August 2023

Learning at Scale: Time and Money

RTI INTERNATIONAL

RTI Project Number
0216840.000.000

Learning at Scale: Time and Money

August 2023

Prepared for

Center for Global Development
1 Abbey Gardens
Westminster, London
SW1P 3SE
United Kingdom

Prepared by

Christine Harris-Van Keuren
SALT Analytics
On behalf of:
RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709
USA

RTI International is a trade name of Research Triangle Institute.
RTI and the RTI logo are U.S. registered trademarks of Research Triangle Institute.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

cwpm	correct words per minute
EGRA	early grade reading assessment
EQUIP-T	Education Quality Improvement Program in Tanzania
IRC	International Rescue Committee
NEI Plus	Northern Education Initiative Plus
SERI	Scaling-Up Early Reading Intervention
USAID	United States Agency for International Development
USD	US dollar

1. INTRODUCTION[1]

Learning at Scale was interested in programs with demonstrated impact on literacy gains at scale.[2] Part of understanding the success of these programs is to explore the costs incurred and the implementation time invested to generate the programs' impact. In this report, "cost" is defined as program expenditures provided by development partners and monetized contributions donated by the government. With regard to "time," the report uses three different lenses—a program's literacy instruction time, a program's implementation time line, and a government's mandated instruction time. The objectives of this research were to:

- Analyze the cost compositions, as well as the cost per student served and cost per teacher trained, associated with each program to achieve success at an appreciable scale.
- Understand the ways in which the impact of these successful programs is associated with the amount of time that students were exposed to the improved instruction and materials supported by the program.
- Consider the contextual variations in recurrent government expenditures, government-mandated instruction time, and the number of primary school-aged children in each of the countries where the programs were implemented.

To achieve these objectives, the first part of this report sets the stage by highlighting the dynamic and differentiated contexts and varied government structures that these programs worked within. While there are a variety of avenues through which to assess contextual differences, we focus on the differences in planned literacy instruction time in the government school schedule and government spending on basic education. Next, we break down the composition of program costs by the categories of development partner funding and government contributions, the cost per student served, the cost per teacher trained and supported, and the largest type of costs. Due to the confidential nature of these data, project-specific findings are anonymized. The report ends with a discussion of the potential implications for governments wishing to scale and sustain development partner-funded programs.

2. METHODOLOGY

Learning at Scale builds on previous cost analysis research conducted by the Abdul Latif Jameel Poverty Action Lab (n.d., 2016, 2019), the Center for Cost-Benefit Cost Studies of

[1] The author thanks the development partners included in this research, who spent a considerable amount of time collecting, reporting, talking with me about their data, and reviewing this report. It is my hope that they can see their comments reflected in this work. Also, thank you to RTI International for its tireless support for this research.

[2] Some of the programs included in the Learning at Scale research were repeated cross-sectional studies with no controls. We use the term "impact" loosely throughout this report to describe the different designs.

Education (Levin, 1975; Levin et al., 2018), Patrick McEwan (2012, 2015), and USAID Education (EducationLinks, n.d.). This research uses an adapted version of the United States Agency for International Development's (USAID) *Cost Reporting Guidance for USAID-Funded Education Activities* (Walls, 2018) and *Cost Analysis Guidance* (Walls, Tulloch, & Harris-Van Keuren, 2021) as its methodological framework. USAID's reporting and analytical protocol provides a systematic framework for USAID evaluation partners, partner organizations, and USAID missions to conduct cost analysis studies. This guidance was selected because of its rigor, specificity for large international education programs, and step-by-step transparency. Additionally, five of the eight Learning at Scale programs were funded in whole or in part by USAID.[3] As a result, there was a higher likelihood that the programs would be familiar with their funder's cost methodology. This familiarity would potentially minimize the program staff's learning curve and time burden, while increasing the accuracy of the submitted data.

To collect the necessary data, Learning at Scale used an adapted version of USAID's *Cost Reporting Guidance*. While a hallmark of this guidance is to collect data concurrent to implementation, this was not possible for all of the programs included in this research. This is due to the timing of the research, as well different forms of cost collection requirements from various development partners. The data used for this portion of the Learning at Scale research were obtained from partner organizations' program reports, through desk research, and by requesting cost information from partner organizations through customized data collection templates. Evaluation and descriptive data were obtained from partner organizations' midline or endline reports and monitoring and evaluation documents. Information on government and program time was gathered through desk research, partner organizations' formative and summative impact reports, and email correspondence, phone calls, and meetings with partner organizations.

