Amazonía Lee* (Amazon Reads), an early-grade reading intervention funded by the U.S. Agency for International Development (USAID). *Amazonía Lee*'s goal is to improve the reading skills of children in two Amazonian regions of Peru: Ucayali and San Martín. The impact evaluation is based on a random assignment design. In each region, schools were randomly assigned to a group that implemented *Amazonía Lee* or to a control group that did not implement the program. In Ucayali, control schools received the usual services provided by the Ministry of Education. In San Martín, most control schools received a program known as *Soporte Pedagógico*, a Ministry initiative that offered teacher training and coaching, pedagogical materials, and remedial tutoring.

The results confirm that the random assignment created groups with similar characteristics at baseline. In 2014, *Amazonía Lee* schools had similar facilities and implemented similar programs than control schools in both regions. In 2014, teachers had similar demographic characteristics and used similar reading activities and instructional approaches. Furthermore, teachers divided their reading instruction time in similar ways. In 2015, first grade students in the *Amazonía Lee* group had similar emergent reading skills than students in the control group in both regions with one exception. In Ucayali, *Amazonía Lee* students had better phonemic awareness.

In both regions, *Amazonía Lee* teachers were more likely to have received training than teachers in the comparison schools, a few months after *Amazonía Lee* training had begun in 2015. But in Ucayali, the difference in training was larger. Despite the training differences, teacher outcome data suggest only a few early impacts of *Amazonía Lee* on teachers use of time and instructional method were found.

This page has been left blank for double-sided copying.

ACKNOWLEDGMENTS

The authors gratefully acknowledge numerous people and organizations that contributed to the implementation of all evaluation activities in Peru and to the production of this report.

The staff at the regional government of the regions of San Martín and Ucayali and their local education administrations, Dirección Regional de Educación (DRE) provided a warm welcome and essential institutional support for the evaluation. Pilar Saavedra, Wilson Guerrero, and Ricardo Quevedo in San Martín and Wagner Sánchez, Victor García, and Zoila García in Ucayali provided important support in coordinating evaluation activities. They shared key information on the schools and communities in the two regions and offered continuous assistance during our data collection efforts.

Universidad Peruana Cayetano Heredia (UPCH), *Amazonía Lee* project director Raquel Villaseca provided us key information on program implementation and valuable input on final timeline for the evaluation design. UPCH's technical and administrative project team in San Martín and Ucayali also provided valuable information and assistance in the coordination of the evaluation activities.

USAID/Washington and USAID/Peru staff have provided guidance and support to the evaluation design and its implementation. They include, from USAID/Washington, Barbara Knox-Seith, Senior Education Advisor Bureau for Latin America and the Caribbean, and while at USAID/Peru, Clifford Brown, Deputy Chief, Governance & Institutional Strengthening; Natalia Petroni, Monitoring and Evaluation Specialist; Sobeida Gonzalez, Project Coordinator Governance and Institutional Strengthening; Miriam Choy, Development Assistance Specialist; Evelyn Rodriguez-Perez, Chief, Health and Education Office; Brindis Ochoa, Education Project Management Specialist; Jo Jean Elenes, Acting Chief, Health and Education Office; and Patricia Perez, Education Advisor, Health and Education Office.

We thank our local evaluation partner, Grupo de Análisis para el Desarrollo (GRADE), and its staff for high quality data collection in the study schools and communities. Miguel Jaramillo ensured careful planning of evaluation activities in Peru, and provided keen oversight of all data collection activities. Elsa Bardalez developed a collaborative relationship with regional staff at the DREs and has been instrumental in obtaining and maintaining local support for the study in the Amazonian region. Juan José Díaz provided overall project management support, oversaw the production of the datasets, and ensured consistency of data programming. Carla Pajares conducted thorough training of field staff, and monitored data collection activities closely. Mariel Bedoya assisted in various study activities including field staff training, data entry, and dataset production.

We are also indebted to Sylvia Linan-Thompson for her thorough and incisive input at different stages of the evaluation. Her expertise in models of reading interventions and ample experience with measurement of early grade reading skills were very useful in designing the evaluation and interpreting baseline findings.

At Mathematica Policy Research, Ivonne Padilla Espinosa, Galina Lapadatova, Irina Cheban, Betty Teller, Jessica Hegedus, and Sheena Flowers provided critical support in data

cleaning, programming, and presentation. We are enormously appreciative of all the school directors and teachers for opening the doors of their classrooms and allowing us to observe the valuable work they do with students. Last, but not least, we thank the students who provided us with the important information we are able to share in this report.

CONTENTS

EXECUTIVE SUMMARY	ES-1
A. Introduction	ES-1
B. Evaluation questions and design	ES-2
C. San Martín baseline findings for schools, teachers, and students	ES-4
D. Ucayali baseline findings for schools, teachers, and students	ES-5
E. San Martín early impact findings in teacher outcomes	ES-7
E. Ucayali early impact findings in teacher outcomes	ES-8
F. Conclusions	ES-9
I. INTRODUCTION	1
A. The <i>Amazonía Lee</i> intervention	2
B. Evaluation questions, design, and limitations	5
C. Data sources and indicators	9
II. FINDINGS: SAN MARTÍN: SCHOOLS, TEACHERS, AND STUDENTS AT BASELINE	13
A. School characteristics	13
B. Teacher characteristics	14
C. Teaching materials and teaching practices	16
D. Student characteristics and emergent literacy assessment	19
III. FINDINGS: UCAYALI: SCHOOLS, TEACHERS, AND STUDENTS AT BASELINE	24
A. School characteristics	24
B. Teacher characteristics	26
C. Teaching materials and teaching practices	28
D. Student characteristics and emergent literacy assessment	31
IV. FINDINGS: EARLY IMPACTS ON TEACHER OUTCOMES IN SAN MARTÍN	35
A. Impacts on participation in teacher training programs	35
B. Early impacts on teachers' instruction	37
V. FINDINGS: EARLY IMPACTS ON TEACHER OUTCOMES IN UCAYALI	43
A. Impacts on participation in teacher training programs	43
B. Early impacts on teachers' instruction	45

VI. CONCLUSIONS AND NEXT STEPS	51
A. Conclusions from comparing schools, teachers, and students at baseline	51
B. Conclusions from early impact analysis on teacher outcomes	52
C. Next steps.....	53
REFERENCES.....	55
APPENDIX A ADDITIONAL INFORMATION ON SCHOOL CHARACTERISTICS.....	A.1
APPENDIX B ADDITIONAL INFORMATION ON TEACHERS CHARACTERISTICS AND FACTORS TEACHERS IDENTIFY AS LIMITING THEIR ABILITY TO TEACH CHILDREN	B.1
APPENDIX C ADDITIONAL INFORMATION ON TEACHERS PRACTICES	C.1
APPENDIX D ADDITIONAL INFORMATION ON STUDENTS CHARACTERISTICS	D.1

TABLES

I.1	Timing of baseline data collection	11
II.1	Emergent literacy skills by group, 2015	22
III.1	Emergent literacy skills by group, 2015	34
IV.1	Teacher participation in training by training organization and group, 2015	35
IV.2	Percentage of teachers who report using reading materials at least 3–4 times per month, 2015	38
IV.3	Teachers' self-reported time spent on reading activities in a typical reading class by group (minutes), 2015	40
IV.4	Teachers' self-reported time use by group: 2015 (hours per week)	41
V.1	Teacher participation in training by training organization and group, 2015	43
V.2	Percentage of schools in which students used reading materials at least 3–4 times per month, 2015	46
V.3	Teachers' self-reported time spent on reading activities in a typical reading class by group, 2015	48
V.4	Teachers' self-reported time use by group, 2015 (hours per week)	49
A.1	Study school characteristics, 2014	A.1
B.1	Teacher characteristics	B.1
C.2	Time use in reading lesson by group (minutes), 2014	C.3
D.2	Emergent reading and writing skills by group, 2014	D.3

This page has been left blank for double-sided copying.

FIGURES

I.1	Percentage of second grade students who achieved the expected (satisfactory) grade-level reading category according to the ECE reading test.....	1
I.2	Map of study locations in Peru	2
I.3	Evaluation design: randomized controlled trial	6
I.4	Content of 1st-grade emergent literacy assessment	10
II.1	School characteristics by group in San Martín, 2014	13
II.2	Support programs in study schools by group in San Martín, 2014	14
II.3	Teachers' educational attainment by group, 2014	15
II.4	Training content areas by group, 2014	16
II.5	Reading instruction approaches teachers reported using by group, 2014	17
II.6	Reading materials available in classrooms by group in San Martín, 2014	18
II.7	Student characteristics by group, 2015.....	20
II.8	Percentage of students with a score of zero by skill and group, 2015	22
III.1	School characteristics by group in Ucayali, 2014	25
III.2	Support programs in study schools by group in Ucayali, 2014	26
III.3	Teachers' educational attainment by group, 2014	27
III.4	Training content areas by group, 2014	28
III.5	Reading instruction approaches teachers reported using by group, 2014	29
III.6	Reading materials available in classrooms by group in Ucayali, 2014	30
III.7	Student characteristics by group, 2015.....	32
III.8	Percentage of students with a score of zero by skill and group, 2015	34
IV.1	Training content areas by group, 2015	36
IV.2	Reading instruction approaches teachers reported using by group, 2015	37
IV.3	Daily assessments of students' reading skills by group, 2015.....	39
V.1	Training content areas by group, 2015	44
V.2	Reading instruction approaches teachers reported using by group, 2015	45
V.3	Daily assessments of students' reading skills by group, 2015.....	47
B.2	Factors teachers identify as limiting their ability to teach children to read, 2014.....	B.3
C.1	Reading skill-building activities observed by group, 2014	C.1

D.1	Student characteristics by group, 2014.....	D.1
D.3	Percentage of students with a score of zero by skill and group, 2014 and 2015	D.4

EXECUTIVE SUMMARY

A. Introduction

This report presents baseline results of the impact evaluation of *Amazonía Lee* (Amazon Reads), an early-grade reading intervention funded by the U.S. Agency for International Development (USAID). *Amazonía Lee*'s goal is to improve the reading skills of children in two Amazonian regions of Peru: Ucayali and San Martín. Ucayali and, to a lesser extent, San Martín lag behind the nation and the metropolitan area of Lima in important student reading achievement indicators.

Amazonía Lee was implemented by the Department of Regional Education Offices (DREs) of Ucayali and San Martín starting in 2015. In San Martín, the program is known as “*Programa Enseñar Es Liderar*” (PEEL). In Ucayali the program is called “*Liderando los Aprendizajes de Niñas y Niños*” (LANN). The two *Amazonía Lee* programs share two main components, described below, and the same technical assistance provider, Universidad Peruana Cayetano Heredia (UPCH).

1. Capacity building component. The aim of this component is to strengthen pedagogical skills and instructional capabilities of teachers, school directors, and regional specialists. At the beginning of 2015, UPCH started providing technical assistance to the DREs in both regions for the implementation of the following five capacity-building mechanisms:

- **Teacher training workshops based on *Amazonía Lee*'s balanced approach.** *Amazonía Lee*'s balanced approach to reading instruction integrates evidence-based foundational skills instruction with the Ministry of Education's communicative-textual approach. Training workshops for 1st- to 3rd-grade teachers aim to improve teachers' approach to reading instruction, class and curricular planning, differentiated instruction based on students' learning needs, and personalized attention for underperforming students.
- **Teacher coaching.** Specialists conduct classroom visits to coach teachers on-site and help them improve content area knowledge and pedagogical strategies for literacy instruction, differentiated instruction, assessment and monitoring of student learning, and effective classroom management.
- **Teacher study groups.** Teachers gather in small groups of 8 to 16 participants to strengthen their learning, review pedagogical materials, share teaching strategies, and plan for classes with support from other teachers.
- **More and more varied pedagogical materials to support reading.** The goal of this component is to improve quality and regional pertinence of pedagogical materials for reading and writing instruction, and increase the availability of classroom book libraries in primary grades.
- **School leadership and management training.** School directors will receive training on pedagogical leadership and school management and support from specialists on issues related to pedagogical matters and institutional management. Regional governments receive technical assistance to improve school planning and budgeting, and execution, and to implement information systems for school management indicators.

2. Regional component. In addition to the five capacity-building mechanisms described above, *Amazonía Lee* includes another three activities envisioned to be implemented at the regional level in both San Martín and Ucayali.

- **Community engagement.** The community engagement component aims to promote the active engagement of the educational community in support of student learning, particularly in gaining reading proficiency during the first three grades of primary education.
- **Regional assessment system.** The goals of this subcomponent are (1) to introduce regional assessments at the beginning and end of the school year to monitor student achievement (on reading comprehension and production of texts); and (2) to develop local capacity in the design, implementation, scoring, data production and interpretation of these assessments.
- **Regional incentive system for teachers.** This subcomponent has regional policies to motivate, encourage, recognize, and reward teacher performance—thereby engaging all stakeholders of the education system (teachers, specialists or facilitators, school directors, students, and others) in efforts to improve student learning and academic achievement.

The program logic is that these components should result in improved teacher training for reading instruction, more relevant reading materials, better school management, increased support of parents and the community for reading. These in turn should lead to students' greater acquisition of literacy skills—the program's ultimate goal.

B. Evaluation questions and design

The impact evaluation, of which this baseline data is a first step, is structured to address the following central evaluation question:

- What was the impact of *Amazonía Lee* programs on early-grade reading and other intermediate outcomes relative to other prevailing practices and programs?

Additionally, it will address the below questions related to program implementation:

- How were *Amazonía Lee* intervention components, activities, and processes implemented in Ucayali and San Martín?
- What services were provided to control schools in Ucayali and San Martín?
- How were *Amazonía Lee* components, activities, and impacts perceived by principals and teachers?
- What were the facilitators and barriers to the implementation of *Amazonía Lee* components?

The central questions of the baseline data collection are:

- What are the characteristics of the study schools and teachers in 2014 (the year before the intervention began)?

- What are the initial characteristics and early reading assessment of first grade students in 2015 who will be followed by the evaluation?
- To what extent did random assignment produce equivalent groups based on measurable characteristics for schools and teachers in 2014 and for students in 2015?
- What are the early impact findings of *Amazonía Lee* on teacher outcomes in 2015 such as participation in training and approaches to reading instruction?

Impact evaluation design. Mathematica is working with USAID/Washington, USAID/Peru, the Ministry of Education in Peru, and especially with the regional offices of San Martín and Ucayali, to conduct a rigorous impact evaluation of *Amazonía Lee*. The impact evaluation employs a randomized controlled trial, the most rigorous evaluative framework for assessing program impacts. We randomly assigned eligible schools in the targeted regions of San Martín and Ucayali to one of two groups: group A was assigned to receive all of the components of *Amazonía Lee* and group B was assigned to receive the services that the Ministry of Education already provides in each region. Both groups will receive the components or activities of *Amazonía Lee* that must be implemented for the whole region.

Because the Ministry of Education was providing different services to the schools assigned to the control group (group B) in each of the two departments at the time of the evaluation, we conducted separate analyses for each department. In Ucayali, control schools received the usual services the Ministry of Education provides in the department. In San Martín, a majority of the control schools received additional assistance in the form of a primary education support program known as *Soporte Pedagógico* (Pedagogical Support), a Ministry of Education initiative implemented in San Martín starting in 2015. *Soporte Pedagógico* aims to improve student learning in the areas of mathematics, communication, science, and social skills in urban elementary schools by strengthening skills of teachers and administrators, providing after-school tutoring for students not meeting the expected learning standards, delivering educational materials and resources, and improving the school and local management capacity.

Timing of baseline data collection. In our original evaluation design, the target student population comprised 1st graders who were enrolled in first grade in the study schools at the time of random assignment—school year 2014 (2014 cohort). We collected baseline data for these students in 2014. However, *Amazonía Lee* implementation did not start during the 2014 school year as initially planned but was delayed until 2015. Therefore, the target student population that the evaluation focused on changed to students enrolled in 1st grade in 2015 (2015 cohort). We collected baseline data for this new student cohort and their teachers in 2015.

Data sources and indicators

The main quantitative data sources for the impact evaluation are student assessments, classroom observations, teacher surveys, and a school infrastructure checklist.

- At baseline, we assessed students’ oral language proficiency and early literacy skills. The student assessment consisted of selected subtasks adapted from two instruments that have been used with dual-language learners in Latin America: the Emergent Literacy Profile (Dickinson et al. 2004); and the Early Grade Reading Assessment (EGRA).

- The classroom observation measured teachers' use of class time, their pedagogical practices, and the classroom literacy environment.
- Teacher surveys gathered self-reported use of instructional practices, proportion of students in the classroom that achieved milestones in reading acquisition, use of formative assessments, participation in professional development or in-service training, education and experience, occupational needs, career expectations, and background characteristics.
- The school infrastructure checklist gathered information on general characteristics of the schools' infrastructure such as the number of classrooms, the distance between the school and the district capital, and the availability of electricity and potable water.

C. San Martín baseline findings for schools, teachers, and students

Amazonía Lee schools had similar facilities to Soporte Pedagógico schools. We did not find any statistically significant differences in school facilities between the groups. In both groups, most of the schools had electricity, blackboards in most classrooms, functional restrooms for students, piped water, an outdoor recreational area, and computers for teachers and students. However, in both groups other important facilities were lacking, such as a library and internet connectivity, and nearly half lacked potable water.

Teachers in both groups had similar experience and demographic characteristics. In 2014, we did not find statistically significant differences between groups on teacher experience, tenure, age, or gender. In both groups, teachers had about 16 years of experience, had been teaching 9 years at the current school, were in the early forties, and three-fourths were women.

Teachers in the *Amazonía Lee* group were less likely to have advanced degrees beyond undergraduate education than teachers in the *Soporte Pedagógico* group. The percentage of teachers with an associate's or technical degree was similar between groups, but a greater proportion of teachers had a graduate degree in the *Soporte Pedagógico* group. The difference in the distribution of educational attainment was statistically significant.

***Amazonía Lee* teachers were less likely to have received in-service reading instruction training in the past year than teachers in *Soporte Pedagógico* schools.** In 2014, before implementation of either program began, teachers at 12 percent of *Amazonía Lee* schools had received in-service training in reading instruction compared to 32 percent at *Soporte Pedagógico* schools.

There were no statistically significant differences between groups on the instructional approach teachers reported using. In 2014, before the intervention started, in both groups, approximately 65 percent of teachers reported using the syllabic method, and 38 percent reported using the phonological-global method. Fifty-eight percent teachers in the *Soporte Pedagógico* group reported using the communicative-contextual method and 48 percent in the *Amazonía Lee* group, but the difference was not statistically significant.

Availability of reading materials was similar in the two groups. We did not find statistically significant differences in availability of reading materials between the groups. Although not statistically significant, *Amazonía Lee* classrooms were more likely than prevailing practice classrooms to have textbooks, workbooks, and picture books with some text.

Teachers used a similar set of reading instruction activities overall. In 2014, we did not find significant differences in teachers' likelihood of using 14 out of 15 reading activities. However, *Amazonía Lee* teachers were significantly more likely to use oral reading and fluency-building exercises at baseline.

Teachers in the two study groups divided their reading instruction time in similar ways. Teachers in both groups spent roughly half the time in the observed period on phonological and phonemic awareness and a quarter of the time on activities involving written texts.

Teachers in both groups spent most of their time teaching the whole class and almost no time teaching small groups of students. We did not find statistically significant differences between the groups on how teachers divided their time among working with the whole class, individual students, and small groups.

***Amazonía Lee* students were more likely to be repeating 1st grade and to be overage than *Soporte Pedagógico* students.** In 2015, students in the *Amazonía Lee* group were 2 percentage points more likely to be older than expected for their grade (age 7 or older in 1st grade) and 0.4 percentage points more likely to be repeating a grade.

Spanish is the dominant language in both groups. There was no significant difference between groups in the likelihood of the students or their families speaking a language other than Spanish. There was also no significant difference between groups in the likelihood Spanish being the main language of instruction at school.

We found no statistically significant differences between groups in emergent reading skills of first grade students in 2015. Students in both groups had similar foundational skills.

- In both groups, students scored roughly 18 out of 24 points on letter identification.
- In both groups, students scored approximately 12 out of 22 on emerging writing skills.
- In both groups, students could identify the initial sound of roughly 3 out of 10 words
- In both groups, students could only read correctly approximately 2 out of 9 familiar words and 2 pseudo-words in one minute.
- In both groups, students responded to roughly 3 questions correctly after the assessor read a passage and asked 6 questions.

D. Ucayali baseline findings for schools, teachers, and students

There were no statistically significant differences in school facilities between the groups. Although not statistically significant, some differences favor the *Amazonía Lee* group, such as internet connectivity (9 percentage points), computer for teachers (17 percentage points), and potable drinking water (9 percentage points). The prevailing practice group, however, is more likely to have computers for students (9 percentage points).

The study programs do not operate in a vacuum. Many schools participate in various support programs. Only one difference between groups was significant. More schools in the

prevailing practice group participated in the One Laptop Per Child program (86 percent, compared to 66 percent in the *Amazonía Lee* group).

Teachers had similar demographic characteristics across groups. We did not find statistically significant differences between groups on teacher age or gender.

Teachers in the *Amazonía Lee* group had more education but less experience teaching than teachers at prevailing practice schools. At *Amazonía Lee* schools, 31.5 percent of teachers had completed a bachelor's or graduate degree, compared to 11.4 percent of teachers at prevailing practice schools. However, at prevailing practice schools, teachers had 14.8 years of teaching experience, compared to 10.7 years at *Amazonía Lee* schools.

We found no significant difference in training received by teachers by study group. In 2014, for the prevailing practice schools, 29 percent of teachers had received training on reading instruction, compared to 23 percent of teachers at *Amazonía Lee* schools. The difference is not statistically significant.

There were no statistically significant differences between groups on the instructional approach teachers reported using. The majority of teachers in both groups reported using the syllabic method for reading instruction in 1st-grade (83 and 69 percent). Also, around half of the teachers in both groups reported using the communicative-textual approach. The phonological-global method was reported using by 49 percent of *Amazonía Lee* teachers and by 37 percent of prevailing practice teachers. However, these differences were not statistically significant.

Availability of reading materials was similar in the two groups. We did not find statistically significant differences in availability of reading materials between the groups.

Teachers in both groups used a similar set of reading instruction activities overall. We did not find significant differences in teachers' likelihood of using 14 out of 15 reading activity types in the class observation. However, *Amazonía Lee* teachers were significantly more likely to teach writing.

Teachers in the two study groups reported similar duration of most reading activities except for phonological and phonemic awareness and vocabulary development activities. *Amazonía Lee* teachers spent more time on phonological and phonemic awareness activities (24 minutes compared to 20). *Amazonía Lee* teachers spent less time on vocabulary development (0.9 minutes compared to 2.7).

Teachers in both groups spent most of their time teaching the whole class and almost no time teaching small groups. We did not find statistically significant differences between the groups on how teachers divided their time among working with the whole class, individual students, and small groups.

***Amazonía Lee* students were more likely to be older than average for 1st grade, but were otherwise similar.** In the *Amazonía Lee* group, 8 percent of students were age 7 years or older, compared to 4 percent in the prevailing practice schools.

Spanish is the dominant language of students in both groups. There was no significant difference between groups in the likelihood of the students or their families speaking a language other than Spanish. There was also no significant difference between groups in the likelihood that Spanish was the main language of instruction at school.

Students in both groups had similar emergent reading skills in 2015 with one exception. Students in *Amazonía Lee* had better phonemic awareness skills. First grade students in both groups had similar skills on letter identification, decoding pseudo-words, emergent reading and writing, and passage comprehension. But we found a statistically significant difference in phonemic awareness favoring *Amazonía Lee* students.

- Students' emergent writing score was on average 17.7 points out of 24 for *Amazonía Lee* and 18.5 for prevailing practice. The difference is not statistically significant.
- Students' emergent writing score was on average 10.2 points out of 22 for *Amazonía Lee* and 10.7 for prevailing practice. The difference is not statistically significant.
- Students could only read on average almost 2 out of 9 familiar words and 2 pseudo-words in a minute in both groups. There was no statistically significant difference between groups.
- Students responded to roughly 2 questions correctly after the assessor read a passage and asked 6 questions. There was no statistically significant difference between groups found.
- *Amazonía Lee* students were able to identify the initial sound of 3 out of 10 words, while prevailing practice students identified only 2.7 initial sounds. The difference of 0.3 is small but statistically significant. Furthermore, more students in the prevailing practice group were unable to correctly identify at least one initial sound than in the *Amazonía Lee* group (7 percentage point difference).

E. San Martín early impact findings in teacher outcomes

Because 2015 data was collected up to three months after *Amazonía Lee* training began, it is possible that teachers would have already changed their instruction based on the training they received. We, therefore, consider the findings for 2015 teacher surveys as potential early impacts of the intervention.

In 2015, *Amazonía Lee* teachers were 26 percentage points more likely to have participated in training than *Soporte Pedagógico* teachers. We found that 80 percent of teachers at *Amazonía Lee* schools and 54 percent of teachers in the *Soporte Pedagógico* group had participated in training, a statistically significant difference of 26 percentage points.

***Amazonía Lee* teachers were more likely than *Soporte Pedagógico* teachers to be trained in most content areas related to reading, but less likely to be trained in other curricular areas.** Teachers in the *Amazonía Lee* group were more likely than teachers in the *Soporte Pedagógico* group to report having received training in most reading areas, and most differences were statistically significant. Two of the largest differences were on phonological awareness (44 percentage points) and letter knowledge (33 percentage points).

In 2015, *Amazonía Lee* teachers were more likely to report using the phonological-global method and less likely to report using the communicative-textual method than the *Soporte Pedagógico* teachers. In 2015, 67 percent of teachers in *Amazonía Lee* and 41 percent

of teachers in *Soprote Pedagógico* reported using the phonological-global method—a statistically significant difference of 26 percentage points. We also found that 39 percent of the teachers in *Amazonía Lee* reported using the communicative contextual approach, versus 62 percent of *Soprote Pedagógico* teachers—a statistically significant difference of 23 percentage points.

We found no significant differences between *Amazonía Lee* and *Soprote Pedagógico* teachers in the type of reading instruction materials used, the use of daily assessments or the type of reading skills assessed.

Teachers from both groups divided their time in similar ways among various reading instruction activities except for reading aloud. We found no statistically significant differences on the amount of time spent on reading activities such as learning letters, practicing phonological or phonemic awareness, copying words, building vocabulary, and narrating stories without text. However, *Amazonía Lee* teachers spent less time practicing reading aloud and reading comprehension than did teachers at *Soprote Pedagógico* schools (32.5 minutes compared to 38.1), and that difference was statistically significant,

We found a positive impact of *Amazonía Lee* on total time spent in teaching activities. *Amazonía Lee* teachers reported spending on average 28 hours per week teaching, whereas *Soprote Pedagógico* teachers reported spending 23 hours per week teaching. The difference of 5 hours per week is statistically significant.