All data collected for this analysis align with programs' midline or endline impact assessments highlighted in the *Learning at Scale: Interim Report* and the program summaries in Section 5 of this report. Imagine a stopwatch that begins when a partner organization starts its program implementation and ends when the impact midline or endline is conducted. All descriptive and cost data fall within this very specific time frame. We recognize that in some instances, this may measure only a portion of a program's overall implementation. However, the intent is to analyze the costs borne during the portion of the program containing the impact assessment.

To collect data on the costs of resources and staff time, students served, and teachers trained and supported, Learning at Scale created a customized Excel workbook for each program. The cost categories and ingredients included in the Learning at Scale workbooks

[3] This includes Chemonics' Lecture Pour Tous, Creative Associates' Nigeria Education Initiative Plus, FHI360's Ghana Learning, the IRC's Pakistan Reading Project, RTI International's Tusome, and, in part, Room to Read's Scaling-Up Early Reading Intervention.

were similar to those used by USAID. The workbooks were sent to each program for completion and were then returned for analysis.

After the findings were drafted, the workbooks were resent to the partner organizations to adjust and correct as needed. Phone calls, virtual meetings, and email correspondence were also conducted to answer questions and clarify points.

The cost data included expenditures incurred by the development partners supporting each program and an estimated cost for the resources provided by the government. The purpose of including the costs borne by the government is to better understand what resources the government provided to a given program versus which resources were funded by development partners that augmented the government's existing system. Not all of the programs were able to include monetized government contributions. Therefore, the division between expenditures and monetized contributions is presented only when both types of data were submitted.

The cost estimates presented do not include development or overhead costs, or costs that fell outside a program's theory of change as it relates to improving student learning outcomes. These costs were purposefully omitted due to the research focus of this study. Therefore, the cost estimates presented should be read as being underestimated from the actual costs incurred by partner organizations. The cost data were converted into 2021 US dollars.

Partner organizations received copies of the drafted report to review and provide feedback at least twice. Adjustments were made accordingly to increase the accuracy and clarity of the report. See Section 10 for a complete list of documents and customized worksheets used in this analysis.

There are several notable limitations to this research.

- Some data were collected retrospectively. Over time, project managers leave; monitoring, evaluation, and learning experts move to different countries; and finance staff focus on new programs. There is an astonishing amount of institutional knowledge that becomes lost at the close of a project. While this research doesn't attempt to address this data loss across programs funded by different development partners, we do recognize that some of the data used in this research include an unmeasurable margin of error.
- Internal accounting systems may not be designed to complement or align with monitoring and evaluation, or impact assessments. Because the programs included in Learning at Scale are sophisticated in their design, it was difficult for some partners to isolate costs to one component or, more specifically, to one grade in one component of a multifaceted program. This is due in part to internal accounting systems that are not designed to complement or align with monitoring and evaluation, or impact assessments. Therefore, while the accounting systems house some of the data needed for cost analysis, they are often structured in such a way that it is difficult and labor intensive to identify the precise data needed for a given cost analysis.

- Monitoring and evaluation systems are not designed to align with cost analysis. Fidelity of implementation studies provide valuable information for cost analysis research. For example, understanding teacher uptake can provide greater specificity regarding not only the cost per teachers trained but also the cost per teachers who implemented the program's pedagogy. These are critical distinctions that can be used to inform policy and practice. Furthermore, documenting resources that contributed to programs, specifically human capital contributions by the government and society, is still a novel concept in the education sector. We, as a sector, justifiably obsess over the sample sizes and statistical models used to measure impact. However, for some implementing partners, it can be more difficult to obtain seemingly basic information, such as the unique count of students served and teachers trained, given the complexities of the programs and the changes that occur in those programs over time.
- Government systems impact partner organization spending and cost analysis results. Government systems impacted Learning at Scale in three ways. First, this research would have benefited from government data on education budgets and expenditures, and general descriptions of the resources currently provided in a country. However, these data are often inaccurate, not publicly available, or out of date. Second, governments often set parameters that can drive up the cost of a program. For example, in-country printing mandates may increase the cost of teaching and learning materials, government-initiated program reforms can drive up the overall cost of a program, and government-set teacher training per diems can quickly inflate the teacher training costs incurred by partner organizations and, consequently, development partners. Third, this research does not attempt to estimate the cost of the programs when adopted by the government. The structure, costs, and impact of a program could change once the government has integrated it into its existing structure.