E. Ucayali early impact findings in teacher outcomes

Because 2015 data was collected up to three months after *Amazonía Lee* training began, it is possible that teachers would have already changed their instruction based on the training. We, therefore, consider the findings for 2015 teacher surveys as potential early impacts of the intervention.

In 2015, *Amazonía Lee* teachers were 69 percentage points more likely to have participated in training than teachers in the prevailing practice group. At *Amazonía Lee* schools, 80 percent of teachers indicated having received teacher training, compared to 11 percent of teachers at prevailing practice schools—a significant difference of 69 percentage points.

***Amazonía Lee* teachers were more likely than prevailing practice teachers to be trained on topics related to reading instruction, and on classroom management.** In 2015, teachers in the *Amazonía Lee* group were more likely than teachers in the prevailing practice group to report receiving training in all reading areas and all differences were statistically significant. We found a 69 percentage point difference for training on reading comprehension, letter knowledge, and phonological awareness, and a difference of over 50 percentage points for training on vocabulary, reading fluency, and creating a class environment to promote reading.

In 2015, teachers in the *Amazonía Lee* group were more likely to report using the phonological-global method than those in the prevailing practice group. Sixty-six percent of *Amazonía Lee* teachers reported using the phonological-global approach and 26 percent of the prevailing practice group did—a statistically significant difference of 40 percentage points.

We found no significant differences between *Amazonía Lee* and prevailing practice teachers in the type of reading instruction materials used, the use of daily assessments or the type of reading skills assessed.

Teachers from both groups divided their time in similar ways among various reading instruction activities. We found no statistically significant differences on the amount of time spent on reading activities such as learning letters, practicing phonological or phonemic awareness, copying words, building vocabulary, and narrating stories without text.

We found no significant differences between groups in the amount of time teachers spent on reading activities during a class period or in teaching reading-related activities. In both groups, teachers reported spending most time practicing reading aloud and reading comprehension, and an average of 28 minutes on each type of reading activity in a typical weekday. We did not find evidence of impacts of *Amazonía Lee* on how teachers spent their time. In both groups, teachers spent on average 23 hours per week teaching, 13 hours preparing for class, and 9 hours on other activities.

F. Conclusions

1. Conclusions from comparing schools, teachers, and students at baseline

The random assignment created similar groups of schools in both regions. In 2014, before the intervention started, schools in both study groups were statistically similar in both regions. No systematic differences were found, but some school facilities, such as the availability of drinking water supply or internet connectivity, were more prevalent in the *Amazonía Lee* group in both San Martín and Ucayali; others, such as computers for students, were more prevalent in the prevailing practice group in Ucayali, but the differences are not statistically significant.

The impact estimation will include school characteristics as covariates in the regression framework to improve the precision of the estimates. No systematic differences on school characteristics were found between groups in both regions. However, because baseline differences in school characteristics could be related to student outcomes and a few differences were found, we will include school characteristics as covariates in the regression model used for impact estimations. The differences in school characteristics are likely to be unrelated to the interventions' outputs or intended outcomes. Therefore, controlling for school characteristics will not bias the result but can improve the statistical precision of our impact estimates.

The random assignment created teacher groups with similar demographic characteristics, reading instructional approaches, and use of time in both regions. In 2014, before the intervention started, in both regions, teachers in both groups had similar age, gender, experience, and tenure, but differed on education. Teachers reported similar approaches to reading instruction before the intervention started. In both regions, the syllabic-method was the instructional approach most commonly reported, followed by the communicative-textual and the phonological global. Furthermore, when teachers were observed, we found no differences in the type of reading instruction activities teachers used (with one exception in San Martín). We also found no differences in the time spent across reading activities types in both regions.

Impacts of *Amazonía Lee* may be diluted in San Martín due to the greater amount of training received by teachers in the *Soporte Pedagógico* group before *Amazonía Lee* started. While we found no statistically significant differences in teacher training across study groups in Ucayali in 2014, in San Martín teachers in the *Soporte Pedagógico* group were more likely to have received training than teachers in the *Amazonía Lee* group. The implication for the evaluation is that the *Amazonía Lee* group will be compared to a group that had already received more teacher training in 2014 prior to the rollout of the intervention, which could potentially dilute impacts of *Amazonía Lee*.

We found one or two differences in students' characteristics in both regions. We, therefore, will include these student characteristics as covariates in the impact estimation model. In San Martín, *Amazonía Lee* students were more likely to be repeating 1st grade and to be overage than *Soporte Pedagógico* students. In Ucayali, the *Amazonía Lee* group, was more likely to be overage. It is possible that being overage or repeating has a direct effect on the final student outcomes; therefore, we will control for these student characteristics in the final impact estimation.

In 2015, students in both groups had similar emergent reading skills in both regions, except for phonemic awareness in Ucayali. In 2015, in both regions, first grade students in both groups had similar skills on letter identification, decoding pseudo-words, emergent reading and writing, and passage comprehension. However, in Ucayali, we found a statistically significant difference in students' phonemic awareness scores favoring the *Amazonía Lee* group. Because 2015 data were collected about two months after the intervention had begun, the 2015 emergent literacy assessments results may have already been affected by the intervention. While it is unlikely that the intervention could affect reading outcomes in such a short time, phonemic awareness could be an exception. Therefore, we cannot be certain if the difference in phonemic awareness is a pre-existing difference unrelated to the intervention or if it is the result of an early effect of the *Amazonía Lee* program.

Student emergent skills that were found to be similar between groups in 2015 will be included as covariates in the impact estimation. Emergent skills for which differences were found in 2015 will not be included as covariates. We plan to use the emergent reading skills scores for which no differences were found in 2015 as covariates in the regressions for final impact results. Including these covariates will allow us to account for initial skills and improve the statistical precision of the impact. However, the difference in students' phonemic awareness skills between groups we found in Ucayali could represent an early impact of the intervention. Therefore, we will not include as covariates reading skills for which differences in 2015 were found.

We will include school averages of student assessments collected in 2014 as covariates in the impact estimation model. Student assessment results from 2014 are not affected by the intervention given that the *Amazonía Lee* roll-out started in 2015. By including the school-level averages from 2014 in the impact estimation models, we can control for initial reading differences between groups and improve the precision of our estimates.

2. Conclusions from early impacts on teacher outcomes

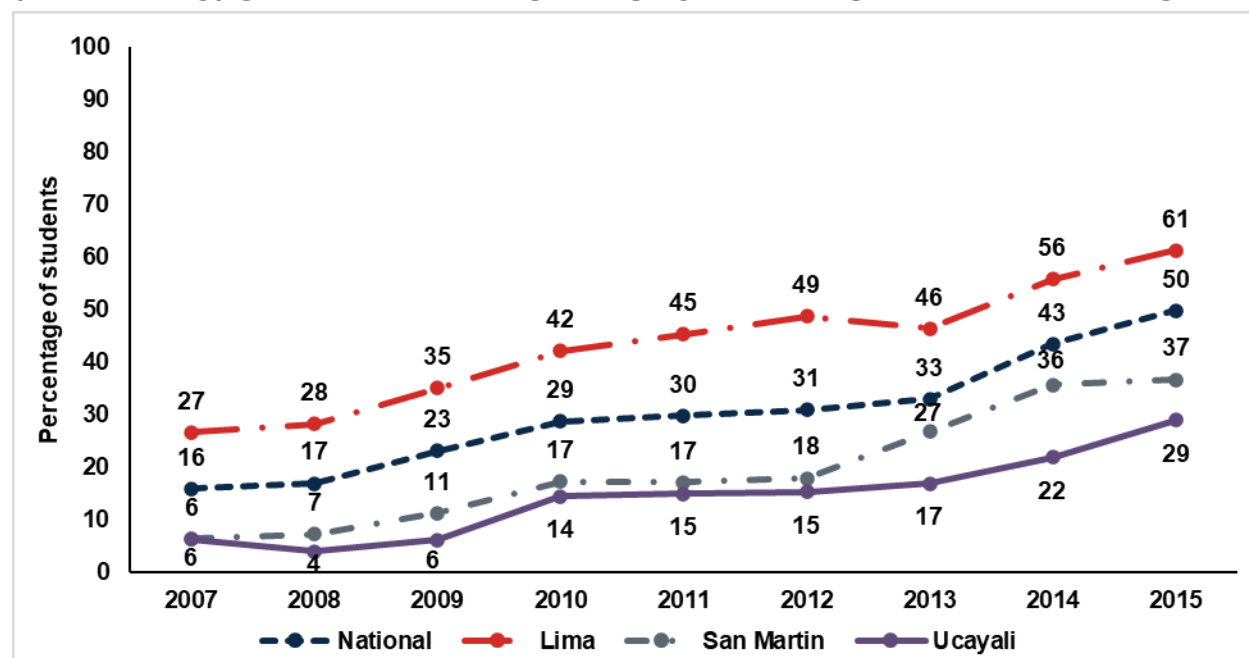
Teacher outcome data suggest a few early impacts on teachers' use of time and instructional method after a few months of teacher training. At the time of data collection in 2015, 80 percent of the teachers in the *Amazonía Lee* group had received training in both regions. In San Martín, slightly more than half of the teachers in the comparison group had received training (mostly from *Soporte Pedagógico*), whereas in Ucayali only a tenth of the comparison teachers had received training. Despite these differences in teacher training between groups, only two significant differences were found on time spent in teaching activities in San Martín and none in Ucayali.

This page has been left blank for double-sided copying.

I. INTRODUCTION

Despite remarkable progress in achieving the Millennium Development Goals of universal primary school completion and gender parity in education for LAC countries (United Nations 2015), important deficits remain in learning and access to education for poor and marginalized populations. In countries where access is limited, the problem is especially acute for remote regions. For example, within Peru, the Amazonian regions of San Martín and Ucayali lag behind the metropolitan area of Lima and other regions in Peru in average years of schooling as well as in learning achievement. According to Peru's national household survey, in 2014, residents of Peru's capital, Lima, had completed 12 years of education on average. The same year, residents of San Martín and Ucayali had completed 8 and 9 years, respectively (Instituto Nacional de Estadística e Informática 2014). On Peru's Student Evaluation Census (ECE, for its acronym in Spanish), which evaluates the reading comprehension and mathematics skills of students in 2nd grade, 56 percent of 2nd-grade students in Lima achieved the satisfactory level of reading comprehension in 2014, compared to 36 percent of students in San Martín, and 22 percent of students in Ucayali. Furthermore, the percentage of second grade students who achieved a satisfactory score on the reading comprehension exam in San Martín and Ucayali has historically been lower compared with the percentage of students at the national level and in Lima according to the ECE results.¹ (See Figure I.1)

Figure I.1. Percentage of second grade students who achieved the expected (satisfactory) grade-level reading category according to the ECE reading test



Source: Reading results from the evaluación censal de estudiantes, ECE. Ministry of Education. Office measuring the quality of learning.

¹ The Ministry of Education reports three levels of achievement for the ECE scores, unsatisfactory, in process, and satisfactory. The definition of satisfactory is that the student achieved the expected learning at the end of second grade and is prepared for the next grade. For the purpose of this evaluation, we refer to the satisfactory level as being at grade-level for second grade.

The geography of the two regions is shown in Figure I.2. San Martín lies in the northern part of the Peruvian Amazon forest; Ucayali is located south of San Martín, bordering Brazil. Both regions are primarily Spanish speaking with a small percentage of the population speaking Quechua, Aymara, or other indigenous languages. In terms of area, Ucayali is nearly twice as large as San Martín, but San Martín is more densely populated. San Martín has approximately 1,302 active primary public education institutions, whereas Ucayali has approximately 796.²

As part of the USAID/LAC Reads effort to strengthen reading capabilities in Latin America and the Caribbean by generating and using information on successful approaches to improving early grade reading, *Amazonía Lee* (Amazon Reads) was designed to improve early grade reading performance and address the learning gaps in these two regions of Peru.

This report presents the baseline analysis findings of an impact evaluation of *Amazonía Lee*, as well as early impact findings on teacher outcomes. The evaluation is one of several evaluations supported by the United States Agency for International Development (USAID), Regional Sustainability Office of the Bureau for Latin America and the Caribbean (LAC) under a contract with Mathematica Policy Research (Mathematica) for independent impact evaluations and analyses of the effectiveness and costs of promising reading and access to education interventions in the LAC region (LAC Reads Evaluation). Evaluations are underway in Guatemala, Peru, Nicaragua, El Salvador, and Honduras.

A. The *Amazonía Lee* intervention

The *Amazonía Lee* intervention aims to improve the quality of reading instruction and learning by children in the early grades of primary in two regions of Peru, San Martín and Ucayali. *Amazonía Lee* is implemented with funding from USAID by the Regional Directorates of Education (DREs, for their name in Spanish) of San Martín and Ucayali with technical assistance from the Universidad Peruana Cayetano Heredia (UPCH), beginning in 2015. In San Martín, the program is known as “*Programa Enseñar es Liderar*”, which means To Teach is to Lead. In Ucayali, the program is known as “*Liderando los Aprendizajes de Niñas y Niños 2014–2016*”, which means Leading the Learning of Girls and Boys.

Figure I.2. Map of study locations in Peru



² Active public primary schools in 2015. Source: ESCALE—Estadísticas de la Calidad Educativa. Ministerio de Educación Perú. Available at <http://escale.minedu.gob.pe/web/inicio/padron-de-iiiec>

The two *Amazonía Lee* programs share the main components: (1) capacity building which includes capacity building of teachers, school directors, and specialists and (2) regional component which includes regional assessment system to monitor students' achievement, a community engagement strategy to support student learning, and a system of incentives to recognize and motivate teachers' performance.

1. Capacity building component. The aim of this component is to strengthen pedagogical skills and instructional capabilities of teachers, school directors, and regional specialists. At the beginning of 2015, UPOCH started providing technical assistance to the DREs in both regions for the implementation of the following five capacity-building mechanisms (subcomponents):

- **Teacher training workshops based on *Amazonía Lee*'s balanced approach.** *Amazonía Lee*'s balanced approach to reading instruction integrates evidence-based foundational skills instruction with the Ministry of Education's communicative-textual approach. Training workshops for 1st- to 3rd-grade teachers aim to improve teachers' approach to reading instruction, class and curricular planning, differentiated instruction based on students' learning needs, and personalized attention for underperforming students. The training is structured in five modules: (1) print environment ; (2) understanding the alphabetic principle (alphabet knowledge, phonological awareness and fluency); (3) reading comprehension; (4) writing and production of text; and (5) leadership and interpersonal relations. Teachers are also trained on the design and use of formative assessments to monitor students' learning, on \ positive classroom climate strategies, and on community engagement strategies in support of students' literacy. Teachers are expected to receive 120 training hours in these workshops, divided in 40-hour blocks throughout the year.
- **Teacher coaching.** Specialists conduct classroom visits to coach teachers on-site and help them improve content area knowledge and pedagogical strategies for literacy instruction, differentiated instruction, assessment and monitoring of student learning, and effective classroom management. These coaches will build teacher capacity through individual demonstration and observation sessions. In demonstration sessions, the coach will serve as a role model for effective classroom practice and demonstrate in the teacher's classroom the teaching strategies introduced during the training workshops. Discussion sessions will be accompanied by discussion of what the teacher learned while observing the coach deliver the lesson. Observation sessions complement the demonstrations. In observations sessions, the coach observes the teacher deliver a live class and offer feedback on areas for improvement.
- **Teacher study groups.** Teachers gather in small groups of 8 to 16 participants to strengthen their learning, review pedagogical materials, share teaching strategies, and plan for classes with support from other teachers. Study groups (or "Grupos de Interaprendizaje") do not have a defined set of contents; rather, they are meant to be flexible spaces for "collective feedback". Coaches set the agenda for the study groups in collaboration with teachers. The content is meant to be responsive to teachers' needs and interests and to the skill-building needs coaches identify during the coaching visits. The goal is for the teacher study groups to become autonomous over time, and the groups are intended to promote teachers' pursuit of continuous professional development in

collaboration with their peers. Teacher study groups are expected to meet once a month for approximately four hours. Teacher groups receive six coaching visits a year, but the frequency can vary depending on teachers' needs.

- **More and more varied pedagogical materials to support reading.** The goal of this component is to improve quality and regional pertinence of pedagogical materials for reading and writing instruction, and increase the availability of classroom book libraries in primary grades. Classroom materials and teacher guides incorporate evidence-based foundational reading skills instruction into the national pedagogical communicative approach promoted by the Ministry of Education in Peru. Materials also emphasize culturally relevant stories and activities for students in the Amazon region. Planned materials include teacher training modules, teacher activity guides and student workbooks, and classroom libraries.
- **School leadership and management training.** School directors will receive training on pedagogical leadership and school management and support from specialists on issues related to pedagogical matters and institutional management. Regional governments receive technical assistance to improve school planning and budgeting, and execution, and to implement information systems for school management indicators.

2. Regional component. In addition to the five capacity-building mechanisms described above, *Amazonía Lee* includes another three activities envisioned to be implemented at the regional level in both San Martín and Ucayali.

- **Community engagement.** The community engagement component aims to promote the active engagement of the educational community in support of student learning, particularly in gaining reading proficiency during the first three grades of primary education. This component focuses on engaging parents and the community at large in promoting reading opportunities for children in and outside the school context through a variety of activities such as book fairs, family oral storytelling, and community workshops. This component also includes elaborating plans for the timely identification of students who may be at risk of academic failure.
- **Regional assessment system.** The goals of this subcomponent are (1) to introduce regional assessments at the beginning and end of the school year to monitor student achievement (on reading comprehension and production of texts); and (2) to develop local capacity in the design, implementation, scoring, data production and interpretation of these assessments.. The DREs of San Martín and Ucayali will lead the design of assessments following the format of the ECE. This component also includes the administration of assessments and the dissemination of assessment results to stakeholders.
- **Regional incentive system for teachers.** This subcomponent has regional policies to motivate, encourage, recognize, and reward teacher performance—thereby engaging all stakeholders of the education system (teachers, specialists or facilitators, school directors, students, and others) in efforts to improve student learning and academic achievement. UPEL will provide technical assistance in the design of a systems of incentives to recognize excellence in teaching (Sistema de Reconocimiento de Buenas Prácticas, SRBPD for its initials in Spanish). Since 2013, The Ministry of Education in

Peru organizes a national contest to recognize excellence in teaching known as “Concurso Nacional de Buenas Prácticas Docentes”. UPCH will use the framework of the national contest as the foundation for the design of the regional system of non-monetary incentives. Initially, the incentive system will be implemented only in schools participating in *Amazonía Lee*. To participate in the SRBPD, teachers must be teaching first, second or third grade (by contract or tenured), and must be continuously engaged in the teacher training activities of *Amazonía Lee*.

In addition to the regional-level efforts described above, UPCH will provide technical assistance to raise awareness and obtain support from regional governments, the Ministry of Education, the private sector, and civil society on efforts to promote, literacy, gender equity, and ongoing assessment of student achievement. Also, UPCH will engage the private sector and teacher training institutions to increase the number of specialists trained in the “balanced approach” to literacy instruction.

B. Evaluation questions, design, and limitations

The impact evaluation, of which this baseline data is a first step, is structured to address the following central evaluation question:

- What was the impact of *Amazonía Lee* programs on early-grade reading and other intermediate outcomes relative to other prevailing practices and programs?

Additionally, it will address the below questions related to program implementation:

- How were *Amazonía Lee* intervention components, activities, and processes implemented in Ucayali and San Martín?
- What services were provided to control schools in Ucayali and San Martín?
- How were *Amazonía Lee* components, activities, and impacts perceived by principals and teachers?
- What were the facilitators and barriers to the implementation of *Amazonía Lee* components?

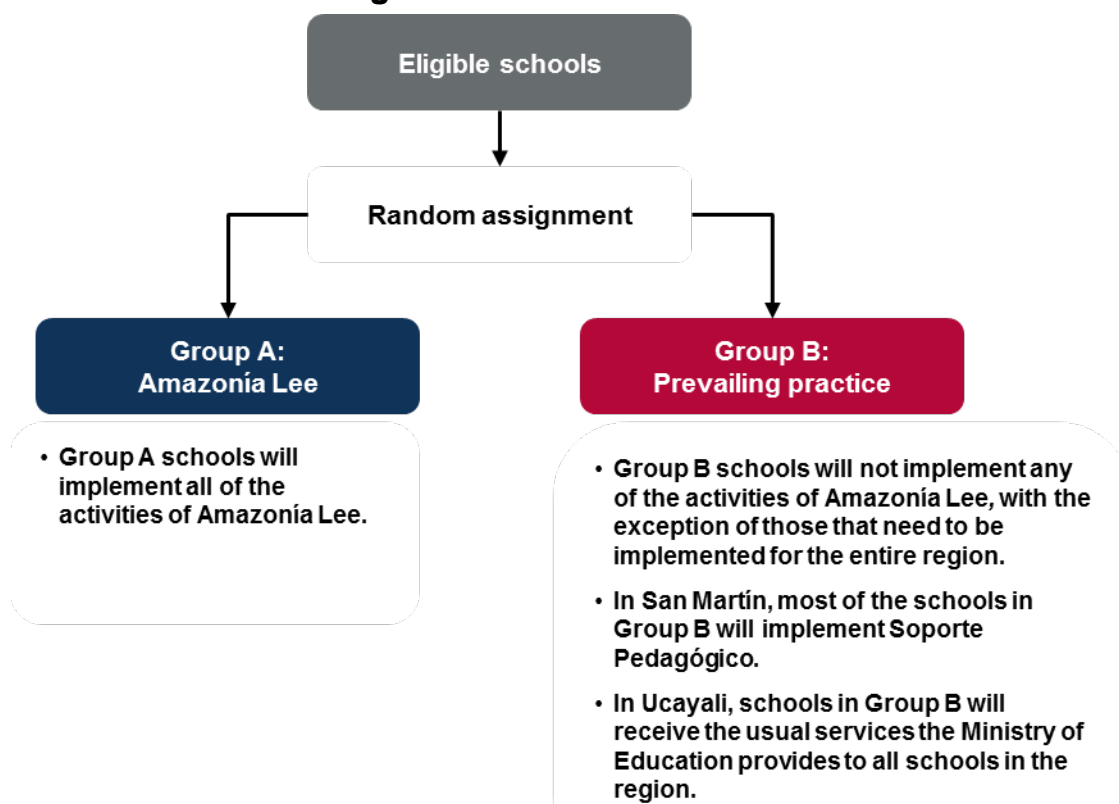
The central questions of the baseline data collection are:

- What are the characteristics of the study schools and teachers in 2014 (the year before the intervention began)?
- What are the initial characteristics and early reading assessment of first grade students in 2015 who will be followed by the evaluation?
- To what extent did random assignment produce equivalent groups based on measurable characteristics for schools and teachers in 2014 and for students in 2015?
- What are the early impact findings of *Amazonía Lee* on teacher outcomes in 2015 such as participation in training and approaches to reading instruction?

Impact evaluation design. Mathematica is working with USAID/Washington, USAID/Peru, and the Ministry of Education in Peru, and especially with the regional offices of

San Martín and Ucayali, to conduct a rigorous impact evaluation of *Amazonía Lee*. The impact evaluation employs a randomized controlled trial, the most rigorous evaluative framework for assessing program impacts. We randomly assigned eligible schools in the targeted regions of San Martín and Ucayali to one of two groups: group A was assigned to receive all of the components of *Amazonía Lee* and group B was assigned to receive the services that the Ministry of Education already provides in each region (“prevailing practice”). Both groups will receive the components or activities of *Amazonía Lee* that must be implemented for the whole region (Figure I.3).

Figure I.3. Evaluation design: randomized controlled trial



In each region, we randomized eligible schools into one of the two groups described above. In San Martín, we randomly assigned 100 schools to each of the two groups in February 2014. In Ucayali, we randomly assigned 35 schools to each of the two groups in September 2014. In both regions, we conducted stratified random assignment to ensure that the two groups would be similar on key characteristics: *Unidad de Gestión Educativa Local* (UGEL), school size, school structure (separate or combined grades), bilingual population, implementation of the Strategic Learning Achievements Program (PELA), and ECE reading scores. We provided full details on random assignment in our evaluation plan (Murray et al. 2016).

The randomized evaluation relies on comparing outcomes between these two groups of schools in each region to estimate the impacts of differences in activities in the schools on learning outcomes. Because the Ministry of Education is providing different services to the schools assigned to group B in each of the two regions, the main evaluation questions differ by

region. In San Martín, a majority of the group B schools will be offered *Soporte Pedagógico*³ (Pedagogical Support), a new initiative from the Ministry of Education whose goal is to improve student learning in urban elementary schools, and which began in 2015. *Soporte Pedagógico*'s main components are training workshops, teachers' peer-to-peer study groups, and virtual educational support for teachers and principals; implementation of learning sessions designed to develop students' communication, math, and social studies skills; efficient mechanisms for timely allocation and use of appropriate teaching materials and educational resources; and capacity building of principals, teachers, and parents to ensure adequate learning conditions for students. In addition, *Soporte Pedagógico* uses teachers who work after regular school hours (*docentes fortaleza*) to provide remedial support to students who are falling behind in achieving expected learning goals. Thus, for San Martín, we will be comparing outcomes for schools receiving the *Amazonía Lee* intervention to schools receiving *Soporte Pedagógico*. As the two programs have similar aims and service components, the main evaluation question in San Martín can be understood as:

- What is the impact on early grade reading and other outcomes of the components of *Amazonía Lee* that can be implemented at the school level, as compared to the impact of *Soporte Pedagógico*?

In Ucayali, schools are not receiving *Soporte Pedagógico*; therefore, the main evaluation question is:

- What is the impact on early grade reading and other outcomes of the components of *Amazonía Lee* that can be implemented at the school level, as compared to prevailing practice?

Timing of baseline data collection. In our original evaluation design, the target student population comprised 1st graders who were enrolled in first grade in the study schools at the time of random assignment—school year 2014 (2014 cohort). We collected baseline data for these students in 2014. However, *Amazonía Lee* implementation did not start during the 2014 school year as initially planned but was delayed until 2015. Therefore, the target student population that the evaluation focused on changed to students enrolled in 1st grade in 2015 (2015 cohort). We collected baseline data for this new cohort in 2015.