See Appendix A for information on data collection, cleaning, and analyzing.

3. UNDERSTANDING THESE FINDINGS

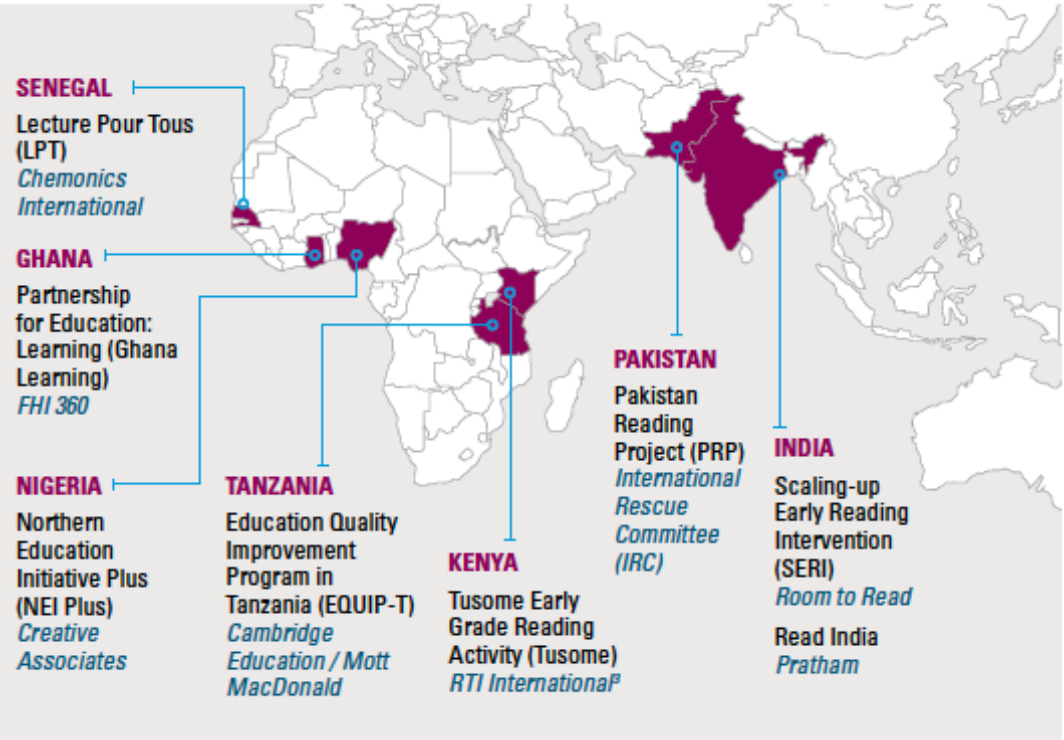
This is not the World Cup, where the most successful at-scale programs compete to determine a "winner."

Before presenting the findings, it is critical to provide a framing. This report seeks to demonstrate the contextual, structural, and financial differences between the successful programs included in this research. The purpose of comparing these programs is not to apply judgment values and determine which program was the most cost-effective. Instead, it is to show how success can look different in different spaces. This is no easy task. There is a justifiable fear that casual readers will compare these programs and misinterpret the findings, which could lead to increased confusion in the field and unintentional harm to the organizations involved. The foundation of this research is *primum non nocere*—first, do no harm. To this end, we provide a general framing for how to look at these findings.