Impact estimation. *Amazonía Lee* is designed to be implemented beginning in the 1st grade, when students begin acquiring foundational literacy skills, and to continue through the 3rd grade when students should have acquired basic reading skills. While the evaluation was originally designed to follow a cohort of students that were in 1st grade in 2014 for three years, implementation delays required us to follow the 2015 cohort of 1st graders instead. *Amazonía Lee* impacts will be assessed using data from the 2015 student cohort that will be collected at the end of the 2016 school year when most of the cohort's students should be in 2nd grade and should have been exposed to the intervention for almost two school years. This decision was made in

³ *Soporte Pedagógico* targets urban areas. *Soporte Pedagógico Intercultural* (SPI) targets bilingual schools located primarily in rural areas. While both programs aim to improve the quality of education and student achievement, SPI emphasizes the implementation of the Ministry of Education's bilingual intercultural education model with strong bilingual literacy acquisition (in mother tongue and in Spanish) and cultural diversity components.

coordination with all evaluation stakeholders and responded to the program logic. The evaluation design also takes into account cost considerations that prevent us from collecting endline data on both the 2014 and 2015 cohorts. The final evaluation timeline plans for impact data to be collected at the end of 2016 which will allow us to present evaluation results in the third quarter of 2017.

Student impacts. In each region, we will estimate the impacts of *Amazonía Lee* on 2nd-grade students' reading skills by comparing the average outcomes of cohort 2015 students in schools in group A and group B, controlling for chance differences in characteristics of schools, teachers, and students before the intervention started (2014). In addition, we will also control for any chance differences in the 2015 first grade cohort's characteristics and test scores observed at baseline. In this report, we discuss the baseline differences found between the study groups for school, teacher, and student characteristics in 2014 and student characteristics in 2015. The group means and differences between groups presented in this report are based on regression models that adjust for the stratified randomization used.

Timing of implementation and baseline analysis. *Amazonía Lee* began implementation in 2015. In San Martín, the first teacher training was conducted in April and the second in August. In Ucayali, the first teacher training was conducted in March and the second in June and August. We might expect to see the early effects of the intervention in the 2015 teacher survey data, since the first teacher training was conducted two to three months before data collection; however, it is unlikely that the intervention had an effect on student reading skills in such a short time. When interpreting the data collected for either year, we will be careful to take into consideration the timing of data collection relative to the intervention.

The baseline analysis will use the data from the 2014 cohort to check whether schools and teachers in groups A and B were similar to each other before the program began in each region. Student data from 2015 will be used to check whether students were similar across study groups in terms of characteristics such as gender, age appropriate for grade, and early reading skills at the beginning of 2015 in our student cohort of interest. We consider this an appropriate measure of baseline equivalence for students because we believe it is unlikely that the intervention could have had an impact on learning within a few months. However, teachers could change their instructions practices within a few months of teacher training. Hence, teacher survey data collected in 2015 are used to assess early impacts of *Amazonía Lee* on access to training and on reading instruction.

Limitation. All study schools are expected to receive the regional component of *Amazonía Lee* (e.g. regional assessments, technical assistance on the use of regional student assessments, parent engagement activities, teacher incentives, leadership training for regional authorities, regional or national campaigns to raise awareness of the importance of literacy education). Because schools in groups A and B are likely to receive all the regional-level subcomponents of *Amazonía Lee*, we don't expect those subcomponents to generate differences between the study groups, and we will not be able to estimate their impacts. Subcomponents implemented at the school level may generate differences between schools in groups A and B; we will estimate the impacts of those subcomponents. The capacity building component of *Amazonía Lee* that will only be made available to schools in group A include the pedagogical materials developed specifically for *Amazonía Lee* schools, teacher training workshops, in-class teacher coaching,

teacher study groups, and leadership training for principals. Schools in group B may have access to similar activities through their participation in other programs, such as *Soporte Pedagógico*, but will not have access to the specific activities designed for the *Amazonía Lee* program. We will gather information on services that study schools received from *Amazonía Lee* or other programs through surveys. Also, we will collect qualitative information from teachers, school directors and regional leadership authorities on the implementation of *Amazonía Lee* activities.

C. Data sources and indicators

1. Data collection instruments

The main data sources for the impact evaluation are student reading assessments, classroom observations, a teacher survey, a school infrastructure survey and enrollment records⁴.

Student assessments. In each study school, one 1st-grade classroom was selected at random for data collection. In each selected classroom, we selected a sample of 1st-grade students at random to participate in the emergent literacy assessment (the sample included 8 students per classroom in San Martín and 10 students per classroom in Ucayali). We first assessed students' oral language proficiency using a screening task based on the PreLas (DeAvila and Duncan 2000)⁵, which included two subtasks: receptive language proficiency and expressive language proficiency. In each subtask, children were asked to follow 10 instructions that are read out loud. The purpose of assessing the students' oral language proficiency was to verify that children were capable of completing assessments of early literacy skills in the language of instruction reported by the teacher. The student language screener also served as a warm-up procedure to build rapport and make children feel comfortable in the testing situation before the assessment of early literacy skills that would come next. After assessing the students' oral language proficiency, we assessed their early literacy skills. We used an instrument designed with a focus on key literacy skills that predict reading comprehension in later grades: alphabet knowledge, phonological awareness/phonics, decoding, listening comprehension, oral reading, and writing. The instrument consisted of selected subtasks adapted from two instruments that have been used with dual-language learners in Latin America⁶: (1) the Emergent Literacy Profile (La Evaluación de Lectoescritura Emergente—LEE [Dickinson et al. 2004]); and (2) the Early Grade Reading Assessment (EGRA, adapted for Guatemala as Evaluación de Lectura para Grados Iniciales [ELGI], [USAID/Reforma Educativa en el Aula 2013]). Figure I.4 shows the content of each instrument used in the emergent literacy assessment.

⁴ Survey instruments for all rounds of data collection are available along with restricted use data files on the USAID Development Data Library.

⁵ If the result of the PreLas indicated that a child did not have sufficient oral language skills in Spanish to take the child assessment, we assessed a different child from the class. If the replacement student also did not pass the PreLas test, we did not look for a second replacement student. All assessments were conducted in Spanish.

⁶ The assessment protocol draws from the experience of the USAID-funded *Reforma Educativa en el Aula* (REAULA) project in adapting Early Grade Reading Assessment (EGRA) instruments for use in Guatemala. We also followed REAULA's experience and recommendations for minimizing the risk of obtaining floor or ceiling effects that could have arisen if the early grade reading assessments were not calibrated at the right difficulty level for our target population. We pre-tested the items in Peru and made adjustments so they were culturally appropriate to the population.

Figure I.4. Content of 1st-grade emergent literacy assessment

Language verification (PreLas)	• Receptive language: “The Grandfather Says” • Productive language: “The Human Body”
LEE	• Alphabet knowledge • Simple word reading (emerging) • Writing (emerging)
ELGI	• Phonemic awareness • Decoding • Oral comprehension

LEE = Evaluación Lectoescritura Emergente (Dickinson et al. 2004)

ELGI = Evaluación de Lectura para Grados Iniciales (USAID/REAULA 2013)

Classroom observation. We developed a classroom observation instrument to be fielded in the Reading and Communication class.⁷ The instrument focused on measuring teachers’ use of class time on different types of literacy-related activities, teachers’ literacy instruction (for example, book reading, fluency, vocabulary development, reading comprehension), teachers use of general pedagogical practices (for example, engage students in class activities, check for understanding, give feedback, and promote positive classroom climate). In addition, the observation protocol measured the quality of the classroom literacy environment (for example, availability of books and print materials, variety and overall quality of books and materials, exposure to print in classroom walls with labeled objects, familiar words, student’s names, bulletin boards) and availability of resources for writing. The main teacher outcomes measured through direct observation of a Reading and Communication class were as follows: time spent on literacy-related activities (time on literacy tasks), use of foundational reading instruction practices, and availability of a print-rich classroom environment.

Teacher survey. Teacher surveys gathered self-reported use of instructional practices, proportion of students in the classroom that achieved milestones in reading acquisition, use of formative assessments, participation in professional development or in-service training, education and experience, occupational needs, career expectations, and background characteristics. The main teacher outcomes measured in the teacher survey were as follows: participation in teacher training (amount and duration) and topics learned during teacher training. The teacher survey was administered to all teachers whose classrooms were observed.

School infrastructure checklist. The school infrastructure checklist gathered information on general characteristics of the schools’ infrastructure such as the number of classrooms, the distance between the school and the district capital, and the availability of electricity and potable water. To complete the infrastructure checklist, enumerators conducted a visual scan of the school premises and buildings and asked a few questions of the school principals.

⁷ In Peru, they refer to this class as Language and Communication.

Enrollment records. Classroom enrollment lists provided by the DRE in each region were used to conduct the random selection of students for reading assessments. These records provided information on students' age, sex, and disability.

2. Timing of data collection.

As Table I.1 shows, we gathered baseline data in 2014 and 2015 for two different cohorts of 1st-grade students and their teachers, but in the same schools. In 2014, we gathered baseline data on schools, teachers, and 1st-grade students, as well as a classroom observation. We refer to this as the 2014 cohort. In 2015, we collected data on a new cohort of 1st-grade students and their teachers—the 2015 cohort.

Table I.1. Timing of baseline data collection

Year	2014		2015	
Semester	1st	2nd	1st	2nd
2014 cohort		Baseline (1st grade)		
Individual emergent literacy assessment		X		
Classroom observation		X		
Teacher survey		X		
School infrastructure checklist		X		
Enrollment records		X		
2015 Cohort			Baseline (1st grade)	
Individual reading assessment			X	
Classroom observation				
Teacher survey			X	
School infrastructure survey				
Enrollment records			X	

Note: The assessments were individually administered. The 2014 cohort of students was assessed in June and July 2014 in San Martín and in September and October 2014 in Ucayali. The 2015 cohort of students was assessed between May and July 2015 in San Martín and in June and July 2015 in Ucayali.

This page has been left blank for double-sided copying.

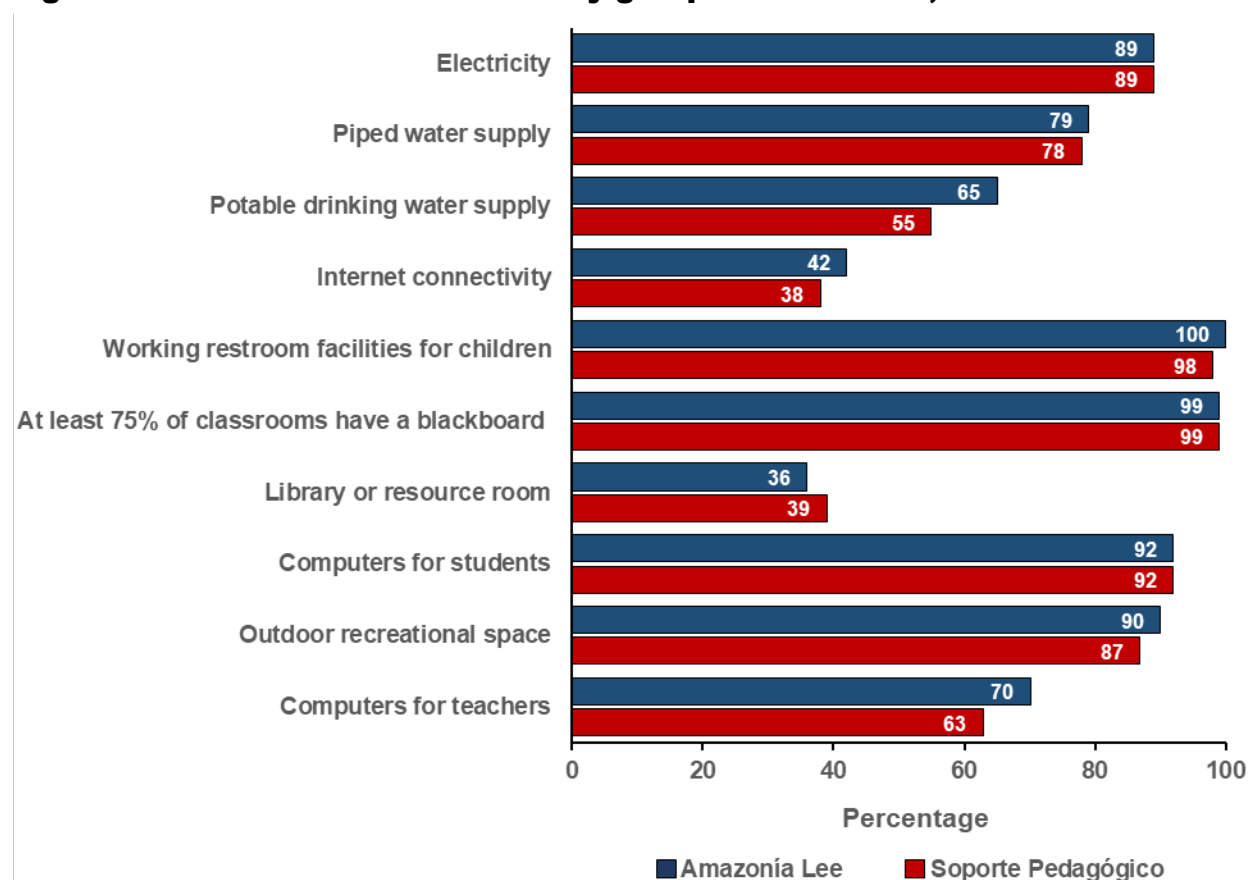
II. FINDINGS: SAN MARTÍN: SCHOOLS, TEACHERS, AND STUDENTS AT BASELINE

This chapter presents the findings for San Martín from school and class observations and teacher surveys collected in 2014, the year before the intervention began, and from the student early reading assessment collected in 2015. We compare the characteristics of the schools, teachers, and students in group A, the *Amazonía Lee* group, to those in group B, the *Soporte Pedagógico* group.

A. School characteristics

Amazonía Lee schools had similar facilities to *Soporte Pedagógico* schools. We did not find any statistically significant differences in school facilities between the groups. However, we found a difference of 10 percentage points in potable drinking water and of 7 percentage points in computers for teachers, both of them in favor of the *Amazonía Lee* group. Although these differences are not statistically significant, they could affect student outcomes. Therefore, we will account for them in our final impact analysis in order to improve the precision of the impact estimates. Figure II.1 shows school characteristics by group. Appendix A.1 provides more information on school characteristics.

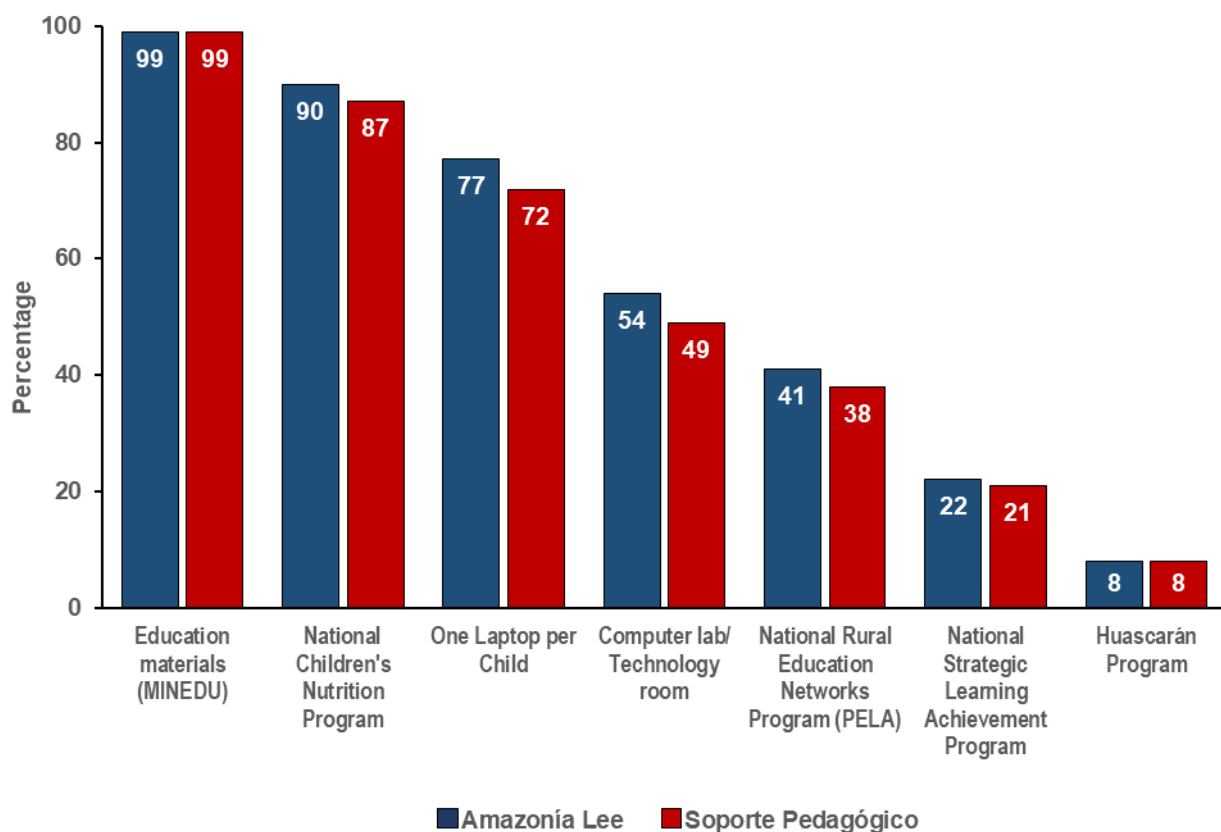
Figure II.1. School characteristics by group in San Martín, 2014



Source: 2014 school infrastructure observation, San Martín.

The study programs do not operate in a vacuum. Many schools participated in various support programs that provide school nutrition, education technology, and other inputs to learning. These are summarized in Figure II.2. The differences in program participation between study groups are small and not statistically significant.

Figure II.2. Support programs in study schools by group in San Martín, 2014



Source: 2014 school infrastructure observation, San Martín.

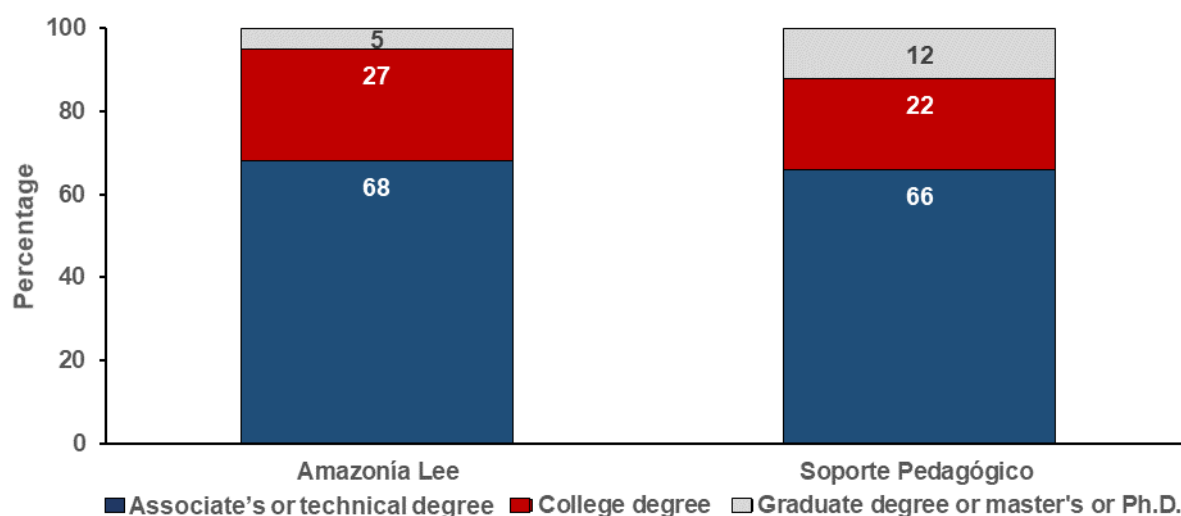
B. Teacher characteristics

Teachers in both groups had similar experience and demographic characteristics. In 2014, we did not find statistically significant differences between groups on teacher experience, tenure, age, or gender. Appendix B.1 presents the results. In both groups, the average teacher in San Martín was experienced, with 16 years of previous teaching experience and 9 years at his or her current school. In both groups, roughly two-thirds of teachers were not in their first year at their current school. Approximately two-thirds of teachers in both groups were tenured; the remaining one-third of teachers had annual contracts. Approximately three-fourths of teachers surveyed in both years were women. On average, these teachers were in their early forties. In both groups, teachers reported missing one day of work on average in the last month for illness or other reasons.

Teachers in the *Amazonía Lee* group were less likely to have advanced degrees beyond undergraduate education than teachers in the *Soporte Pedagógico* group. The percentage of

teachers with an associate's or technical degree was similar between groups, but a greater proportion of teachers with higher education had a graduate degree in the *Soporte Pedagógico* group (12 versus 5). The difference in the distribution of educational attainment was statistically significant at the 5 percent level. These results are shown in Figure II.3.

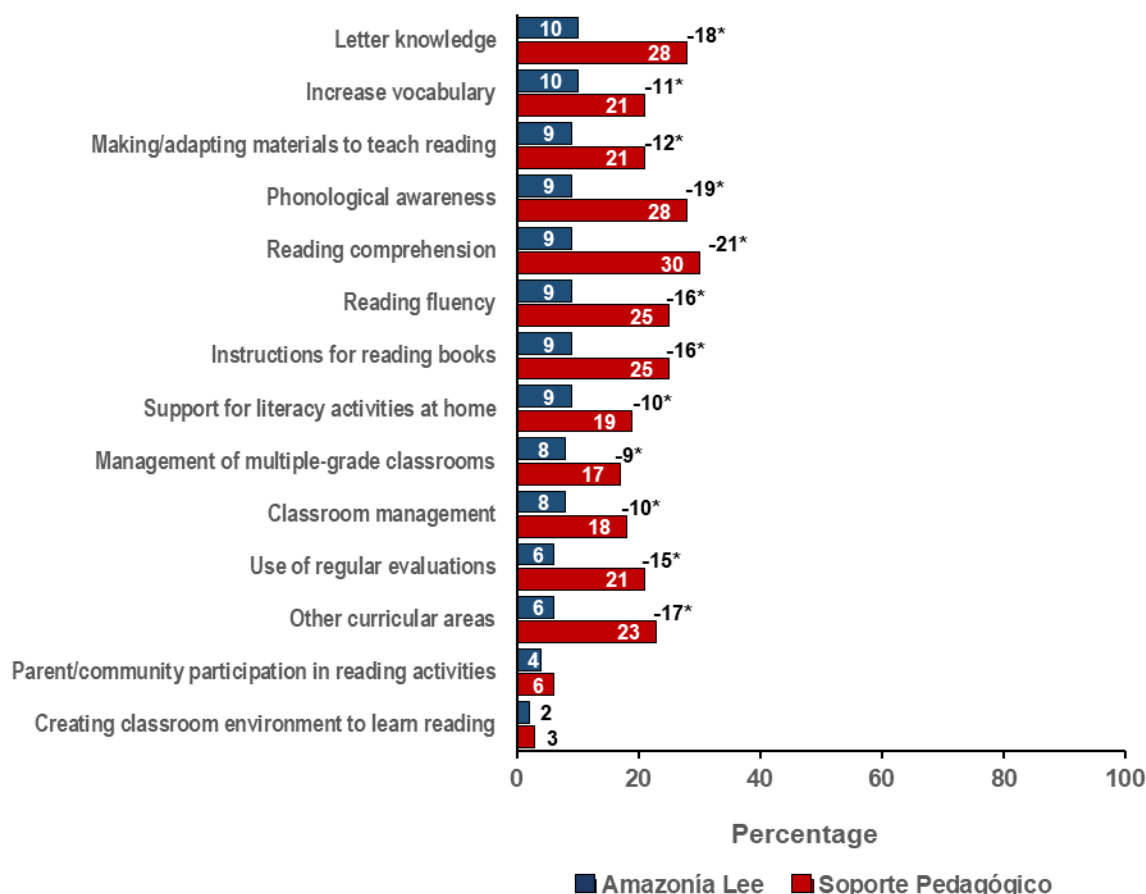
Figure II.3. Teachers' educational attainment by group, 2014



Source: 2014 1st-grade teacher surveys, San Martín.

Note: Differences in the distribution of educational levels were statistically significant at the 5 percent level.

***Amazonía Lee* teachers were less likely to have participated in in-service reading instruction training in the past year than teachers in *Soporte Pedagógico* schools.** In 2014, before implementation of either program began, teachers at 12 percent of *Amazonía Lee* schools had received in-service training in reading instruction in the previous year compared to 32 percent at *Soporte Pedagógico* schools (a statistically significant difference) and had received fewer hours of training on average (5 hours compared to 14). Figure II.4 shows participation in training by subject for each group. Training was most likely to cover reading comprehension, letter knowledge, and phonological awareness. Appendix B.2 shows details on factors teachers identified as limiting their ability to teach students to read.

Figure II.4. Training content areas by group, 2014

Source: 2014 1st-grade teacher surveys, San Martín.

Note: Differences that are statistically significant at the 5 percent level are labeled next to each pair of bars and marked with an asterisk.

C. Teaching materials and teaching practices

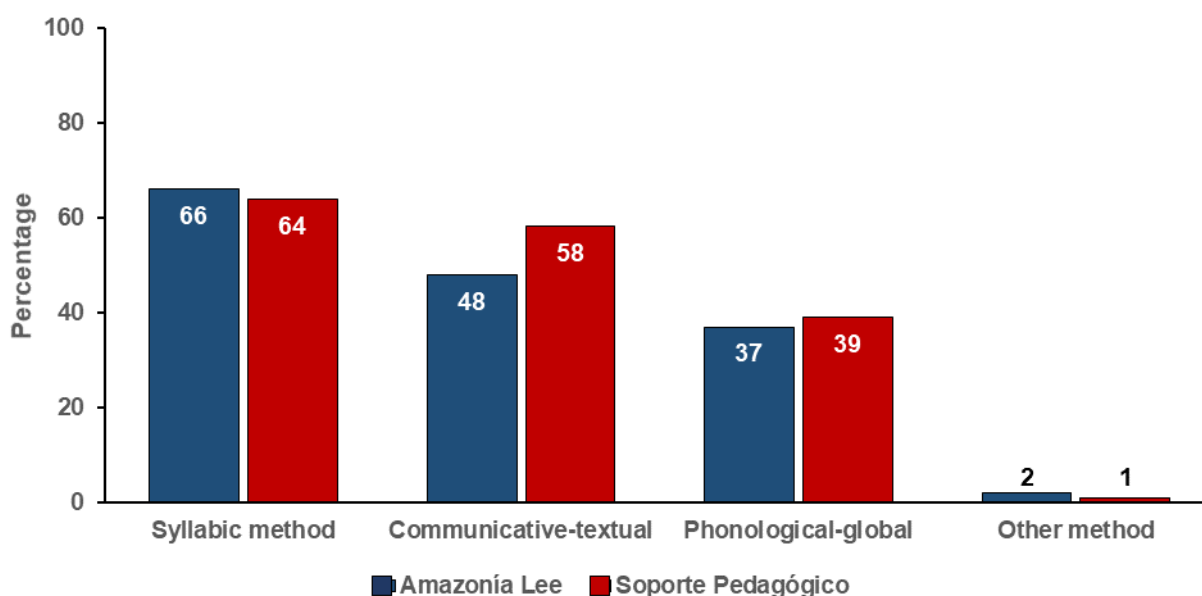
In combination with teacher surveys, classroom observations (conducted in 2014 only) allow us to describe the learning environment and teaching approaches used in study schools before the intervention began, and identify any differences between the groups.

Teachers reported using one or more of three main pedagogical approaches: **syllabic, communicative-textual, and phonological-global**. In Peru, the syllabic method has been used historically for reading instruction in Spanish; it focuses on the progressive learning of vowels, sounds, and consonant-vowel combinations building into words. The communicative-textual approach to early grade reading instruction is the national approach currently supported by the Ministry of Education. It is a two-pronged approach that focuses on developing students' communicative competencies in meaningful contexts, and their interactions with different types of texts that respond to a variety of communicative situations. The phonological-global method integrates whole language with phonological analysis, emphasizing global word recognition and

understanding of the correspondence between letters and their sounds. This method typically promotes a range of foundational reading skills like alphabet knowledge, phonological awareness, decoding, vocabulary, fluency, and comprehension.

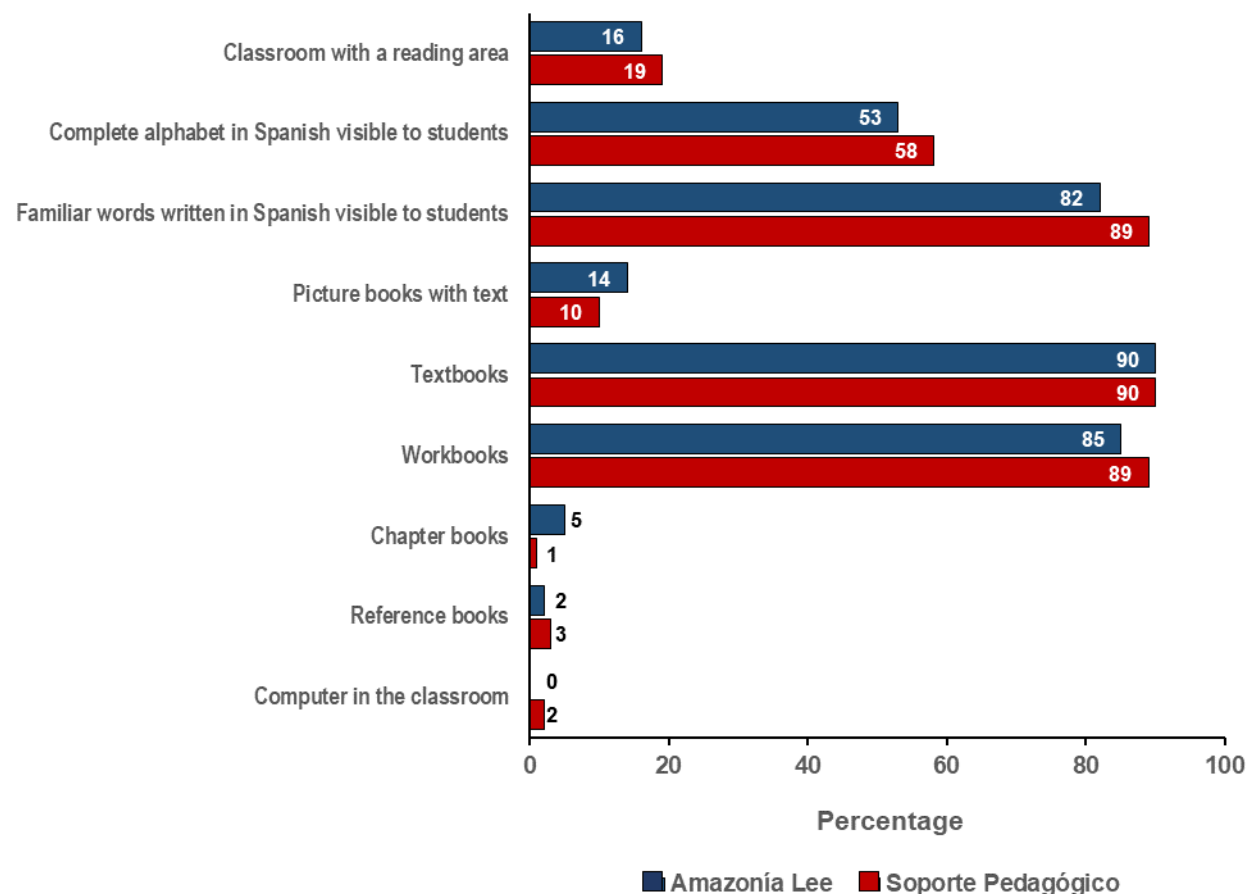
There were no differences between groups on the pedagogical approach teachers reported using for reading instruction. In 2014, approximately 65 percent of teachers reported using the syllabic method and 38 percent reported using the phonological-global method. There were no statistically significant differences between groups on the methods that teachers reported using. Despite more teachers reporting using the communicative-textual approach in the *Soporte Pedagógico* group than in the *Amazonía Lee* group (58 percent versus 48), the difference was not statistically significant.

Figure II.5. Reading instruction approaches teachers reported using by group, 2014



Source: 2014 1st-grade teacher surveys, San Martín.

Availability of reading materials was similar in the two groups. We did not find statistically significant differences in availability of reading materials between the groups. Most schools had familiar words in Spanish on display and roughly half had the full Spanish alphabet on the wall. Less than 20 percent of schools had a reading area and only two schools had computers in the classroom (see Figure II.6). Classrooms observed had books available for students, but these were predominantly textbooks or workbooks. Other types of materials such as picture books and reference books were less common.

Figure II.6. Reading materials available in classrooms by group in San Martín, 2014

Source: 2014 1st-grade classroom observations, San Martín.

Note: None of the differences between groups were significant at the 5 percent level.

Teachers were observed using diverse activities to teach their students to read and write. Over 70 percent of teachers were observed providing instruction on writing words, encouraging students to recognize and identify letters, engaging students in fluency-building activities, and providing explicit instructions for phonological and phonemic awareness. We present detailed results of reading instruction activities in Appendix C.1.

Teachers in both groups used a similar set of reading instruction activities overall. We did not find significant differences in teachers' likelihood of using 14 out of 15 reading activities in the class observation. However, *Amazonía Lee* teachers were significantly more likely to use oral reading and fluency-building exercises at baseline. This difference was significant, but on average, teachers in both groups spent less than 3 minutes per class on fluency-building activities.

Teachers in the two study groups divided their reading instruction time in similar ways, generally spending more time on phonological and phonemic awareness than on other activities. Teachers in both groups spent roughly two-thirds of their time on phonological

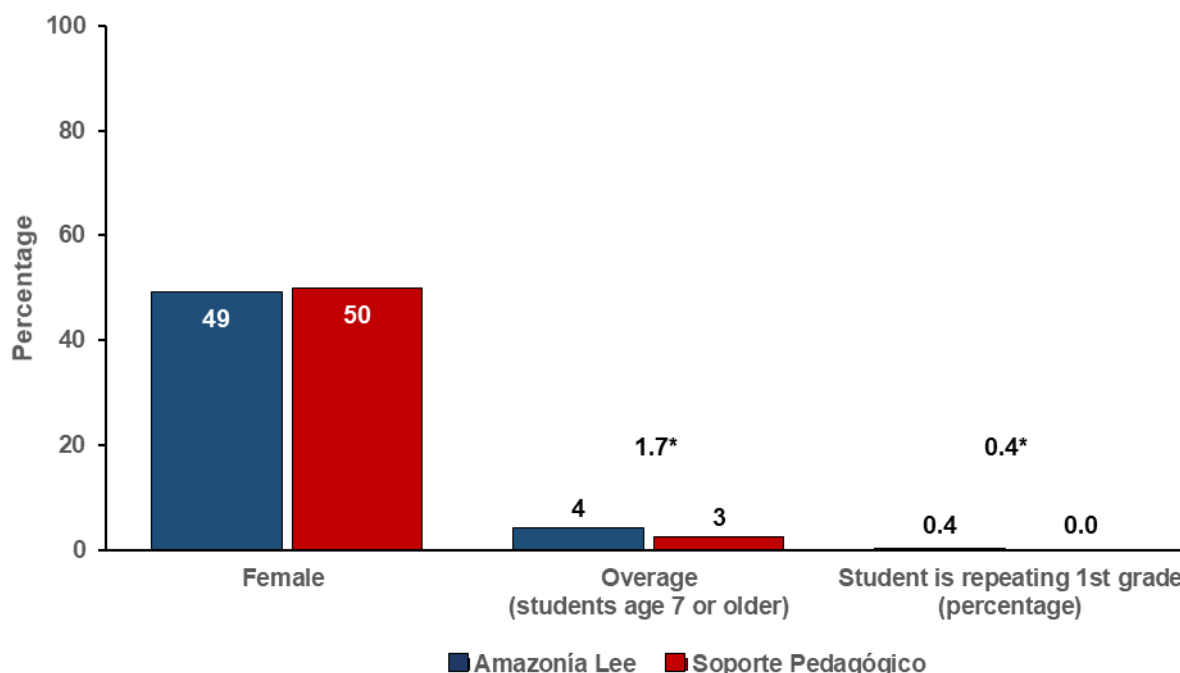
and phonemic awareness. Of roughly 38 minutes of reading instruction time observed, teachers in both groups spent most of their time on phonological and phonemic awareness activities. This finding is expected given that at the beginning of first grade, when this data was collected, teachers spend most of their time teaching emergent reading skills such as phonological and phonemic awareness. Detailed results on time use are provided in Appendix C.2.

Teachers in both groups spent most of their time teaching the whole class and almost no time teaching small groups. We did not find statistically significant differences between the groups on how teachers divided their time among working with the whole class, individual students, and small groups. Teachers spent just over half of the time observed teaching the entire class; in most of the remaining time, they worked with individual students. Teachers spent little time working with small groups or on activities not related to reading; differences between groups were small and not statistically significant (results not shown)

D. Student characteristics and emergent literacy assessment

Baseline data for the 2015 cohort of 1st-grade students were drawn from enrollment records as well as the emergent literacy assessments in 2015. The 2015 student cohort is the cohort that will be tracked to evaluate the impact of the *Amazonía Lee* intervention at the end of 2nd grade. We will use the 2015 baseline student characteristics and reading assessment scores as control variables for the sample of children that we follow in the analysis of learning impact. Data on the 2014 cohort are presented in Appendix D.1–D.3 to provide context on the characteristics of students in the study schools and student levels of achievement in the year before the intervention began.

In the 2015 1st grade cohort, *Amazonía Lee* students were more likely to be repeating 1st grade and to be overage than *Soporte Pedagógico* students. We found that in 2015, students in the *Amazonía Lee* group were 2 percentage points more likely to be old for their grade (age 7 or older in 1st grade) and 0.4 percentage points more likely to be repeating a grade. Although these differences are statistically significant, they are small in magnitude. In both groups, approximately half the students were girls. Because these differences would have not resulted from exposure to the intervention but could affect the final outcome, we will control for them when conducting our impact analysis. Figure II.7 shows student characteristics by group in 2015.

Figure II.7. Student characteristics by group, 2015

Source: 2015 school enrollment administrative records reported in the student assessment, San Martín.

Note: The sample is restricted to include students who agreed to participate in the emergent literacy assessment.

Differences between groups that are significant at the 5 percent level are above the bars and marked with an asterisk.

In both groups, Spanish is the dominant language. There was no significant difference between groups in the likelihood of the students or their families speaking a language other than Spanish. There was also no significant difference between groups in the likelihood that Spanish was the language of instruction. In both groups, just over 1 percent of students reported that their families spoke a language other than Spanish at all, and just under 1 percent of students reported speaking a language other than Spanish themselves. The language of instruction was Spanish in all but about 3 percent of schools (results not shown).

Most students agreed to participate in the emergent literacy assessment, and participation was similar across groups. In 2015, 91 percent of students selected for the study passed the language screener and attempted the emergent literacy assessment. However, some of the students that attempted the assessment got a score of zero in assessment tasks that required more advanced skills, like decoding. We present the results of the emergent literacy assessment in Table II.1. In Figure II.8, we show the percentage of students who had a score of zero for each task in the emergent literacy assessment.

We found no statistically significant differences between groups in the emergent reading skills assessed. In 2015, first grade students in both groups had similar skills on letter identification, phonemic awareness, decoding pseudo-words, emergent reading and writing, and passage comprehension (Table II. 1). Below, we describe in more detail the findings for each

skill. We also did not find significant differences between groups in emergent reading skills in the 2014 cohort (Appendix D.3).

Students could identify most but not all letters. For this activity, a student received two points for each of 12 letters he or she identified correctly when shown a card with four printed letters. If the student did not identify a letter, the evaluator asked the student to find the letter on the card, saying, for example, “Where is the letter *m*?” If the child found the *m* on the card, he or she received one point for the letter. Students in the *Amazonía Lee* group and the *Soporte Pedagógico* group scored roughly 18 out of 24 points on letter identification. No statistically significant difference found (Table II.1). Less than 2 percent of students were unable to identify any letters (Figure II.8).

Students had some emergent writing skills but had not yet mastered skills like writing their own name or simple words correctly. Students were asked to write their own name, to write two simple words of their own choosing, and to write three simple words, with two or three syllables, dictated by the evaluator. Students lost points if they wrote nonexistent words, misspelled words, wrote in a direction other than from left to right, or held their pencil incorrectly. Students in both groups had on average a score of almost 12 out of 22 on emerging writing skills (Table II. 1). No statistically significant difference found. Less than 2 percent of students had a score of zero (Figure II.8).

Students struggled to identify initial sounds or read familiar or pseudo-words. In both groups, students could identify the initial sound of roughly 3 out of 10 words and could read 2 out of 9 familiar words (Table II. 1), such as *sol*, *mama*, or *casa*. They could read 2 pseudo-words or invented words, in one minute. Averaging across groups, around 16 percent of students could not identify a single initial letter sound and 21 percent of students could not read any of 9 simple words shown to them (Figure II.8). Just over three-quarters of students were unable to read a single pseudo-word or invented word.

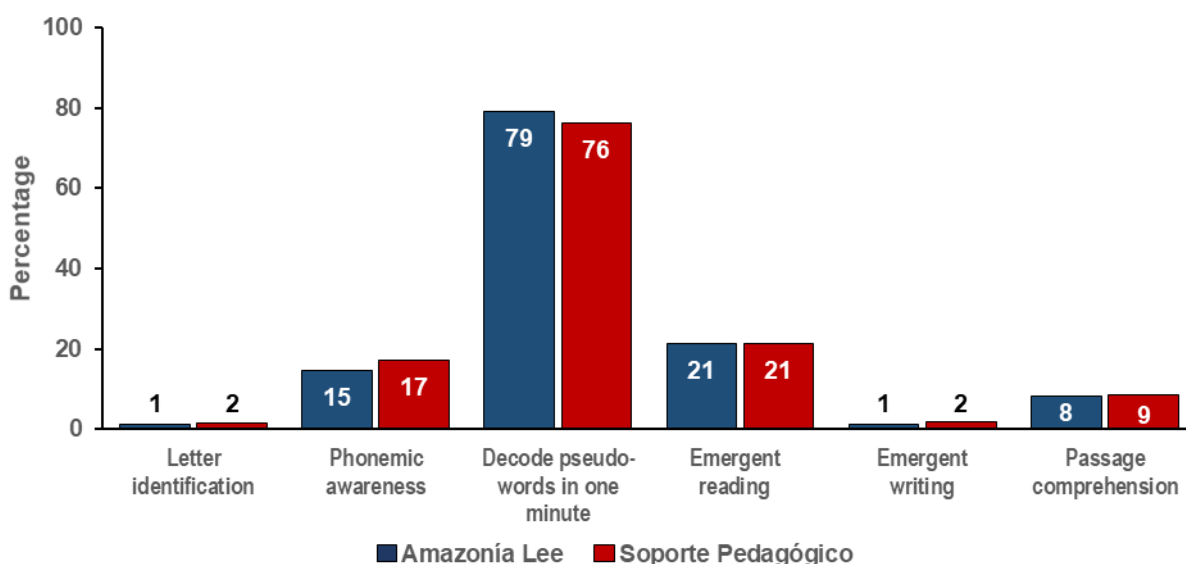
Students demonstrated emerging oral comprehension skills but struggled to understand a simple passage. The assessor read a nine-line passage to the student, and asked the student six oral comprehension questions about the story. In both groups, students responded to roughly three questions correctly (Table II. 1).

Table II.1. Emergent literacy skills by group, 2015

	2015		
	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Letter identification	18.3	18.5	-0.2 (0.472)
Phonemic awareness	3.4	3.3	0.1 (0.448)
Decoded pseudo-words in one minute	1.7	2.0	-0.3 (0.172)
Emergent reading	2.2	2.3	-0.1 (0.247)
Emergent writing	11.7	11.9	-0.2 (0.457)
Passage comprehension	2.8	2.7	0.0 (0.840)

Source: 2015 1st-grade emergent literacy assessments, San Martín.

Note: None of the differences between groups were significant at the 5 percent level.

Figure II.8. Percentage of students with a score of zero by skill and group, 2015

Source: 2015 1st-grade emergent literacy assessments, San Martín.

Note: None of the differences between groups were significant at the 5 percent level.

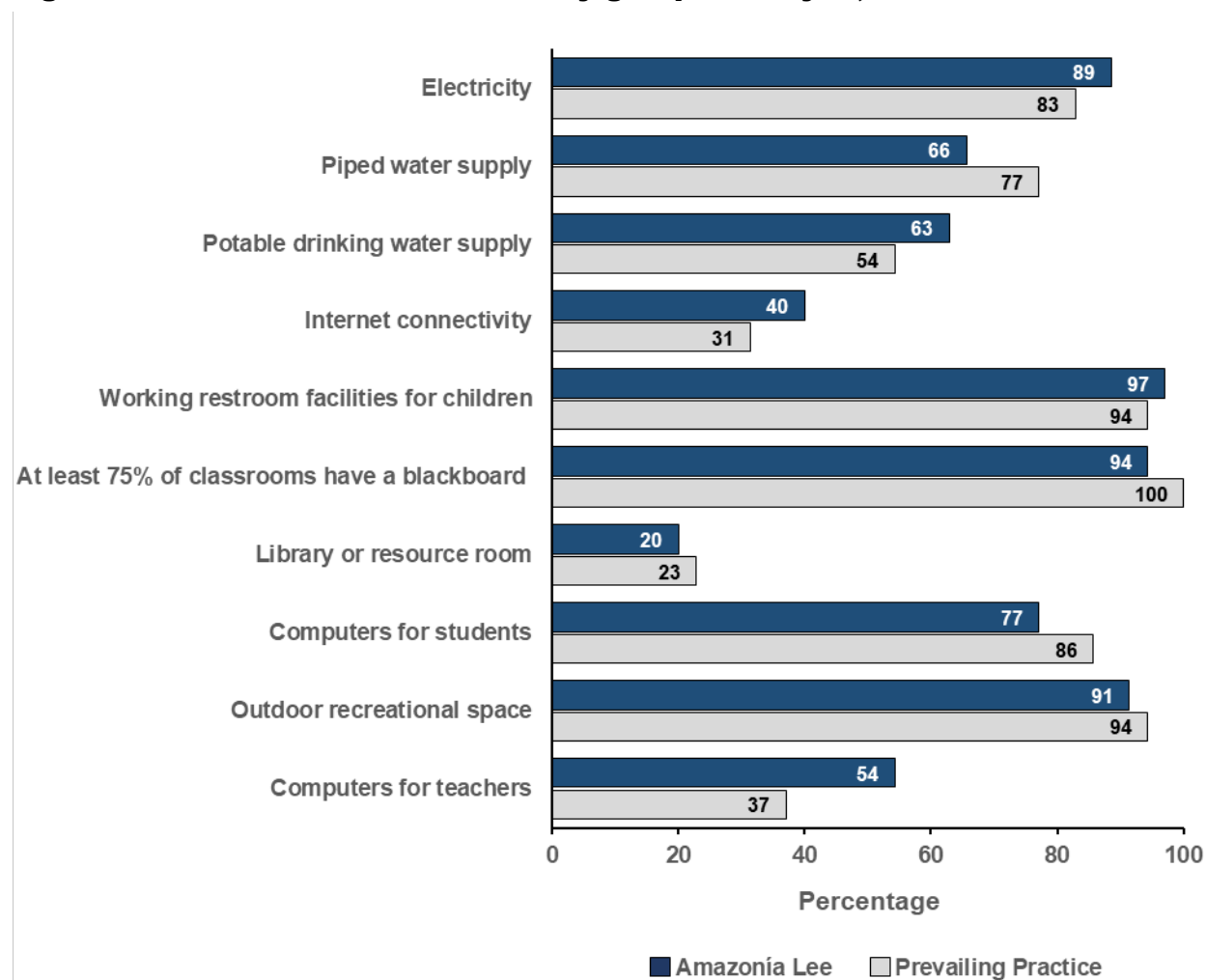
This page has been left blank for double-sided copying.

III. FINDINGS: UCAYALI: SCHOOLS, TEACHERS, AND STUDENTS AT BASELINE

This chapter is parallel to Chapter II. It presents the findings for Ucayali from school and classroom observations and teacher surveys collected in 2014, the year before the intervention began in Ucayali, and the findings from the student early reading assessment collected in 2015.

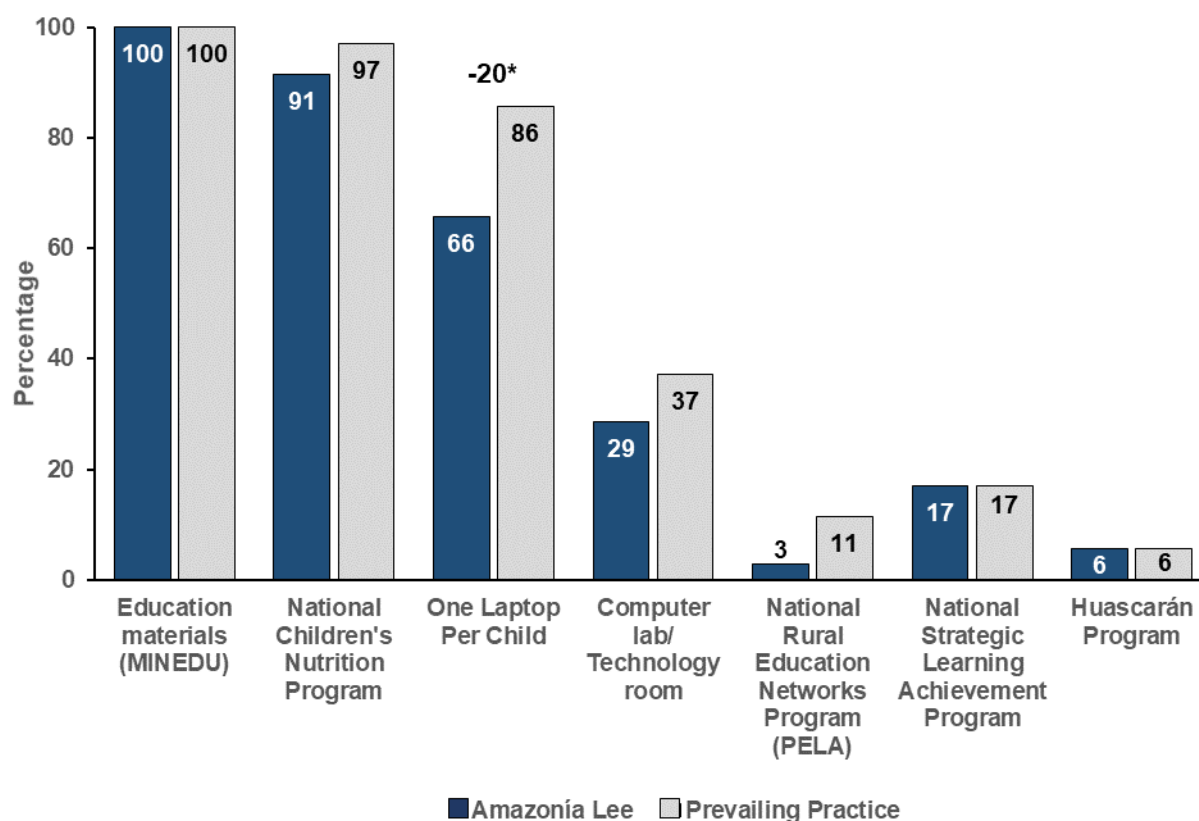
A. School characteristics

Amazonía Lee schools had similar facilities to prevailing practice schools. We did not find any statistically significant differences between the groups in school facilities. Figure III.1 shows school characteristics by group. Even if not statistically significant some differences favor the *Amazonía Lee* group, such as internet connectivity (9 percentage points), computer for teachers (17 percentage points), and potable drinking water (9 percentage points). The prevailing practice group, however, is more likely to have computers for students (9 percentage points). Appendix A.1 provides information on other school characteristics. Because school facilities are not influenced by the intervention itself but could affect final outcomes, we will include school characteristics as covariates in our final impact estimations.

Figure III.1. School characteristics by group in Ucayali, 2014

Source: 2014 school infrastructure observation, Ucayali.

The study programs do not operate in a vacuum. Many schools participate in various support programs that provide school nutrition, education technology, and other inputs to learning. Participation in these programs is summarized in Figure III.2. Only one difference between groups was significant. More schools in the prevailing practice group participated in the One Laptop Per Child program (86 percent, compared to 66 percent in the *Amazonía Lee* group). The last finding might help explain our finding that schools in the prevailing practice group had more computers than *Amazonía Lee* schools. We, therefore, will control for the availability of support programs in 2014 in our final impact estimation.

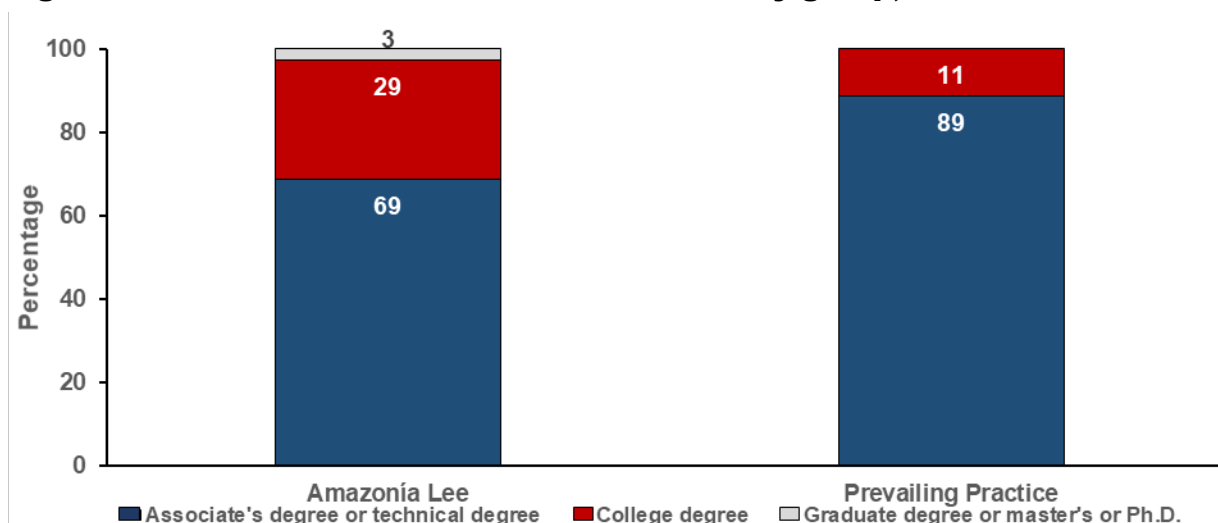
Figure III.2. Support programs in study schools by group in Ucayali, 2014

Source: 2014 school infrastructure observation, Ucayali.

B. Teacher characteristics

Teachers had similar demographic characteristics across groups. In 2014, we did not find statistically significant differences between groups on teacher age or gender. We present these results and results on teacher experience in Appendix B.1.

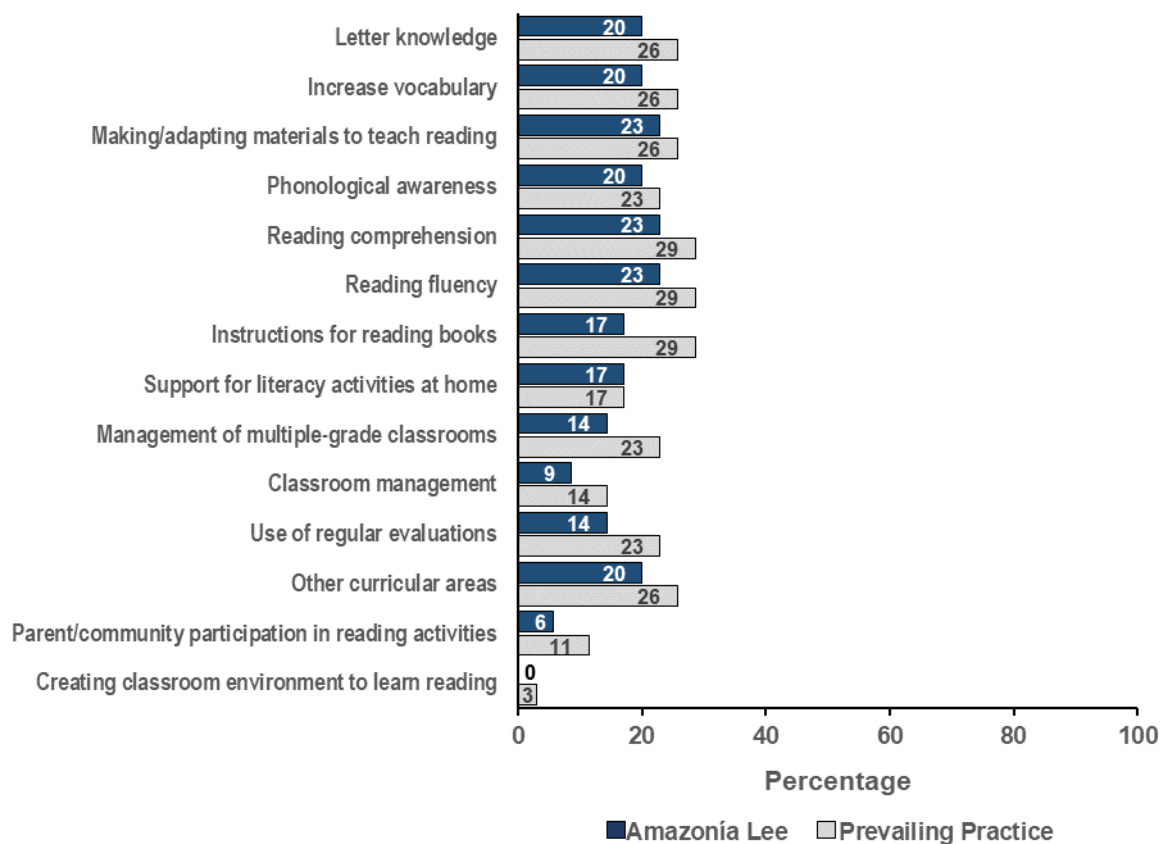
Teachers in the *Amazonía Lee* group had more education but less experience teaching than prevailing practice teachers. In 2014, at *Amazonía Lee* schools, 31.5 percent of teachers had completed a bachelor's or graduate degree, compared to 11.4 percent of teachers at prevailing practice schools. However, at prevailing practice schools, teachers had 14.8 years of teaching experience, compared to 10.7 years at *Amazonía Lee* schools. Figure III.3 shows the distribution of educational attainment by group.

Figure III.3. Teachers' educational attainment by group, 2014

Source: 2014 1st-grade teacher surveys, Ucayali.

Note: Differences in distributions between groups is significant at the 5 percent level.

We found no significant difference in training received by teachers by study group. In 2014, 29 percent of teachers in the prevailing practice group had received training on reading instruction in the previous year, compared to 23 percent of teachers at *Amazonía Lee* schools. The 6 percentage point difference is not statistically significant, nor are the differences in likelihood that teachers had training on specific activities related to reading instruction. Figure III.4 shows training content areas by group. Training was most likely to cover making/adapting materials to teach reading, reading comprehension, letter knowledge, and phonological awareness. Appendix B.2 shows details on factors teachers identified as limiting their ability to teach students to read.

Figure III.4. Training content areas by group, 2014

Source: 2014 1st-grade teacher surveys, Ucayali.

Note: None of the differences between groups is significant at the 5 percent level.

C. Teaching materials and teaching practices

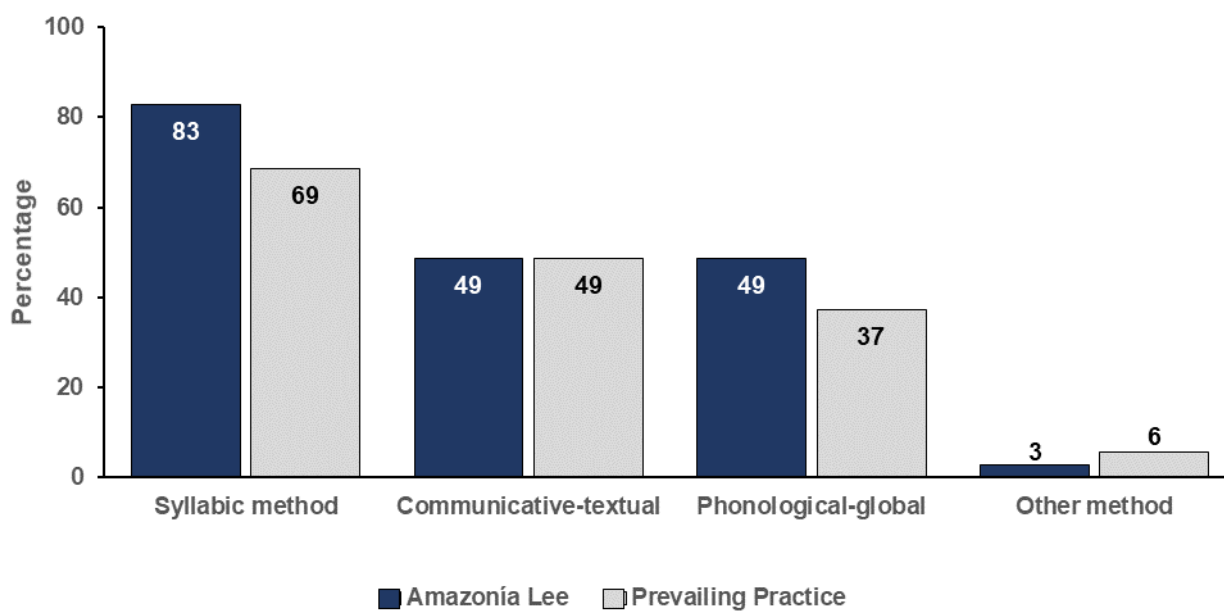
In combination with teacher surveys, classroom observations (conducted in 2014 only) allow us to describe the learning environment and teaching approaches used in study schools before the intervention began, and identify any differences between the groups.

Most teachers reported using one or more of three main pedagogical approaches: syllabic, communicative-textual, and phonological-global. Sixty percent of *Amazonía Lee* teachers reported using more than one of these methods and 46 percent of teachers in prevailing practice schools. The difference is not statistically significant. The syllabic method has been used traditionally for reading instruction in Spanish; it focuses on the progressive learning of vowels, sounds, and consonant-vowel combinations building into words. The communicative-textual approach to early grade reading instruction is the national approach currently supported by the Ministry of Education. It is a two-pronged approach that focuses on developing students' communicative competencies in meaningful contexts, and their interactions with different types of texts that respond to different communicative situations. The phonological-global method integrates whole language with phonological analysis, emphasizing global word recognition and understanding of the correspondence between letters and their sounds. This method typically

promotes a range of foundational reading skills like alphabet knowledge, phonological awareness, decoding, vocabulary, fluency, and comprehension.

There were no statistically significant differences between groups on the instructional approach teachers reported using. The majority of teachers in groups A and B (83 and 69 percent, respectively) reported using the syllabic method to teach reading to their 1st-grade students (Figure III.5). Also, around half of the teachers in both groups reported using the communicative-textual approach. The phonological-global method was reported using by 49 percent of *Amazonía Lee* teachers and by 37 percent of prevailing practice teachers. However, these differences were not statistically significant. .

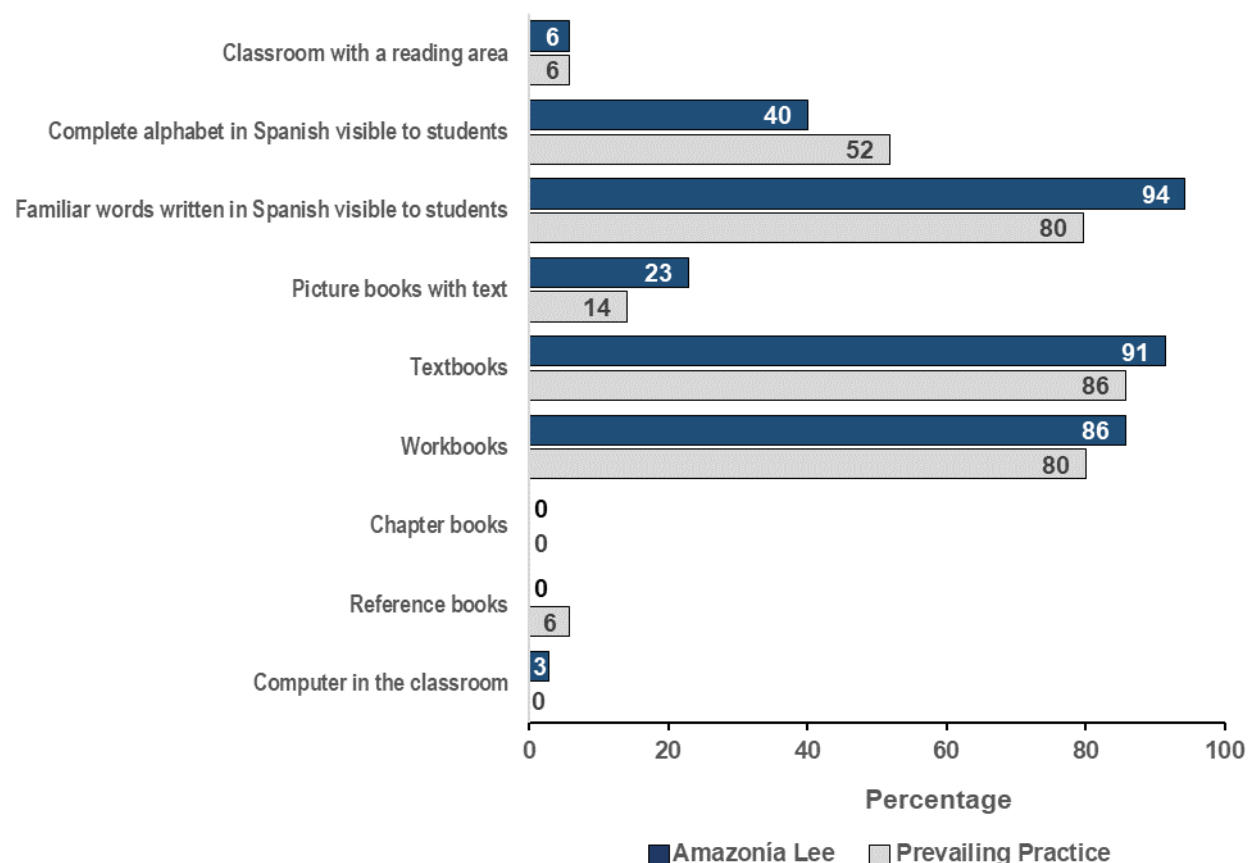
Figure III.5. Reading instruction approaches teachers reported using by group, 2014



Source: 2014 1st-grade teacher surveys, Ucayali.

Note: None of the differences between groups is significant at the 5 percent level.

Availability of reading materials was similar in the two groups. We did not find statistically significant differences in availability of reading materials between the groups (Figure III.6). Most schools in both groups had familiar words in Spanish on display, and roughly half had the full Spanish alphabet on the wall. It was rare, however, for classrooms to have a dedicated area for reading, and a computer in the classroom.. Although not statistically significant, *Amazonía Lee* classrooms were more likely than prevailing practice classrooms to have textbooks, workbooks, and picture books with some text.

Figure III.6. Reading materials available in classrooms by group in Ucayali, 2014

Source: 2014 1st-grade classroom observations, Ucayali.

Note: None of the differences between groups is significant at the 5 percent level.

Teachers were observed using diverse activities to teach their students to read and write. Activities focused on basic skills like letter recognition were most common. Most teachers were observed providing instruction on word writing and recognizing letters. This finding is consistent with teachers focusing on building emergent skills for reading at the beginning of first grade, when this data was collected. Fluency-building activities and reading comprehension strategies were also common. We present the detailed results on literacy instruction practices by group in Appendix C.1.

Teachers in both groups used a similar set of reading instruction activities overall. We did not find significant differences in teachers' likelihood of using 14 out of 15 reading activities in the class observation. However, *Amazonía Lee* teachers were significantly more likely to teach writing.

Teachers generally spend more time on phonological and phonemic awareness than on other activities. Teachers were observed for an average period of 39 minutes while they provided reading instruction to 1st-grade students during a typical language or reading class.

Teachers in both groups spent roughly half their time on phonological and phonemic awareness and a quarter of the time on activities involving written texts. As before, this is expected given that the data was collected in the first part of first grade when teachers are focused on building emergent literacy skills such as phonological and phonemic awareness. Teachers spent little time on oral language activities, vocabulary building, or activities not related to literacy instruction. Detailed results on time use are provided in Appendix C.2.

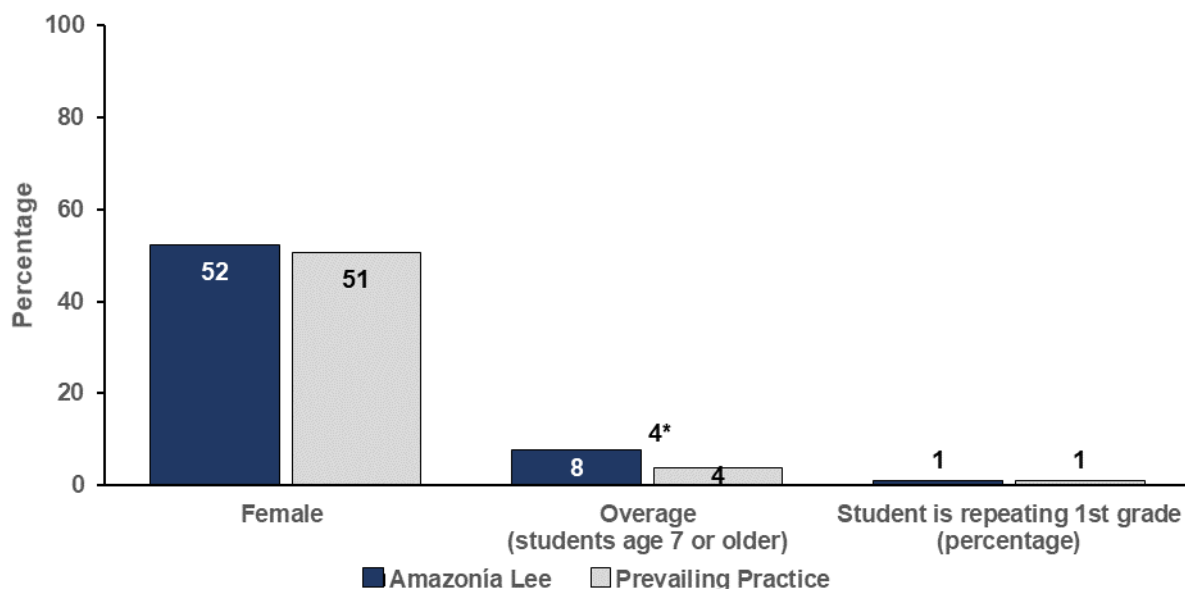
Teachers in the two study groups spent similar time in most activities except for time spent on vocabulary development. *Amazonía Lee* teachers spent less time on vocabulary development (0.9 minutes compared to 2.7); this difference was statistically significant at the 5 percent level. *Amazonía Lee* teachers spent more time on phonological and phonemic awareness activities (24 minutes compared to 20); but this difference was significant only at the 10 percent level.

Teachers in both groups spent most of the time teaching the whole class, and relatively little time working with small groups. Averaging over the two groups, out of 42 minutes observed, teachers spent 22 minutes teaching the whole class, 13 minutes working with individual students, 4 minutes working with small groups, and 2 minutes on activities unrelated to literacy instruction. We did not find significant differences between the groups.

D. Student characteristics and emergent literacy assessment

Here we present baseline results for the 2015 cohort of 1st-grade students. This section follows the same structure as the student characteristics and emergent literacy assessment section presented in Chapter II for San Martín. As described in Chapter II, we present 2014 data in Appendix D.1–D.3 to provide context and assess equivalence of school, teacher, and student characteristics before the intervention began. Here, we present results for the 2015 1st-grade cohort, because this is the cohort we will follow for our endline assessment.

***Amazonía Lee* students were more likely to be old for 1st grade, but were otherwise similar.** In the *Amazonía Lee* group, 8 percent of students were age 7 years or older, compared to 4 percent in the prevailing practice schools (Figure III.7). In both groups almost half the students were female and only 1 percent were repeating first grade. It is possible that being overage has a direct effect on the final student outcomes; therefore, we will control for this student characteristic in the final impact estimation.

Figure III.7. Student characteristics by group, 2015

Source: 2015 school enrollment data reported in emergent literacy assessment

Note: Data are restricted to children who agreed to participate in the emergent literacy assessment. Differences between groups that are significant at the 5 percent level are above the bars and marked with an asterisk.

In the two groups, it was rare for students or their families to speak a language other than Spanish, or for the language of instruction to be a language other than Spanish. Five percent of students in the 2015 cohort reported that they themselves spoke a language other than Spanish, and nine percent reported that their family spoke a language other than Spanish. In both groups, nine percent of students were in a classroom in which the language of instruction was not Spanish. There were no significant differences between groups in student or family language or language of instruction in either year (results now shown).

Most students agreed to participate in the child assessment and participation was similar between groups. In 2015, 85 percent of students selected for the study passed the language screener and attempted the emergent literacy assessment. However, some of the students that attempted the assessment got a score of zero in assessment tasks that required more advanced skills, like decoding. We did not find a statistically significant difference in participation rates between groups. We present the results of the student assessment in Table III.1. In Figure III.8, we show the percentage of students with a score of zero for each task in the emergent literacy assessment.

Students in both groups had similar emergent reading skills except for phonemic awareness. In 2015, first grade students in both groups had similar skills on letter identification, decoding pseudo-words, emergent reading and writing, and passage comprehension. But we found a statistically significant difference in phonemic awareness favoring *Amazonía Lee* students. *Amazonía Lee* students successfully identified the initial sound of 3 out of 10 words, whereas in prevailing practice students identified 2.7 words. This difference is small but statistically significant ($p = 0.013$, Figure III.8). Furthermore, more students in the prevailing practice group were unable to identify correctly the initial sound of 10 words than in the

Amazonía Lee group (26 percent versus 19 percent, respectively). The difference of 7 percentage points is statistically significant. These differences could be due to some initial differences in students skills unrelated to the intervention or they may be an early impact. Because we cannot rule out that the differences are due to the intervention, we will not control for baseline differences in phonemic awareness, since this may bias our impact estimates downward. We will, however, control for other baseline differences when estimating the impact of *Amazonía Lee*. Below, we describe in more detail the findings for each skill in 2015. We also did not find significant differences between groups in emergent reading skills in the 2014 cohort with the exception of phonemic awareness favoring prevailing practice group. (Appendix D.3).

Students in both groups could identify most, but not all letters. For this activity, a student received 2 points for each of 12 letters he or she identified correctly when shown a card with 4 letters. If the student did not identify a letter, the evaluator asked the student to find the letter on a card, saying, for example, “Where is the letter *m*?” If the child found the *m* on the card, he or she received one point for the letter. Students’ letter identification scores were 17.7 out of 24 points for students in *Amazonía Lee* and 18.5 for students in prevailing practice. The difference is not statistically significant (Table III.1). Less than 1 percent of students overall were unable to identify any letters (Figure III.8).

Students in both groups had emerging writing skills but had not yet mastered skills like writing their own name or simple words. Students were asked to write their own name, to write two simple words of their own choosing, and to write three simple words with two or three syllables, dictated by the evaluator. Students lost points for writing nonexistent words, misspelled words, writing in a direction other than from left to right, or holding their pencil incorrectly. Students’ emergent writing score was on average 10.2 points out of 22 for students in *Amazonía Lee* and 10.7 for students in prevailing practice. The difference is not statistically significant (Table III.1). Less than 2 percent of students overall were unable to write anything (Figure III.8).

Students in both groups struggled to read familiar or pseudo-words. Students could read almost 2 out of 9 familiar words, such as *sol*, *mama*, or *casa* in both groups. They could read on average 2 pseudo-words or invented words in a minute. We found no statistical differences between groups. Figure III.8 presents the percent of students that could not read any of 9 simple words shown to them, more than 20 percent in each group. Furthermore, around four in five students were unable to read a single pseudo-word.

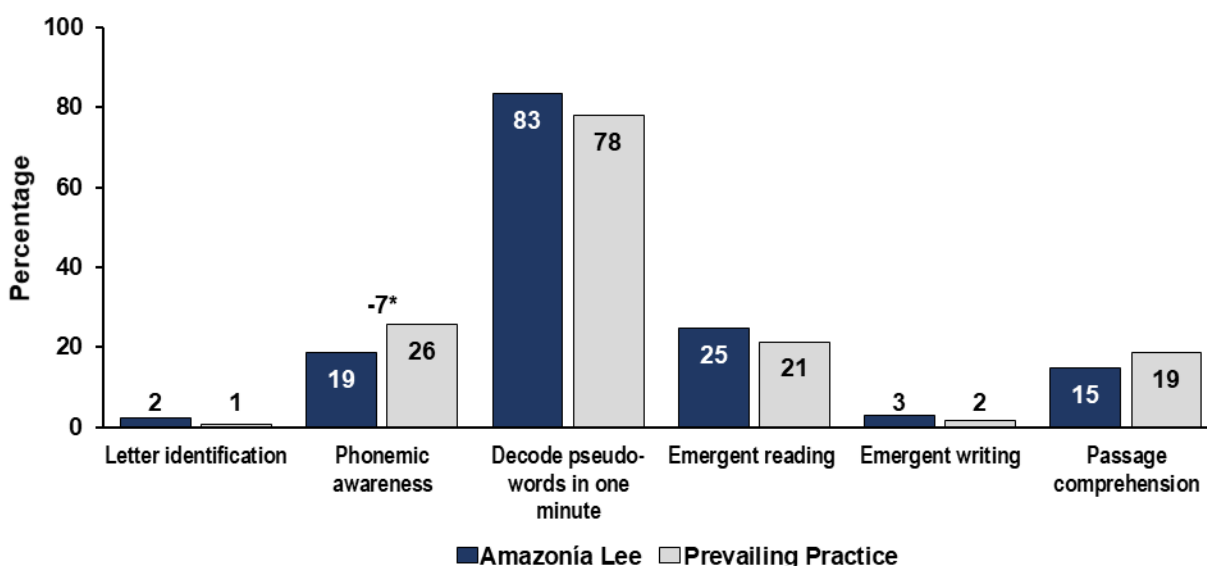
Students in both groups demonstrated emerging oral comprehension skills, but understanding a simple passage remained a challenge. The assessor read a nine-line passage to the student and asked the student to respond to six questions about the passage. Students in both groups were able to correctly answer two questions, on average. The difference is not statistically significant.

Table III.1. Emergent literacy skills by group, 2015

	2015		
	<i>Amazonía Lee</i>	Prevailing practice	Difference
Letter identification	17.7	18.5	-0.8 (0.060)
Phonemic awareness	3.0	2.7	0.3* (0.019)
Decoded pseudo-words in one minute	1.8	1.9	-0.1 (0.750)
Emergent reading	2.0	2.1	-0.2 (0.451)
Emergent writing	10.2	10.7	-0.5 (0.294)
Passage comprehension	2.2	2.1	0.1 (0.446)

Source: 2015 1st-grade emergent literacy assessments, Ucayali.

Note: P-values are shown below differences. Differences that are significant at the 5 percent level are marked with an asterisk.

Figure III.8. Percentage of students with a score of zero by skill and group, 2015

Source: 2015 1st-grade emergent literacy assessments, Ucayali.

Note: Differences between groups that are statistically significant at the 5 percent level are labeled above the bars and are marked with an asterisk.

IV. FINDINGS: EARLY IMPACTS ON TEACHER OUTCOMES IN SAN MARTÍN

In this section, we present results of the 2015 teacher surveys in San Martín to assess early impacts of the *Amazonía Lee* program, such as increased access to training and differences in teaching approach, as compared to the *Soporte Pedagógico* schools. Both *Amazonía Lee* and *Soporte Pedagógico* began in 2015. The first *Amazonía Lee* teacher training in San Martín took place in May. We collected data from May to July 2015, up to three months after *Amazonía Lee* training began. We, therefore, expected to find some differences in access to training and possibly in teacher behavior in 2015 due to some exposure to both programs.

A. Impacts on participation in teacher training programs

In 2015, *Amazonía Lee* teachers were 26 percentage points more likely to have participated in training than *Soporte Pedagógico* teachers. Table IV.1 shows participation rates in training in 2015. We found that 80 percent of teachers at *Amazonía Lee* schools and 54 percent of teachers in the *Soporte Pedagógico* group had participated in training, a statistically significant difference of 26 percentage points. Furthermore, *Amazonía Lee* teachers received more hours of training than *Soporte Pedagógico* teachers, 38 hours compared to 22 (results not shown). As expected, *Amazonía Lee* teachers were more likely to receive training from the DRE, the office responsible for offering the *Amazonía Lee* training, than were teachers in *Soporte Pedagógico*—a statistically significant difference of 67 percentage points. In contrast, teachers in *Soporte Pedagógico* schools were more likely to have received training from the *Soporte Pedagógico* program than were *Amazonía Lee* teachers—a statistically significant difference of -34 percentage points. Both programs began implementation in 2015, so we expected that most teachers in both groups would participate in training in that year. However, because training activities were just beginning at the time data collection began, it is not surprising that some schools had not yet participated in training when the teachers were surveyed.

Table IV.1. Teacher participation in training by training organization and group, 2015

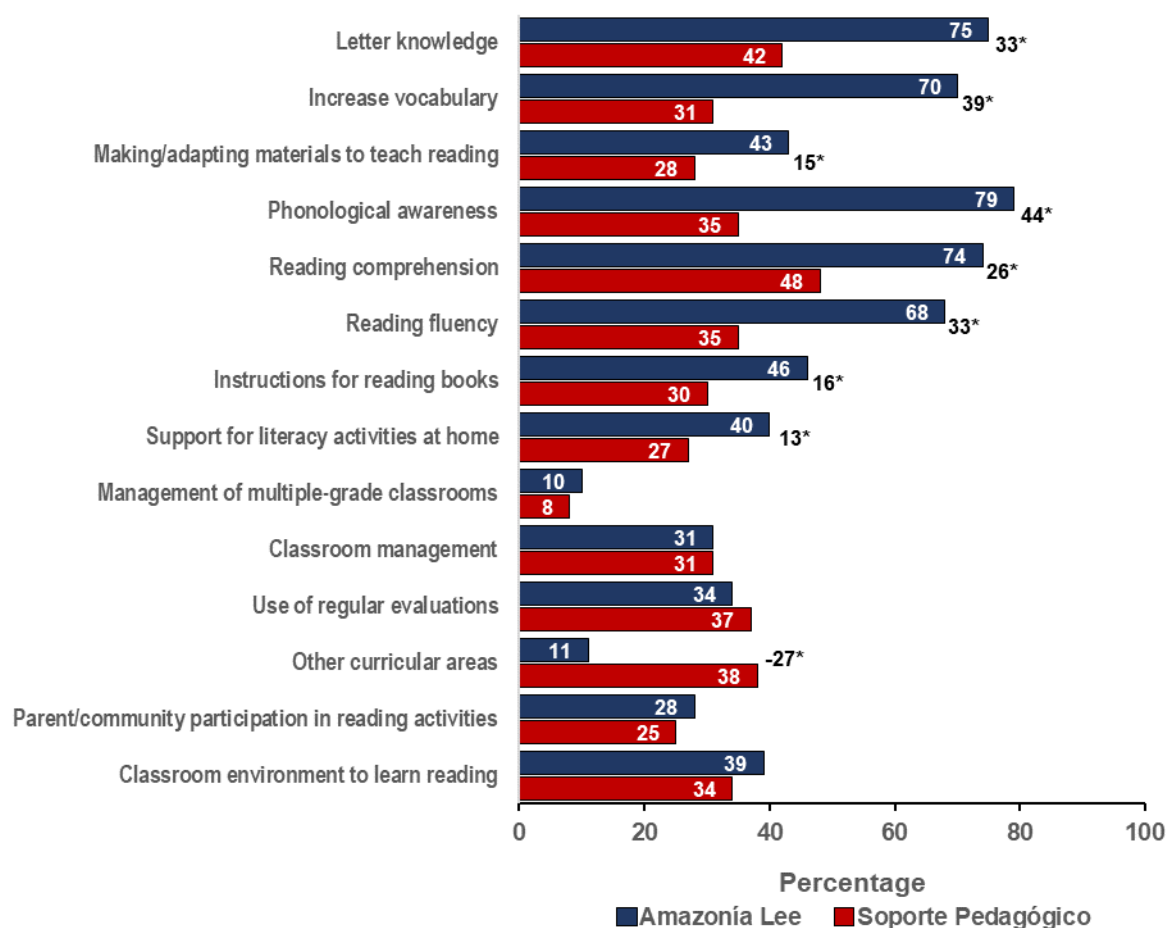
	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Any institution	80.0	54.0	26.0* (0.000)
<i>Soporte Pedagógico</i>	2.0	36.0	-34.0* (0.000)
DRE	69.0	2.0	67.0* (0.000)
UGEL specialist	6.0	11.0	-5.0 (0.198)
Other institution	4.0	5.0	-1.0 (0.741)

Source: 2015 1st-grade teacher survey, San Martín.

Note: P-values are shown below differences. Differences that are statistically significant at the 5 percent level are marked with an asterisk.

Amazonía Lee teachers were more likely than *Soporte Pedagógico* teachers to be trained in most content areas related to reading, but less likely to be trained in other curricular areas. Figure IV.1 presents the percentage of teachers receiving training in each content area by study group. Teachers in the *Amazonía Lee* group were more likely than teachers in the *Soporte Pedagógico* group to report receiving training in most reading areas, and most differences were statistically significant. Two of the largest differences were on phonological awareness (44 percentage points) and letter knowledge (33 percentage points). These differences were consistent with the emphasis that *Amazonía Lee* teacher training places on phonological awareness. Teachers in the *Soporte Pedagógico* group were more likely than *Amazonía Lee* teachers to receive training in other curricular areas (a statistically significant difference of 27 percentage points). This difference is likely due to the fact that the *Soporte Pedagógico* program includes training on math and social studies instruction, but *Amazonía Lee* does not.

Figure IV.1. Training content areas by group, 2015



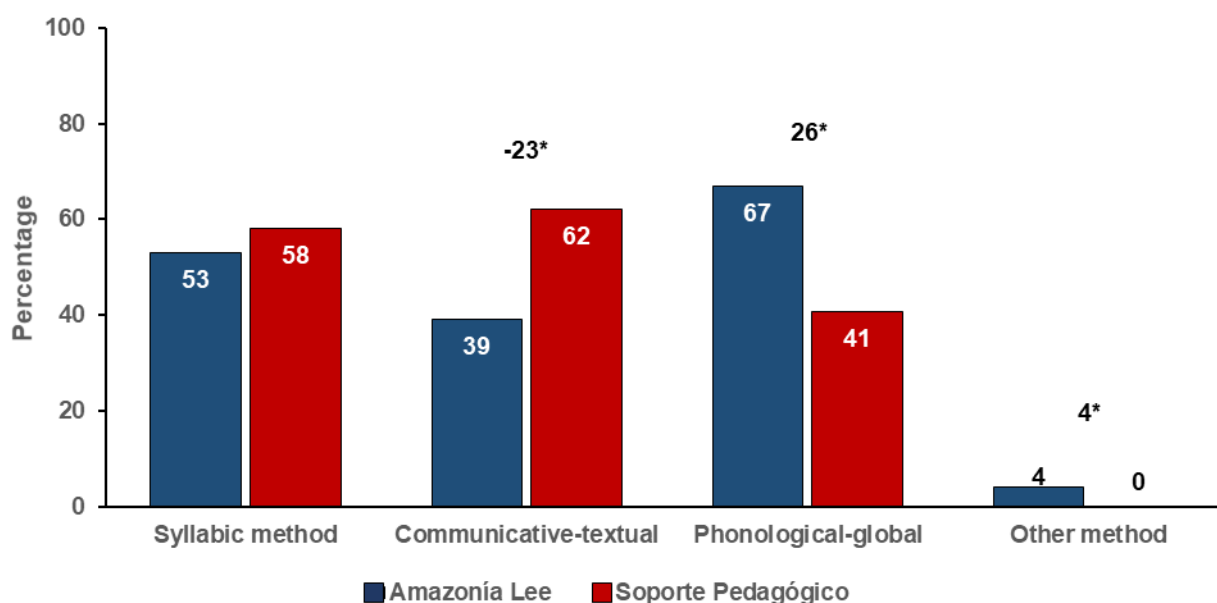
Source: 2015 1st-grade teacher surveys, San Martín.

Note: Differences that are statistically significant at the 5 percent level are labeled next to each pair of bars and marked with an asterisk.

B. Early impacts on teachers' instruction

Amazonía Lee teachers were more likely to report using the phonological-global method and less likely to report using the communicative-textual method than *Soporte Pedagógico* teachers. Figure IV.2 shows the prevalence of each method reported by teachers in each group. We should note that teachers could report using more than one method. We found that 67 percent of teachers in *Amazonía Lee* and 41 percent of teachers in *Soporte Pedagógico* reported using the phonological-global method—a statistically significant difference of 26 percentage points. We also found that 39 percent of the teachers in *Amazonía Lee* reported using the communicative contextual approach, versus 62 percent of *Soporte Pedagógico* teachers—a statistically significant difference of -23 percentage points. This is consistent with the training offered in each group, since *Soporte Pedagógico* program's trainings focused on the use of the communicative-textual method. *Amazonía Lee* trainings' focus on instruction of foundational reading skills, which has commonalities with the phonological-global approach. The syllabic method remained common in both groups, with 56 percent of teachers using it. Forty-eight percent of *Amazonía Lee* teachers and 45 percent of *Soporte Pedagógico* teachers reported more than one method.

Figure IV.2. Reading instruction approaches teachers reported using by group, 2015



Source: 2015 1st-grade teacher surveys, San Martín.

Note: Differences that are statistically significant at the 5 percent level are above each pair of bars and marked with an asterisk.

We found no significant differences in the type of reading instruction materials used between *Amazonía Lee* and *Soporte Pedagógico* teachers. Table IV.2 shows the percentage of teachers using each type of reading material at least three to four times per month in 2015, by group. We found no statistically significant differences on the type of materials used between the study groups. In both groups, workbooks or exercises guides, school textbooks, and cards with

syllables were used around 90 percent of the time. In principle, we did not expect to find a difference in use of materials so early in 2015; however, we could find it in the final data collection.

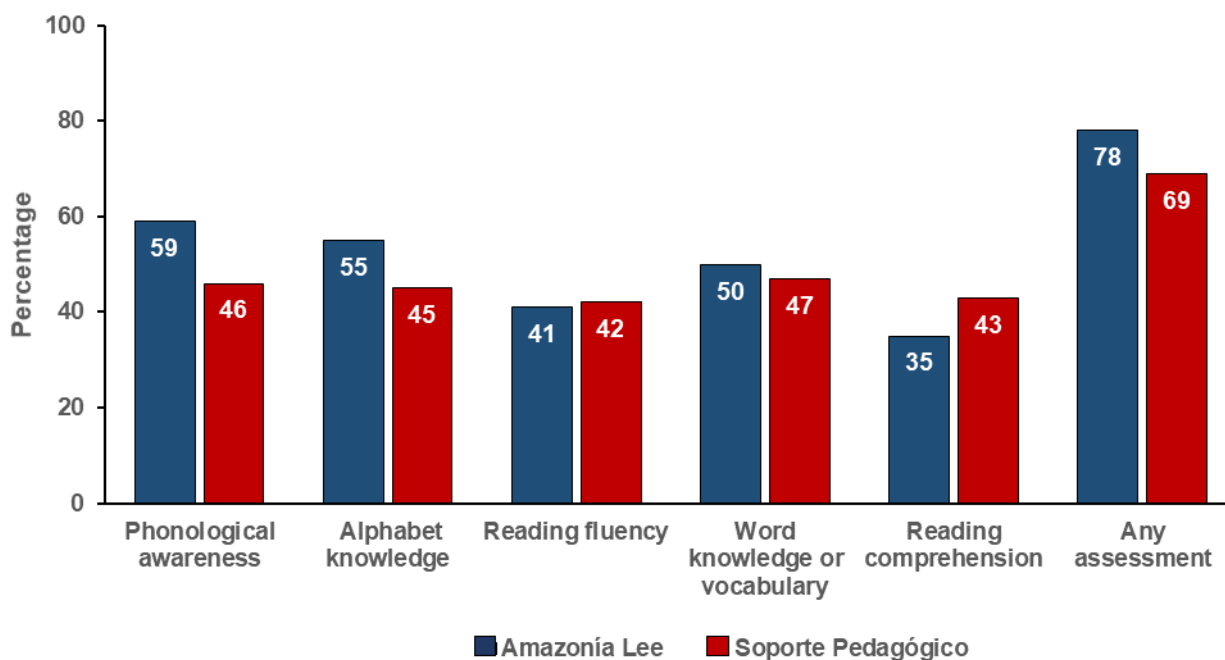
Table IV.2. Percentage of teachers who report using reading materials at least 3–4 times per month, 2015

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Materials from MINED	57.0	64.0	-7.0 (0.288)
School textbooks	90.0	89.0	1.0 (0.820)
Workbooks or exercise guides	93.0	94.0	-1.0 (0.765)
Simple-reading books	78.0	82.0	-4.0 (0.519)
Books to name objects or letters of the alphabet	86.0	78.0	8.0 (0.131)
Cards with syllables or words	93.0	89.0	4.0 (0.348)
Picture or artwork books	82.0	81.0	1.0 (0.854)

Source: 2015 1st-grade teacher survey, San Martín.

Note: P-values are shown below differences. Differences that are significant at the 5 percent level are marked with an asterisk.

We found no significant differences in the use of daily assessments or on the type of reading skills assessed between *Amazonía Lee* and *Soporte Pedagógico* teachers. Figure IV.3 provides details of the type of daily assessments by group. The results of the teacher survey indicate that regular student assessment is widespread in both groups. In the *Amazonía Lee* group, 78 percent of teachers reported assessing their students daily, compared to 69 percent of teachers in the *Soporte Pedagógico* group; the difference is not statistically significant. The skills that teachers were most likely to test daily were alphabet knowledge and phonological awareness, with roughly half of teachers testing these skills daily. Teachers in both groups were least likely to test students daily on reading comprehension or reading fluency. The *Amazonía Lee* training included guidance on developing and using formative assessments regularly, but we did not find any significant difference across the groups. It is possible that teachers need more time to change their practice in this area, or the *Amazonía Lee* training may not have covered those topics yet at the time of data collection.

Figure IV.3. Daily assessments of students' reading skills by group, 2015

Source: 2015 1st-grade teacher surveys, San Martín.

Note: None of the differences were significant at the 5 percent level.

Teachers from both groups divided their time in similar ways among various reading instruction activities except for reading aloud. Table IV.3 shows the distribution of time by group. We found no statistically significant differences on the amount of time spent on reading activities such as learning letters, practicing phonological or phonemic awareness, copying words, building vocabulary, and narrating stories without text. However, *Amazonía Lee* teachers spent statistically significantly less time practicing reading aloud and reading comprehension than did teachers at *Soporte Pedagógico* schools (32.5 minutes compared to 38.1). Nonetheless, the results in Table IV.3 and IV.4 should be interpreted with caution. These tables show teachers' self-reported time on activities that may not be mutually exclusive or exhaustive, and are subject to recall bias.

Table IV.3. Teachers' self-reported time spent on reading activities in a typical reading class by group (minutes), 2015

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Learning to identify the letters of the alphabet	33.9	32.1	1.9 (0.530)
Practicing phonological or phonemic awareness activities	36.1	36.3	-0.2 (0.939)
Copying letters or words	33.4	34.5	-1.1 (0.645)
Taking dictation	29.8	30.9	-1.2 (0.666)
Building vocabulary and learning new words	35.4	36.5	-1.2 (0.621)
Narrating familiar stories without reference to text	25.8	29.5	-3.8 (0.098)
Practicing reading aloud and reading comprehension	32.5	38.1	-5.6* (0.017)
Total minutes reported	221.1	233.4	-12.3 (0.253)

Source: 2015 1st-grade teacher surveys, San Martín.

Note: P-values are shown below differences. Differences that are significant at the 5 percent level are marked with an asterisk.

We found positive impacts of *Amazonía Lee* on total time spent in teaching activities. *Amazonía Lee* teachers reported spending almost five more hours per week teaching all subjects than *Soporte Pedagógico* teachers. Table IV.4 shows how teachers reported spending their time. *Amazonía Lee* teachers reported spending on average 28 hours per week teaching, whereas *Soporte Pedagógico* teachers reported spending 23 hours per week teaching. The difference of 5 hours per week is statistically significant and is driven by statistically significant differences in time spent teaching reading (1.8 hours) and math (1.3 hours). Furthermore, when we sum time teachers reported spending on teaching, class preparation, supporting students, administrative tasks, and other activities, we find that *Amazonía Lee* teachers spent 51 hours a week on all of those activities and *Soporte Pedagógico* teachers spent 45 hours per week; the difference of 6.1 hours per week is statistically significant and is driven by the significant difference in reported time spent teaching. We found no differences in time allocated to class preparation, supporting students, administrative tasks, or other activities.

Table IV.4. Teachers' self-reported time use by group: 2015 (hours per week)

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Teaching (all subjects)	27.8	22.9	4.9* (0.010)
Teaching reading	10.5	8.8	1.8* (0.032)
Teaching math	8.8	7.5	1.3* (0.040)
Class preparation	12.6	11.8	0.9 (0.425)
Supporting students	7.1	6.3	0.8 (0.244)
Administrative tasks	1.6	2.0	-0.3 (0.501)
Other activities	3.8	4.3	-0.5 (0.447)
Total hours per week	51.3	45.2	6.1* (0.024)

Source: 2015 1st-grade teacher surveys, San Martín.

Note: The “supporting students” category includes meeting with students, monitoring students during recess or free time, providing support to students with low academic achievement, and preparing students for national exams.

Some of teachers' time may have appeared in multiple categories. For example, one of the activities grouped under “other” is preparing students for national tests. This time may have included time reported as teaching time.

This page has been left blank for double-sided copying.

V. FINDINGS: EARLY IMPACTS ON TEACHER OUTCOMES IN UCAYALI

In this section, we present results of the 2015 teacher surveys in Ucayali to assess early impacts of the *Amazonía Lee* program, such as increased access to training, and differences in teaching approach as compared to the prevailing practice schools. The first *Amazonía Lee* teacher training in Ucayali took place in March 2015. We collected data in June and July 2015, three to four months after *Amazonía Lee* training began. We, therefore, expected to find some differences in access to training and possibly in teacher behavior in 2015 due to some exposure to the program.

A. Impacts on participation in teacher training programs

In 2015, *Amazonía Lee* teachers were 69 percentage points more likely to have participated in training than teachers in the prevailing practice group. Table V.1 shows participation rates in training in 2015. At *Amazonía Lee* schools, 80 percent of teachers indicated having participated in teacher training, compared to 11 percent of teachers at prevailing practice schools—a significant difference of 69 percentage points. *Amazonía Lee* teachers were more likely to have received training from the DRE, the office responsible for offering the *Amazonía Lee* training, than from others. Teachers at *Amazonía Lee* schools had also received more hours of training than teachers at the prevailing practice schools (40 hours compared to 2, results not shown). These results are consistent with the fact that *Amazonía Lee* began implementation about four months before data collection and that prevailing practice schools were not implementing any program in particular. This is an important difference from San Martín where, as we have shown, over half of the schools in group B were implementing other programs, in particular *Soporte Pedagógico*.

Table V.1. Teacher participation in training by training organization and group, 2015

	Amazonía Lee	Prevailing practice	Difference
Any institution	80.0	11.4	68.6* (0.000)
<i>Soporte Pedagógico</i>	2.9	0.0	2.9 (0.324)
DRE	68.6	0.0	68.6* (0.000)
UGEL specialist	0.0	2.9	-2.9 (0.324)
UGEL or other institution	8.6	8.6	0.0 (1.000)

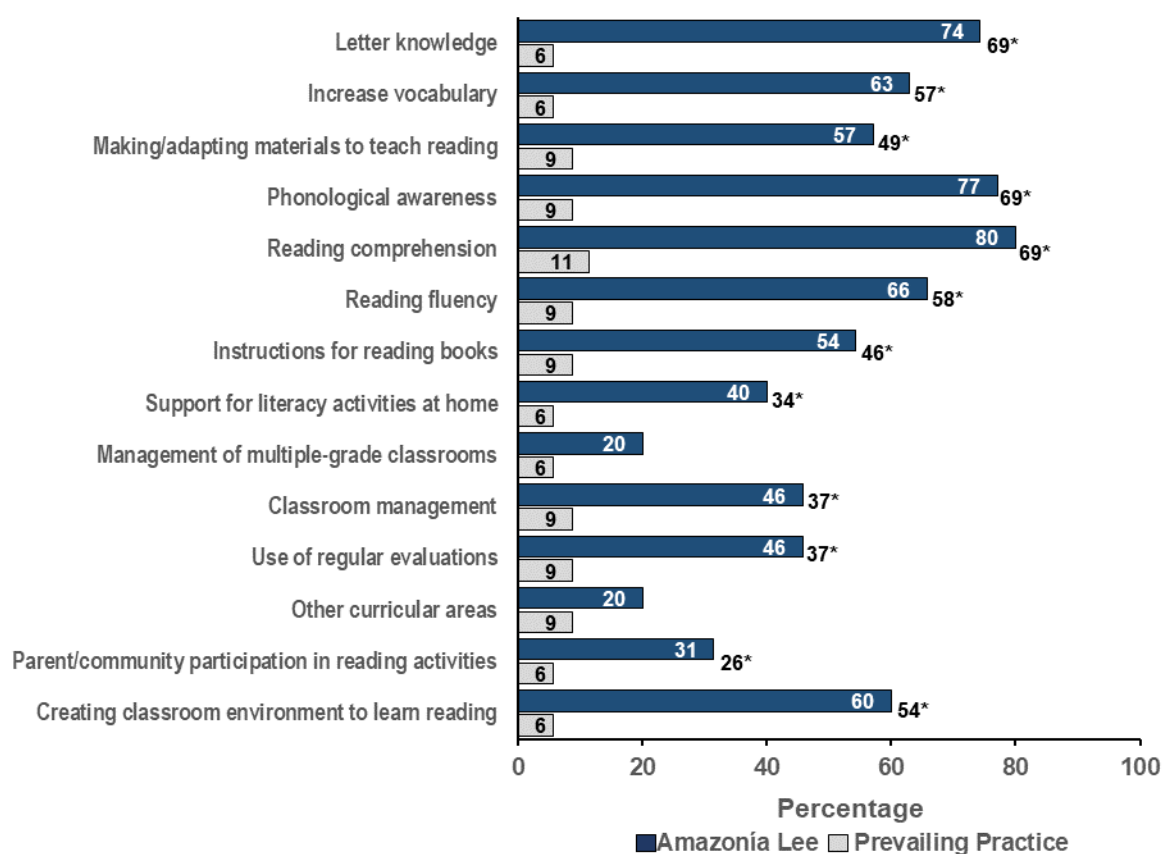
Source: 2015 teacher surveys, Ucayali.

Note: P-values are shown below differences. Differences that are significant at the 5 percent level are marked with an asterisk.

Amazonía Lee teachers were more likely than prevailing practice teachers to be trained on all content areas related to reading instruction, and on classroom management. The differences were as small as 26 percentage points and as large as 69 percentage points. Teachers in the *Amazonía Lee* group were more likely than teachers in the prevailing practice group to

have had training on all reading content areas, and all of the differences were statistically significant. We find a 69 percentage point difference for training on reading comprehension, letter knowledge, and phonological awareness, and a difference of over 50 percentage points for training on vocabulary, reading fluency, and creating a class environment to learn reading. Training on activities outside the school, such as training for supporting literacy activities at home and parent participation in reading activities, have the lowest differences, 34 and 26 percentage points, respectively. The only areas in which *Amazonia Lee* teachers were not significantly more likely to have received training are the management of multigrade classrooms and teaching other curricular areas (see Figure V.1). These findings are consistent with the emphasis of *Amazonia Lee* on reading instruction and with the fact that the prevailing practice group had low rates of training in all areas.

Figure V.1. Training content areas by group, 2015



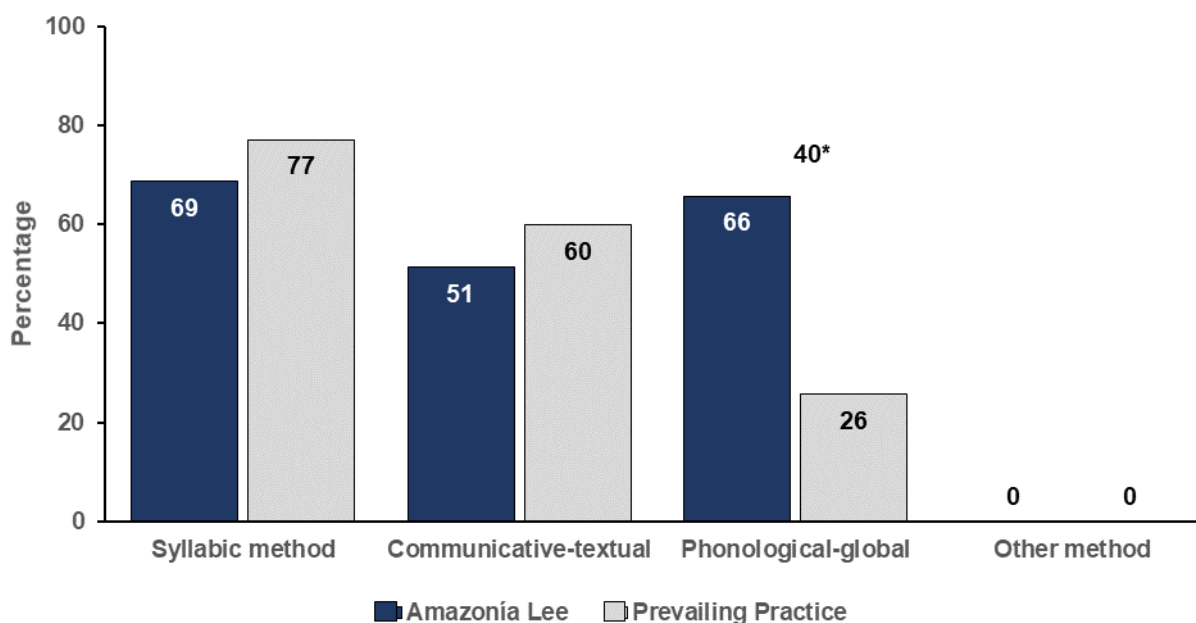
Source: 2015 1st-grade teacher surveys, Ucayali.

Note: Differences that are statistically significant at the 5 percent level are labeled next to each pair of bars and marked with an asterisk.

B. Early impacts on teachers' instruction

Teachers in the *Amazonía Lee* group were 40 percentage points more likely to report using the phonological-global method than those in the prevailing practice group. Figure V.2 shows the prevalence of each method reported in each group. We should note that teachers could report using more than one method. Sixty-six percent of *Amazonía Lee* teachers reported using the phonological-global approach and 26 percent of the prevailing practice group did—a statistically significant difference of 40 percentage points. This is consistent with *Amazonía Lee* trainings' focus on instruction of foundational reading skills, which has commonalities with the phonological-global approach. We found no statistically significant differences in the use of the communicative-textual approach and the syllabic method between the study groups.

Figure V.2. Reading instruction approaches teachers reported using by group, 2015



Source: 2015 1st-grade Teacher Surveys, Ucayali.

Note: Differences that are statistically significant at the 5 percent level are labeled next to each pair of bars and marked with an asterisk.

We found no significant differences between groups on the types of materials used for reading instruction. Table V.2 shows the percentage of teachers using each type of text at least three to four times per month in each group in each year. We found no statistically significant differences between groups. In both groups, teachers used workbooks or exercise guides and cards with syllables and words more frequently than any other type of text. Materials from the Ministry of Education were least commonly used in both groups.

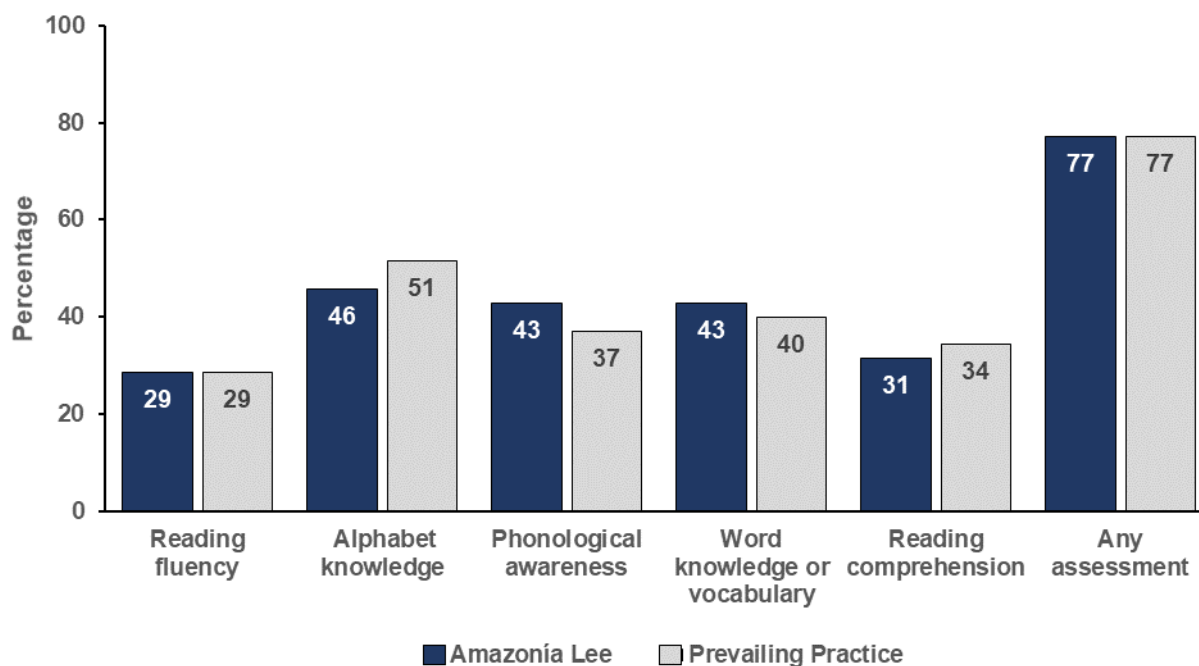
Table V.2. Percentage of schools in which students used reading materials at least 3–4 times per month, 2015

	<i>Amazonía Lee</i>	Prevailing practice	Difference
Materials from MINED	68.6	71.4	-2.9 (0.800)
School textbooks	91.4	85.7	5.7 (0.487)
Workbooks or exercise guides	97.1	94.3	2.9 (0.571)
Simple-reading books	80.0	80.0	0.0 (1.000)
Books to name objects or letters of the alphabet	88.6	85.7	2.9 (0.744)
Cards with syllables or words	91.4	94.3	-2.9 (0.571)
Picture or artwork books	91.4	91.4	0.0 (1.000)

Source: 2015 1st-grade teacher surveys, Ucayali.

Note: P-values are shown below differences. Differences that are significant at the 5 percent level are marked with an asterisk.

We found no significant differences between *Amazonía Lee* and prevailing practice teachers on the use of daily assessments and the type of assessment used. Figure V.3 provides details by type of assessment used in each group. The use of daily student assessment is widespread in both groups. In both groups, teachers at 77 percent of schools assessed their students daily. In 2015, the skills that teachers in both groups were most likely to test daily were alphabet knowledge, word knowledge, and phonological awareness. We found no statistically significant differences between groups. Teachers in both groups were least likely to test students daily on reading comprehension or reading fluency.

Figure V.3. Daily assessments of students' reading skills by group, 2015

Source: 2015 1st-grade teacher surveys, Ucayali.

Note: None of the differences are significant at the 5 percent level.

We found no significant differences between groups in the amount of time teachers spent on reading activities within a class. Table V.3⁸ shows the distribution of time by group. We found no statistically significant differences in the time teachers reported spending on each of the reading activities. In both groups, teachers reported spending most time practicing reading aloud and reading comprehension, and an average of 28 minutes on each activity on a typical day. Although the difference is not statistically significant, teachers at *Amazonía Lee* schools reported spending at least 5 more minutes each day on both identifying the letters of the alphabet and practicing phonological and phonemic awareness. This is consistent with the *Amazonía Lee* emphasis on teaching emergent reading skills.

⁸ The results in Table V.3 and V.4 should be interpreted with caution. These tables show teachers' self-reported time on activities that may not be mutually exclusive or exhaustive, and are subject to recall bias.

Table V.3. Teachers' self-reported time spent on reading activities in a typical reading class by group, 2015

	<i>Amazonía Lee</i>	Prevailing practice	Difference
Learning to identify the letters of the alphabet	27.6	22.6	5.0 (0.193)
Practicing phonological or phonemic awareness activities	30.9	24.1	6.8 (0.161)
Copying letters or words	32.4	31.9	0.6 (0.899)
Taking dictation	29.3	25.8	3.5 (0.475)
Building vocabulary and learning new words	28.7	26.8	1.9 (0.582)
Narrating familiar stories without reference to text	27.1	26.1	0.9 (0.810)
Practicing reading aloud and reading comprehension	33.0	34.1	-1.2 (0.800)
Total minutes reported	198.7	186.9	11.9 (0.534)

Source: 2015 1st-grade teacher surveys, Ucayali.

Note: P-values are reported in parentheses. None of the differences were significant at the 5 percent level.

We found no significant differences in the amount of time teachers spent on teaching-related activities. Table V.4 shows the distribution of time by group. In both groups, teachers spent on average 23 hours per week teaching, 13 hours preparing for class, and 9 hours on other activities such as supporting students, administrative tasks, and other. We did not find evidence of impacts of *Amazonía Lee* on how teachers spent their time.

Table V.4. Teachers' self-reported time use by group, 2015 (hours per week)

	<i>Amazonía Lee</i>	Prevailing practice	Difference
Teaching	22.5	22.9	-0.4 (0.776)
Class preparation	12.7	13.4	-0.7 (0.648)
Supporting students	4.9	6.4	-1.5 (0.106)
Administrative tasks	1.8	2.0	-0.2 (0.857)
Other activities	3.6	3.4	0.2 (0.900)
Total hours per week	43.6	46.0	-2.4 (0.437)

Source: 2015 1st-grade teacher surveys, Ucayali.

Note: P-values are reported in parentheses. None of the differences were significant at the 5 percent level.

The "supporting students" category includes meeting with students, monitoring students during recess or free time, providing support to students with low academic achievement, and preparing students for national exams.

Some of teachers' time may have appeared in multiple categories. For example, one of the activities grouped under "other" is preparing students for national tests. This time may have included time reported as teaching time.

This page has been left blank for double-sided copying.

VI. CONCLUSIONS AND NEXT STEPS

A. Conclusions from comparing schools, teachers, and students at baseline

The random assignment created similar groups of schools in both regions. In 2014, before the intervention started, schools in both study groups were statistically similar in San Martín and Ucayali. Schools in both groups had basic facilities, although the levels of availability were generally lower in Ucayali, which is consistent with the fact that this is one of the more disadvantaged regions in Peru. Some school facilities, such as the availability of drinking water supply or internet connectivity, favor the *Amazonía Lee* group in both regions; others, such as computers for students, favor prevailing practice in Ucayali, but the differences are not large enough to be considered statistically significant.

While no systematic differences in school characteristics were found between groups, the impact estimation will include school characteristics as covariates in the regression framework. Because baseline differences on school characteristics could affect student outcomes and a few differences were found, we will include school characteristics as covariates in the regression model used for impact estimations. The differences in school characteristics are unrelated to the intervention outcomes. Therefore, controlling for school characteristics will not bias the impact results but can improve the statistical precision of our impact estimates.

The random assignment created teacher groups with similar demographic characteristics, reading instructional approaches, and use of time in both regions. In 2014, before the intervention started, in both regions, teachers in both groups had similar age, gender, experience, and tenure, but differed on years of education. In San Martín, teachers in the *Amazonía Lee* group had less years of education than teachers in group B; the opposite was found in Ucayali. We did not find evidence that the teachers were systematically different between the study groups in either region. Furthermore, teachers reported and we observed similar approaches to reading instruction before the intervention started. In both regions, the syllabic-method was the instructional approach teachers' most reported, followed by the communicative-textual and the phonological global. Furthermore, we did not find differences in the type of reading instruction activities teachers used during the classroom observation periods (with one exception in San Martín). We also found no differences in teachers' use of class time for reading activities in either region.

***Amazonía Lee* impacts may be diluted in San Martín due to the greater amount of training received by teachers in the *Soporte Pedagógico* group before the intervention started.** While we found no statistically significant differences in teacher training across study groups in Ucayali in 2014, in San Martín teachers in the *Soporte Pedagógico* group were more likely to have had training than teachers in the *Amazonía Lee* group. The implication for the evaluation is that the *Amazonía Lee* group will be compared to a group that had already received more teacher training in 2014 prior to the rollout of the intervention, which could potentially dilute impacts of *Amazonía Lee*.

The impact estimation models will include student characteristics as covariates where baseline differences were found. In 2015, we found two statistically significant differences in student characteristics in San Martín. Students in the *Amazonía Lee* group were more likely to be

repeating 1st grade and to be overage than students in the *Soporte Pedagógico* group. In Ucayali, we found that students in the *Amazonía Lee* group were more likely to be overage than students in the prevailing practice group. In our final analysis for each region, we will control for these student characteristics, because age and grade repetition could directly affect a student's final reading level. Therefore, including them will allow us to account for the direct effect of these variables in the reading outcomes and reduce the overall variance of the estimation increasing the precision of the estimates.

In 2015, students in both groups had similar emergent reading skills in both regions, except for phonemic awareness in Ucayali. In 2015, in both regions, first grade students in both groups had similar skills on letter identification, decoding pseudo-words, emergent reading and writing, and passage comprehension. However, in Ucayali, we found a statistically significant difference in students' phonemic awareness scores favoring the *Amazonía Lee* group. Because 2015 data were collected about two months after the intervention had begun, the 2015 emergent literacy assessments results may have already been affected by the intervention. With the exception of phonemic awareness, it is unlikely that the intervention would have measurable effects on other reading outcomes in such a short time. Phonemic awareness is a foundational reading skill that could potentially change after a few months of practice. Therefore, we cannot be certain whether the observed difference in phonemic awareness is due to a pre-existing difference unrelated to the intervention or if it is the result of an early effect of the *Amazonía Lee* program.

Student emergent skills that were found to be similar between groups in 2015 will be included as covariates in the impact estimation. Emergent skills for which differences were found in 2015 will not be included as covariates. We plan to use the emergent reading skills scores for which no differences were found in 2015 as covariates in the regressions for final impact results. Including these covariates will allow us to account for initial skills and improve the statistical precision of the impact. However, the difference in students' phonemic awareness skills between groups that we found in Ucayali could represent either a pre-existing difference unrelated to the *Amazonía Lee* program, or early impact of the intervention. If we could be certain this was a pre-existing difference unrelated to the intervention, including the covariate in the regression model will help us improve the precision of the estimates. However, if the difference was due to the intervention (early impact), including it as covariate in the model, will reduce the overall impacts attributed to the intervention (Schochet 2008). Therefore, we will not include as covariates reading skills for which differences in 2015 were found.

We will include school averages of student assessments collected in 2014 as covariates in the impact estimation model. Student assessment results from 2014 are not affected by the intervention given that *Amazonía Lee* started in 2015. By including the school-level averages from 2014 in the impact estimation models, we can control for initial reading differences between groups and improve the precision of our estimates which will allow us to statistically detect smaller effects.

B. Conclusions from early impact analysis on teacher outcomes

Teacher outcome data suggest a few early impacts on teachers' use of time and instructional method after a few months of *Amazonía Lee* teacher training. At the time of

data collection in 2015, 80 percent of the teachers in the *Amazonía Lee* group received training in both regions. In San Martín, slightly more than half of the teachers in the comparison group received training (mostly from *Soporte Pedagógico*), whereas in Ucayali only a tenth of the comparison teachers received training. Despite these differences in teacher training between groups, we found only two significant differences on time spent in teaching activities in San Martín and none in Ucayali. In San Martín, *Amazonía Lee* teachers reported spending almost five more hours per week teaching all subjects than did *Soporte Pedagógico* teachers. *Soporte Pedagógico* teachers reported spending, on average, five more minutes in reading aloud and reading comprehension practice during a reading class.

C. Next steps

Our cohort of interest, the cohort that was in 1st grade in 2015, will be in 2nd grade in the 2016 school year in both regions. We are focusing on this cohort because of our interest in the impacts of the *Amazonía Lee* intervention on early grade reading. The 2015 cohort of first graders is the youngest cohort that benefited from *Amazonía Lee*. In the second half of the 2016 school year, we will conduct classroom observations, teacher surveys, and a school infrastructure survey. Most importantly, we will conduct student assessments that will provide us with the information on our main outcomes to estimate the impact of *Amazonía Lee* on student reading achievement. To complement the survey data collection, we will carry out an implementation study, which will include document review, key informant interviews, and focus groups with teachers. We will analyze endline data in 2017 and expect to have a presentation of preliminary results to share with stakeholders in mid-2017 in English and Spanish. The final report will be published in English and will be available at the end of 2017.

This page has been left blank for double-sided copying.

REFERENCES

- Dickinson, D.K., J.P. Brady, and J.A. Hirschler. “Evaluacion Lectoescritura Emergente: Versión adaptada de la versión en inglés.” Centro para el Desarrollo Educativo y Carolyn Chaney. Adaptation by Proyecto MEDIR (USAID / Juárez & Asociados), 2004.
- Ferrer, G. (2007). “Enseñanza de la lectura y de la escritura: aportes para un marco teórico-metodológico.” Universidad Peruana Cayetano Heredia, Primera edición.
- Ministerio de Educación. “Estadística y Censo de la Calidad Educativa,” 2014. Available at <http://escale.minedu.gob.pe/web/inicio/padron-de-ieee>.
- Ministerio de Educación. “Resultados por Región/DRE de la Evaluación Censal Estudiantil; 2013,” 2016. Available at https://sistemas15.minedu.gob.pe:8888/evaluacion_censal_publico. Accessed August 24, 2016.
- Murray, Nancy, Larissa Campuzano, Steve Glazerman, and Camila Fernández. “Latin America and Caribbean Reading and Education Evaluation Services: *Amazonía Reads* Evaluation Plan.” Report prepared for USAID. Washington, DC: Mathematica Policy Research.
- National Reading Panel. *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction*. Washington, DC: National Institute of Child Health and Human Development, 2000. Available at <http://www.nichd.nih.gov/publications/pubs/nrp/Pages/smallbook.aspx>. Accessed June 3, 2014.
- Schochet, Peter. “The Late Pretest Problem in Randomized Control Trials of Education Interventions.” Institute of Education Sciences, October 2008.
- United Nations “The Millennium Development Goals Report 2015.” Available at [http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20\(July%201\).pdf](http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf).
- USAID/REAULA. “Evaluación de Lectura para Grados Iniciales.” Misión Guatemala (USAID/G), with Juárez y Asociados, 2013. Unpublished instrument.

This page has been left blank for double-sided copying.

APPENDIX A
ADDITIONAL INFORMATION ON SCHOOL CHARACTERISTICS

This page has been left blank for double-sided copying.

A.1. Study school characteristics, 2014

a. San Martín

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Number of primary grade classrooms in use	8.1	8.0	0.1 (0.821)
Number of students in second grade	20.6	21.0	-0.5 (0.572)
School has multigrade classrooms (percent)	25.0	24.0	1.0 (0.765)

Source: 2014 School Infrastructure Observation, San Martín.

Note: P-values are below the differences. None of the differences were significant at the 5 percent level.

b. Ucayali

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Number of primary grade classrooms in use	8.5	8.8	-0.2 (0.878)
Number of students in second grade	24.6	23.7	0.8 (0.689)
School has multigrade classrooms (percent)	42.9	37.1	5.7 (0.571)

Source: 2014 School Infrastructure Observation, Ucayali.

Note: P-values are below the differences. None of the differences were significant at the 5 percent level.

This page has been left blank for double-sided copying.

APPENDIX B
ADDITIONAL INFORMATION ON TEACHERS CHARACTERISTICS AND
FACTORS TEACHERS IDENTIFY AS LIMITING THEIR ABILITY TO TEACH
CHILDREN

This page has been left blank for double-sided copying.

B.1. Teacher characteristics

a. San Martín, 2014

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Female (percent)	75.0	79.0	-4.0 (0.482)
Age	43.6	42.4	1.2 (0.277)
Years of teaching experience	16.4	16.1	0.2 (0.863)
Years of teaching at current school	9.2	9.7	-0.5 (0.703)
First year at current school	32.0	33.0	-1.0 (0.874)
Tenured	68.0	64.0	4.0 (0.519)

Source: 2014 Teacher Survey, San Martín.

Note: P-values are below the differences. None of the differences were significant at the 5 percent level.

b. San Martín, 2015

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Female (percent)	79.0	76.0	3.0 (0.615)
Age	44.0	45.0	-1.0 (0.417)
Years of teaching experience	16.9	17.4	-0.4 (0.745)
Years of teaching at current school	9.3	9.1	0.2 (0.881)
First year at current school	39.0	28.0	11.0 (0.094)
Tenured	65.0	67.0	-2.0 (0.770)

Source: 2015 Teacher Survey, San Martín.

Note: P-values are below the differences. None of the differences were significant at the 5 percent level.

c. Ucayali, 2014

	<i>Amazonía Lee</i>	Prevailing practice	Difference
Female (percent)	82.9	82.9	0.0 (1.000)
Age	38.9	41.3	-2.5 (0.203)
Years of teaching experience	10.7	14.8	-4.0* (0.014)
Years of teaching at current school	4.1	5.3	-1.3 (0.340)
First year at current school	54.3	45.7	8.6 (0.475)
Tenured	42.9	48.6	-5.7 (0.600)

Source: 2014 Teacher Survey, Ucayali.

Note: P-values are below the differences. Differences that were significant at the 5 percent level are marked with an asterisk.

d. Ucayali, 2015

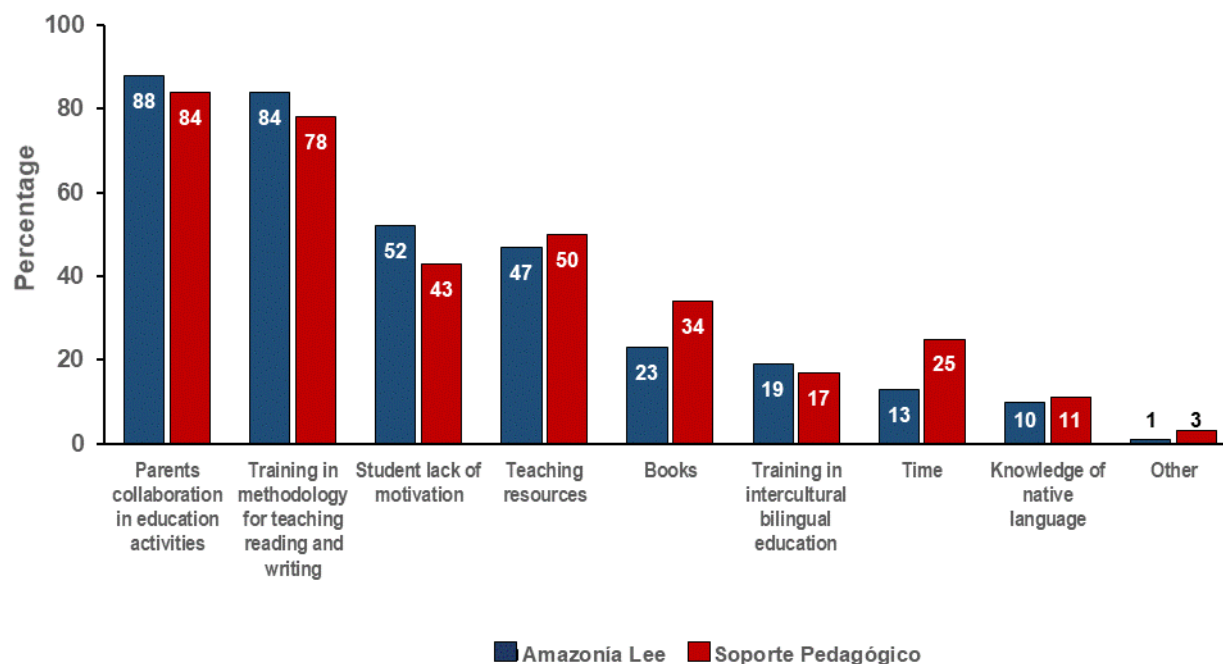
	<i>Amazonía Lee</i>	Prevailing practice	Difference
Female (percent)	74.3	77.1	-2.9 (0.786)
Age	41.4	45.2	-3.8* (0.036)
Years of teaching experience	13.8	14.9	-1.1 (0.572)
Years of teaching at current school	3.8	7.9	-4.2* (0.035)
First year at current school	48.6	37.1	11.4 (0.353)
Tenured	40.0	60.0	-20.0 (0.070)

Source: 2015 Teacher Survey, Ucayali.

Note: P-values are below the differences. Differences that were significant at the 5 percent level are marked with an asterisk.

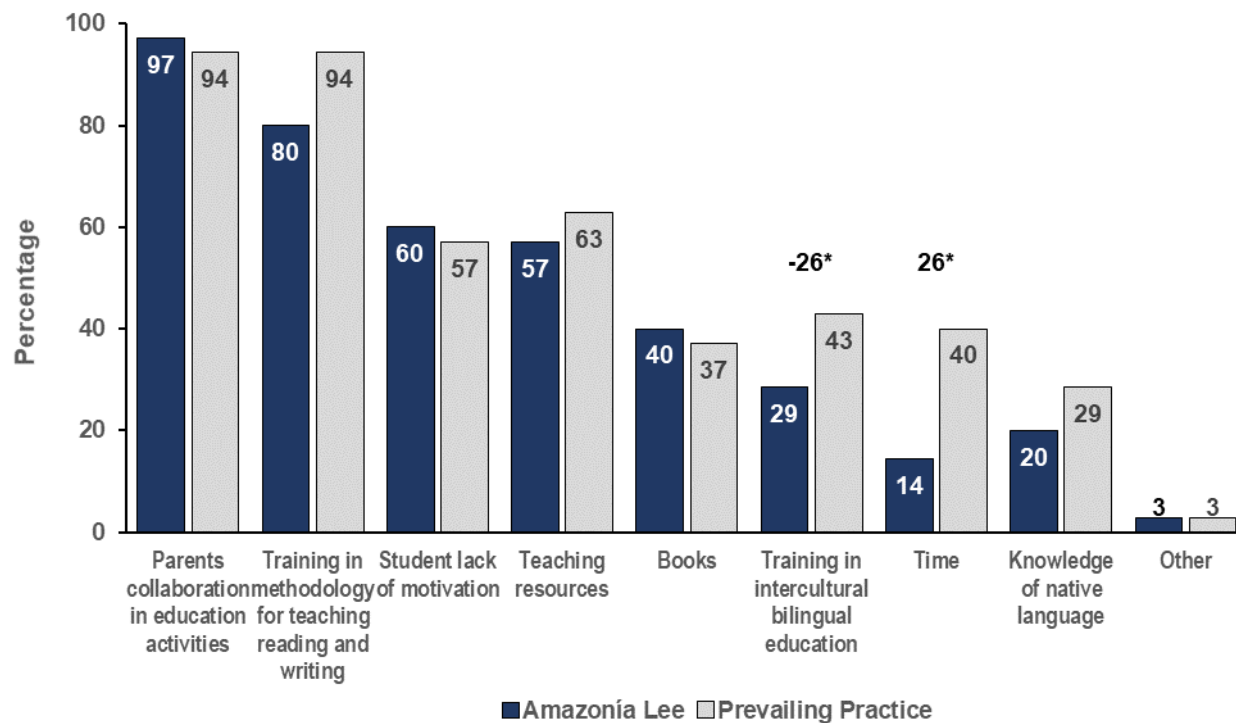
B.2. Factors teachers identify as limiting their ability to teach children to read, 2014

a. San Martín



Source: 2014 Teacher Survey, San Martín.

Note: None of the differences were significant at the 5 percent level.

b. Ucayali

Source: 2014 Teacher Survey, Ucayali.

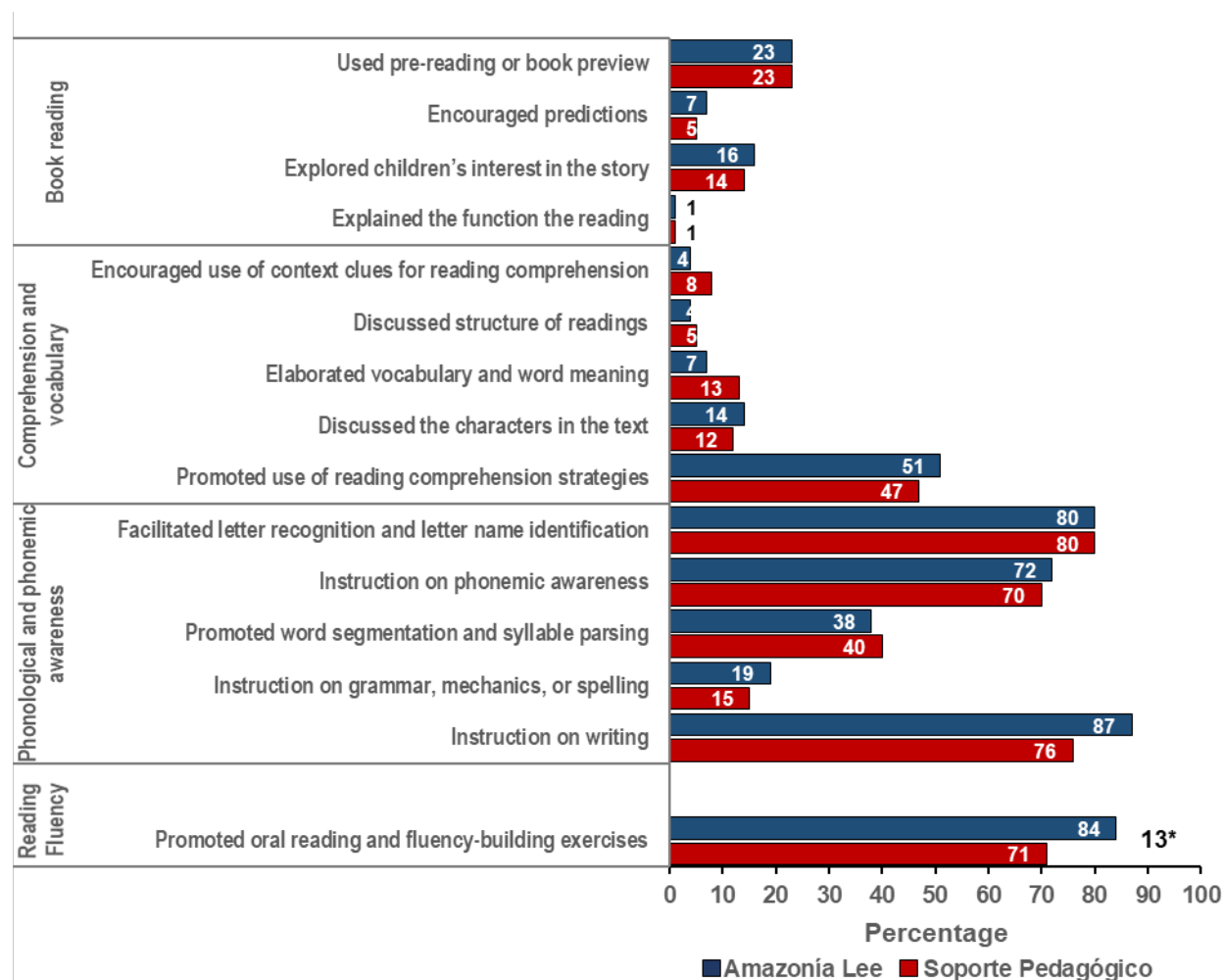
Note: Differences that were significant at the 5 percent level are labeled above each pair of bars, and are marked with an asterisk.

APPENDIX C
ADDITIONAL INFORMATION ON TEACHERS PRACTICES

This page has been left blank for double-sided copying.

C.1. Reading skill-building activities observed by group, 2014

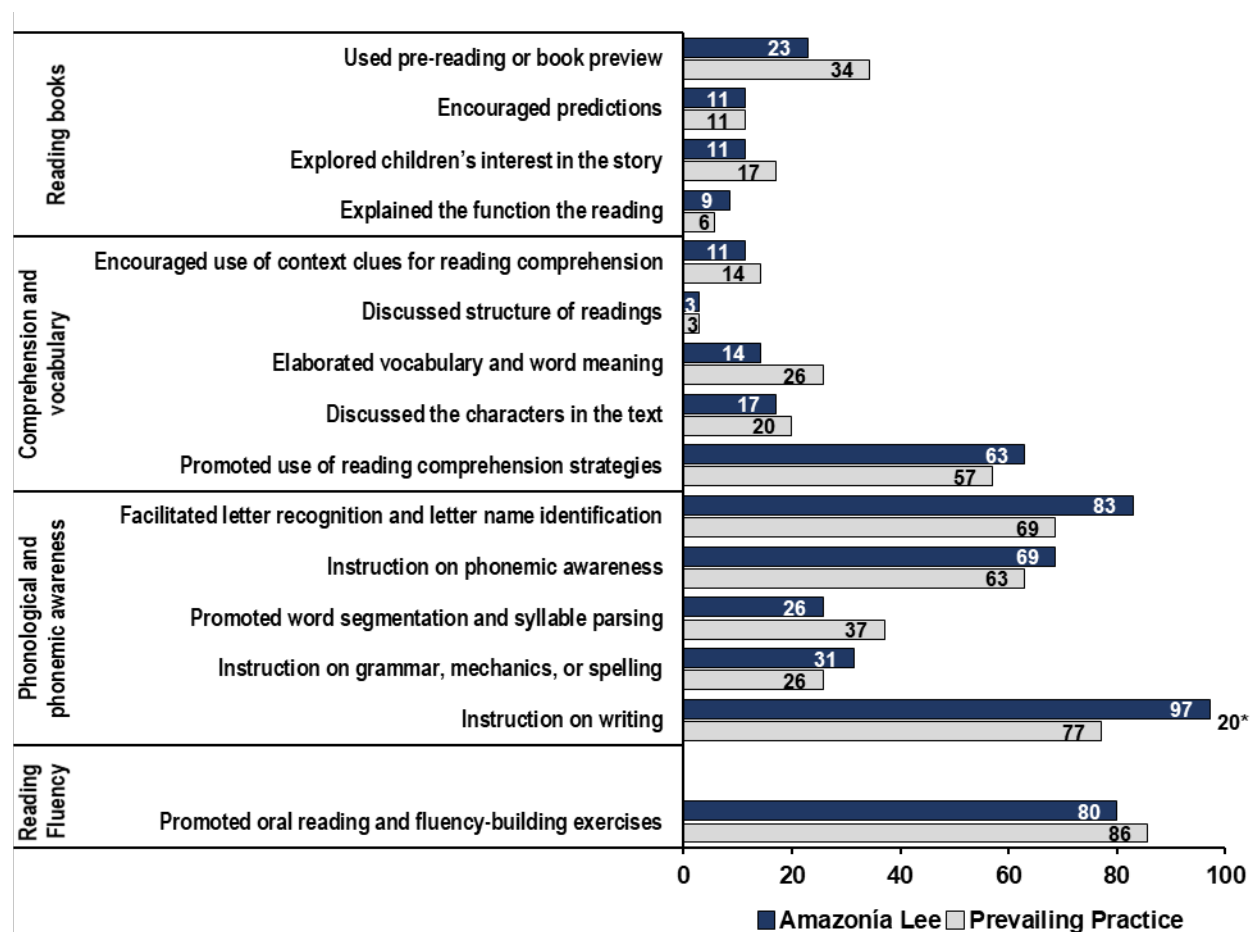
a. San Martín



Source: 2014 1st-grade classroom observations, San Martín. This figure shows the percent of schools in which each activity was used during the time observed.

Note: Differences that are statistically significant at the 5 percent level are labeled next to each pair of bars and marked with an asterisk.

b. Ucayali



Source: 2014 1st-grade classroom observations, Ucayali. This figure shows the percent of schools in which each activity was used during the time observed.

Note: Differences between groups that are significant at the 5 percent level are above the bars and marked with an asterisk.

C.2. Time use in reading lesson by group (minutes), 2014

a. San Martín

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
TOTAL OBSERVED	38.0	37.6	0.4 (0.509)
Written text activities	6.6	6.7	-0.1 (0.937)
Oral language activities	2.1	2.6	-0.6 (0.425)
Phonological or phonemic awareness activities	24.2	23.4	0.8 (0.618)
Vocabulary development activities	0.5	0.5	0.0 (0.972)
Other reading activities	2.4	2.0	0.5 (0.329)
Other activities not related to reading	2.3	2.8	-0.5 (0.478)

Source: 2014 First Grade Classroom Observation, San Martín.

Note: P-values are below the differences. None of the differences were significant at the 5 percent level.

b. Ucayali

	<i>Amazonía Lee</i>	<i>Prevailing practice</i>	Difference
Total observed	39.1	39.4	-0.3 (0.535)
Written text activities	9.5	10.4	-0.9 (0.578)
Oral language activities	1.7	1.3	0.4 (0.604)
Phonological or phonemic awareness activities	24.0	19.8	4.2 (0.062)
Vocabulary development activities	0.9	2.7	-1.9* (0.044)
Other reading activities	1.8	2.1	-0.4 (0.703)
Other activities not related to reading	1.3	3.1	-1.8 (0.099)

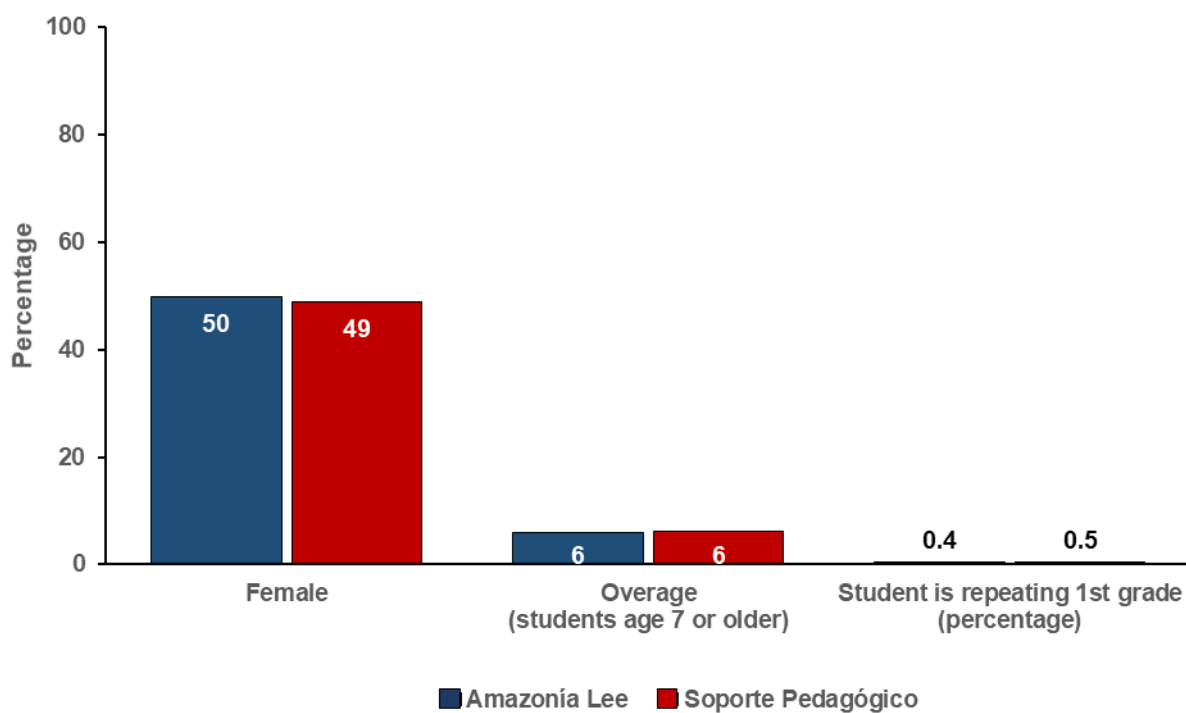
Source: 2014 First Grade Classroom Observation, Ucayali.

Note: P-values are below the differences. Differences that are significant at the 5 percent level are marked with an asterisk.

This page has been left blank for double-sided copying.

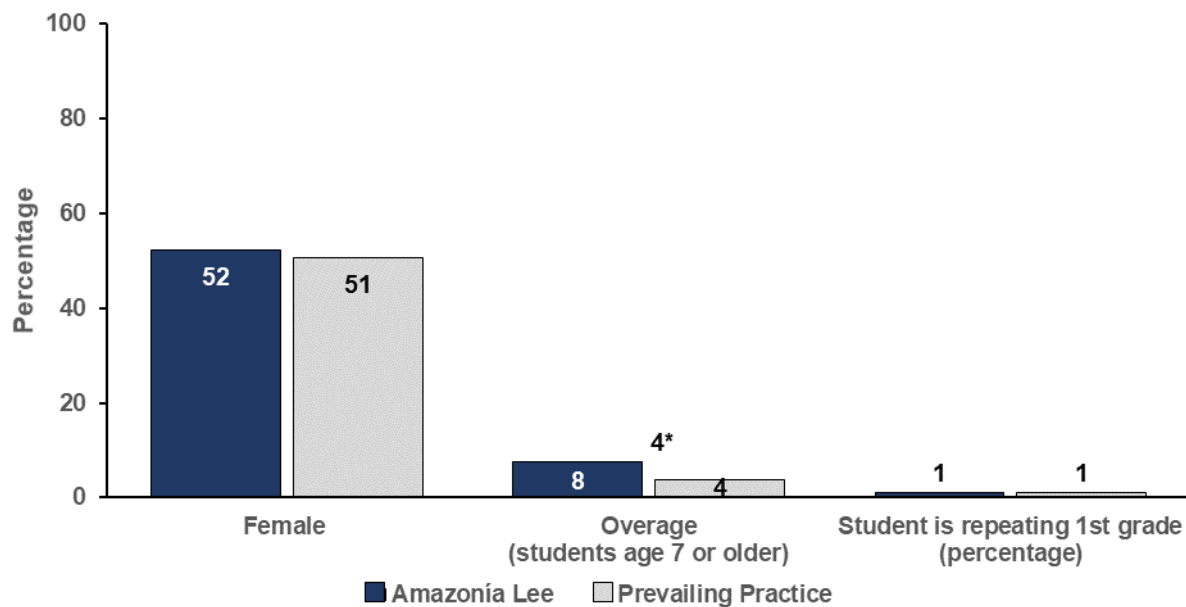
APPENDIX D
ADDITIONAL INFORMATION ON STUDENTS CHARACTERISTICS

This page has been left blank for double-sided copying.

D.1. Student characteristics by group, 2014**a. San Martín**

Source: 2014 1st-grade child assessment data, San Martín. Data are restricted to children who agreed to participate in the child assessment.

Note: None of the differences were significant at the 5 percent level.

b. Ucayali

Sources: 2014 1st-grade child assessment data, Ucayali. Data are restricted to children who agreed to participate in the child assessment.

Note: Differences between groups that are significant at the 5 percent level are above the bars and marked with an asterisk.

D.2. Emergent reading and writing skills by group, 2014

a. San Martín

	<i>Amazonía Lee</i>	<i>Soporte Pedagógico</i>	Difference
Letter identification	18.6	18.6	0.0 (0.932)
Phonemic awareness	3.3	3.3	0.0 (0.693)
Decoded pseudo-words in one minute	2.3	2.2	0.2 (0.481)
Emergent reading	2.5	2.4	0.1 (0.597)
Emergent writing	12.9	12.9	0.0 (1.000)
Passage comprehension	2.9	2.7	0.1 (0.058)

Source: 2014 1st-grade student assessments, San Martín.

Note: None of the differences between groups were significant at the 5 percent level.

b. Ucayali

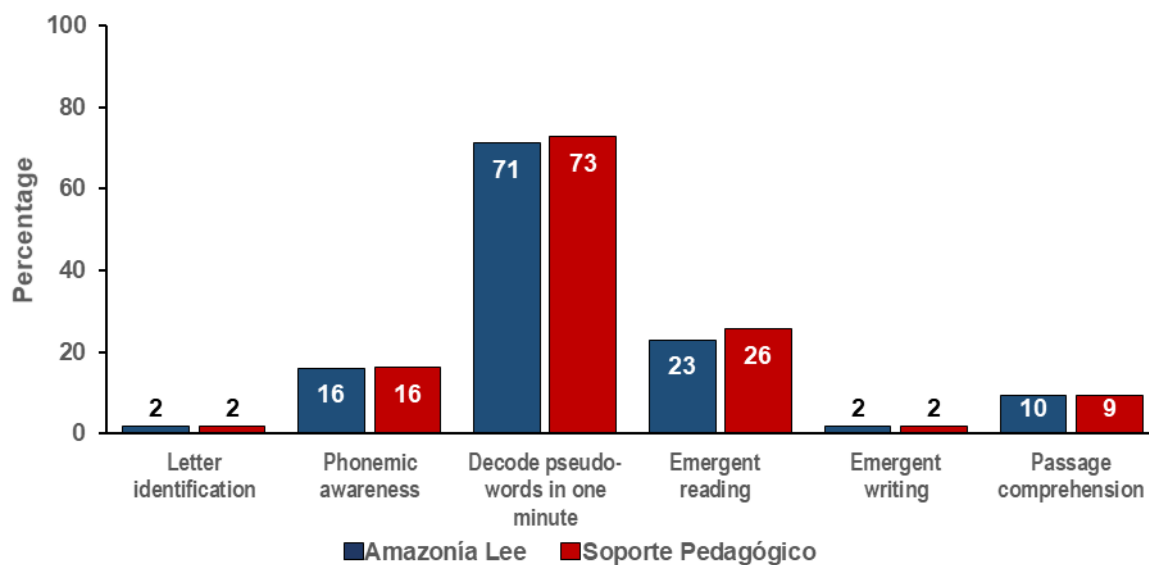
	<i>Amazonía Lee</i>	Prevailing practice	Difference
Letter identification	19.4	19.7	-0.3 (0.521)
Phonemic awareness	3.0	3.5	-0.4* (0.010)
Decoded pseudo-words in one minute	4.4	4.6	-0.2 (0.719)
Emergent reading	3.3	3.5	-0.2 (0.417)
Emergent writing	13.1	13.6	-0.5 (0.336)
Passage comprehension	2.6	2.5	-0.1 (0.488)

Source: 2014 1st-grade student assessments, Ucayali.

Note: P-values are shown below differences. Differences that are significant at the 5 percent level are marked with an asterisk.

D.3. Percentage of students with a score of zero by skill and group, 2014 and 2015

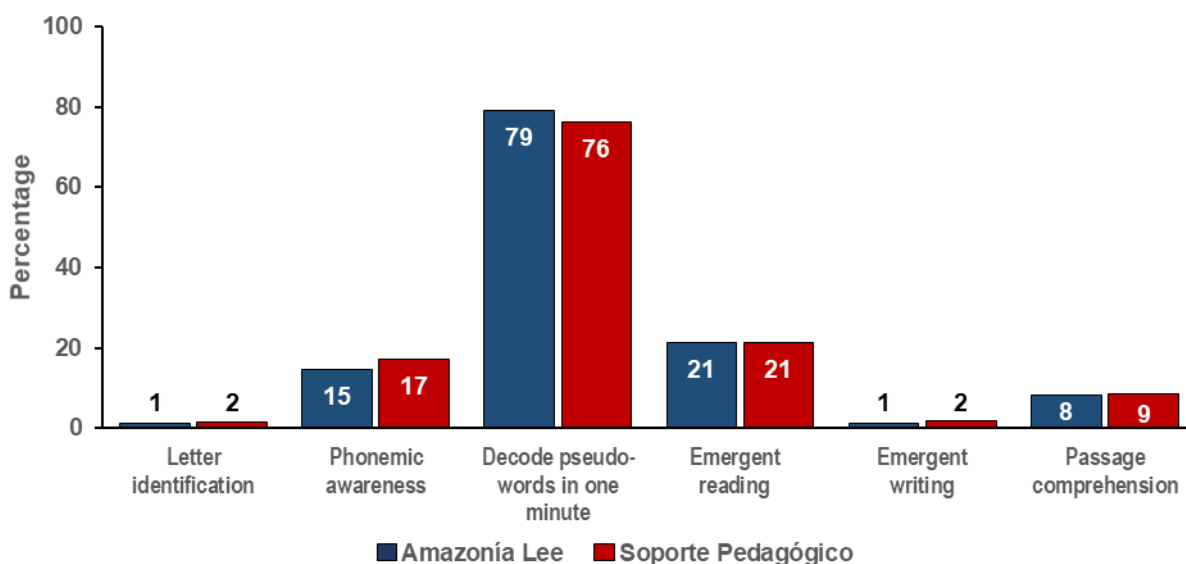
a. San Martín, 2014



Source: 2014 1st-grade student assessments, San Martín.

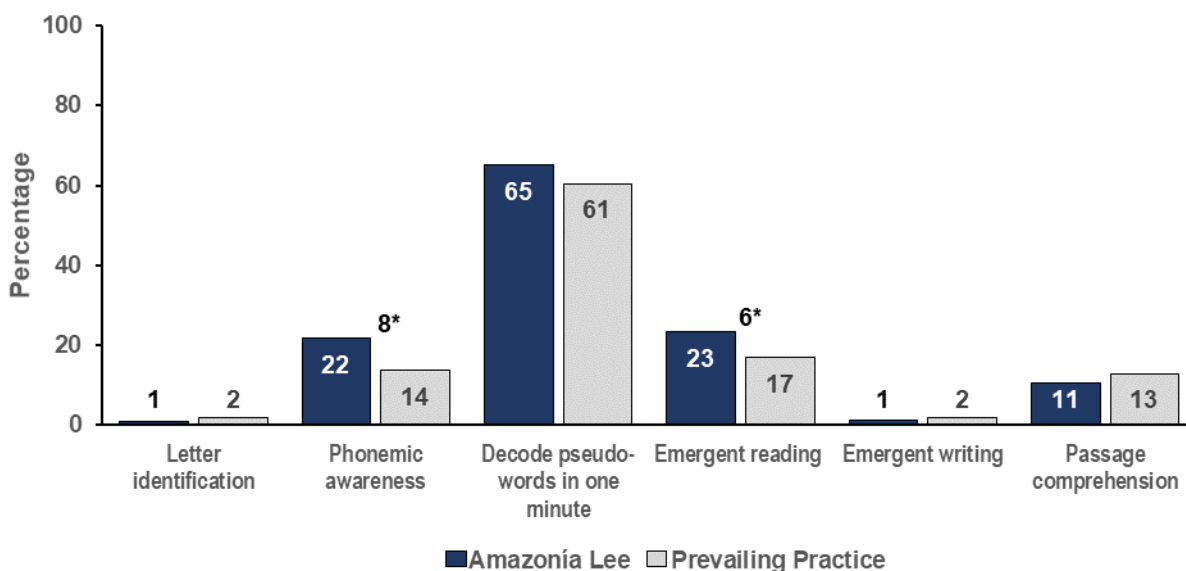
Note: None of the differences between groups were significant at the 5 percent level.

b. San Martín, 2015



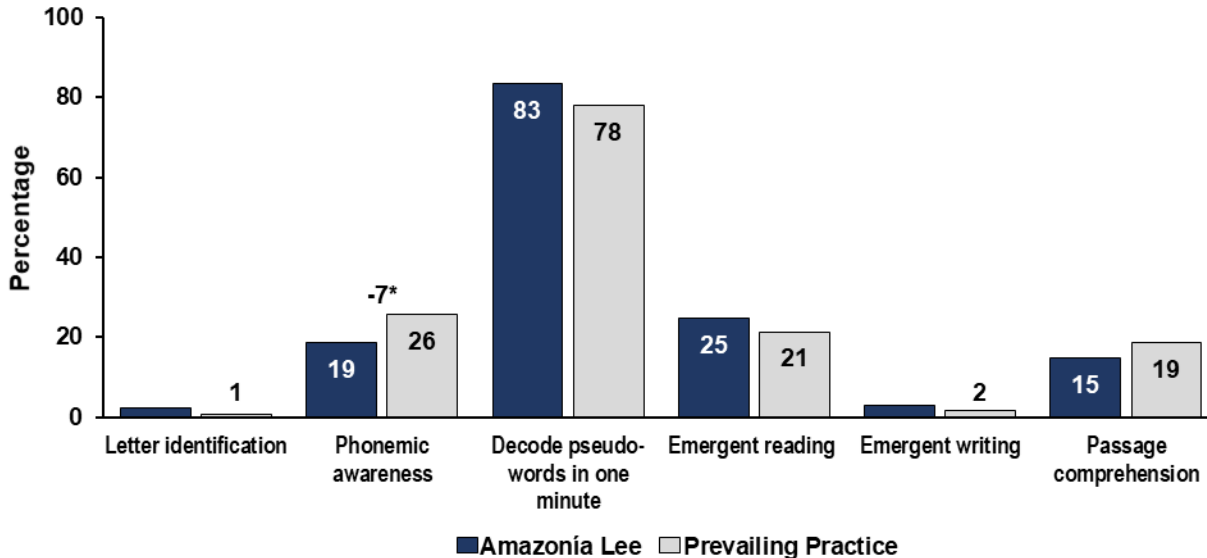
Source: 2015 1st-grade student assessments, San Martín.

Note: None of the differences between groups were significant at the 5 percent level.

c. Ucayali, 2014

Source: 2014 1st-grade student assessments, Ucayali.

Note: Differences between groups that are statistically significant at the 5 percent level are labeled above the bars and are marked with an asterisk.

d. Ucayali, 2015

Source: 2015 1st-grade student assessments, Ucayali.

Note: Differences between groups that are statistically significant at the 5 percent level are labeled above the bars and are marked with an asterisk.

This page has been left blank for double-sided copying

U.S. Agency for International Development

1300 Pennsylvania Avenue, NW

Washington, DC 20523

Tel: (202) 712-0000

Fax: (202) 216-3524

www.usaid.gov