It can be helpful to liken the programs included in this research to the construction of eight different houses in eight different countries.[4] While each house may have the same fundamental features, such as bathrooms, kitchens, bedrooms, plumbing, electricity, and so on, it's the details about each house, the regulations the construction company follows, the cost of the resources, and the location-specific factors (among other variables) that contribute to the costs. Furthermore, because some of the programs had been implemented and evolving for years prior to this analysis, they were at different phases. Some resembled "new construction," while others were conducting "minor renovations," "gut renovations," or even "additions." The programs included in this study all focused on primary school literacy and had some of the same components, such as teacher training, teaching and learning materials, and teacher support. However, they were implemented in different countries, reached different numbers of teachers and students, provided instruction in different languages, required different levels of support staff to implement and monitor them, and were customized for the contexts in which they worked. In addition, the cost of resources differs depending on the context. A local monitoring and evaluation person may cost more in Nigeria than in India due to supply and demand. Furthermore, the programs were all at different stages of maturity; some still needed refining while others were well established. That said, the purpose of this analysis is to look at the costs of the different programs, in an anonymized fashion, and consider the largest costs and, when possible, the cost of the resources contributed by the funding agency as compared to the local government. These findings should not be used to assess whether one program was more or less "cost-effective" than another. Indeed, even within a given context for a specific program, there is a non-linear relationship between costs and impact. Therefore, care should be taken not to assume that increased costs would be associated with concomitantly increased impact.[5] At the time of this writing, as shown in Figure 1, we had eight data points on the global map estimating the costs of programs being implemented at scale. As more programs collect and report on costs as well as other valuable information such as time, the data will become nuanced and rich. For now, these findings can be used to help the education sector understand why seemingly similar programs may use different amounts of resources and can help us assess the potential strengths and weaknesses of such programs' scalability and sustainability.

[4] A similar analogy is provided in Walls, Tulloch, & Harris-Van Keuren (2021).
[5] In fact, some cost analyses round up the results to the nearest $100 spent and adjust the impact to align with the cost. This is erroneous, as there is not a linear relationship between cost and impact.

Figure 1. Data points for at-scale programs



4. TIME, MONEY, AND MAGNITUDE

Government school systems are dynamic and have differentiated structures.

The purpose of this section is to demonstrate the differences in contexts where the various programs worked. Here we look at the differences in "time, money, and magnitude." Time refers to the amount of literacy instructional time planned (see notes on this below) by each country for primary school education. Money is an estimate of the government's per-student expenditures for primary school, and magnitude is the estimated number of students enrolled in government-run primary schools. An additional consideration is whether the school system provides instruction in a primary and/or secondary language.

In short, time, money, and magnitude frame each context as "the government allows this much time and this much money to help this many primary school-aged children in public government schools achieve literacy." The information shown in this section represents recent information on the systems in which the programs worked. It is not a historical perspective depicting how the government systems looked when the programs were being implemented. We use current-day depictions for two reasons. First, government systems

were disrupted in an unprecedented manner due to the COVID-19 pandemic, and attempting to compare disrupted systems falls outside the scope of this research. Second, assessing different systems in different time periods is less helpful than comparing different systems across the same period of time.

This information is helpful in at least three ways. First, it seeks to put development actors in the seat of Ministry of Education officials by providing a glimpse into the challenges that governments face. Second, it visually depicts the differences among the countries in these three areas (i.e., time, money, and magnitude), and third, it provides some parameters that future programs can consider when constructing their interventions for scale and sustainability. A program's ability to meet time, money, and magnitude thresholds provides some evidence of the government's potential challenges in adopting and scaling successful interventions.

While each of the Learning at Scale programs worked within government-run primary schools, what is considered "primary school" varies by country and is subject to frequent changes. At a fundamental level, this means that what is understood as "primary school" in one context can be different than what is considered "primary school" in another context (see Appendix B). However, the differences among the systems become more pronounced when we take a deeper look at government-mandated literacy instructional time.

To compare the differences within and between government systems, we looked at the intersection of primary school literacy instructional time planned by the government, per-student expenditures for primary school, and the number of primary school-aged children. Planned primary school literacy instructional time mandated by the government was calculated as:

$$= number\ of\ years\ of\ primary\ school * number\ of\ instructional\ weeks\ per\ academic\ year * number\ of\ literacy\ lessons\ per\ week * number\ of\ hours\ of\ instruction\ per\ lesson$$

See Appendix C for these calculations.

Per-student government expenditures for primary school were calculated as:

$$= \frac{\left(country's\ 2021\ GDP * education\ spending\ as\ a\ \%\ of\ GDP * \%\ of\ education\ expenditures\ on\ primary\right) * (years\ of\ primary\ school)}{total\ primary\ enrollment\ in\ government\ schools}$$

Source: World Bank Education Statistics (n.d.)

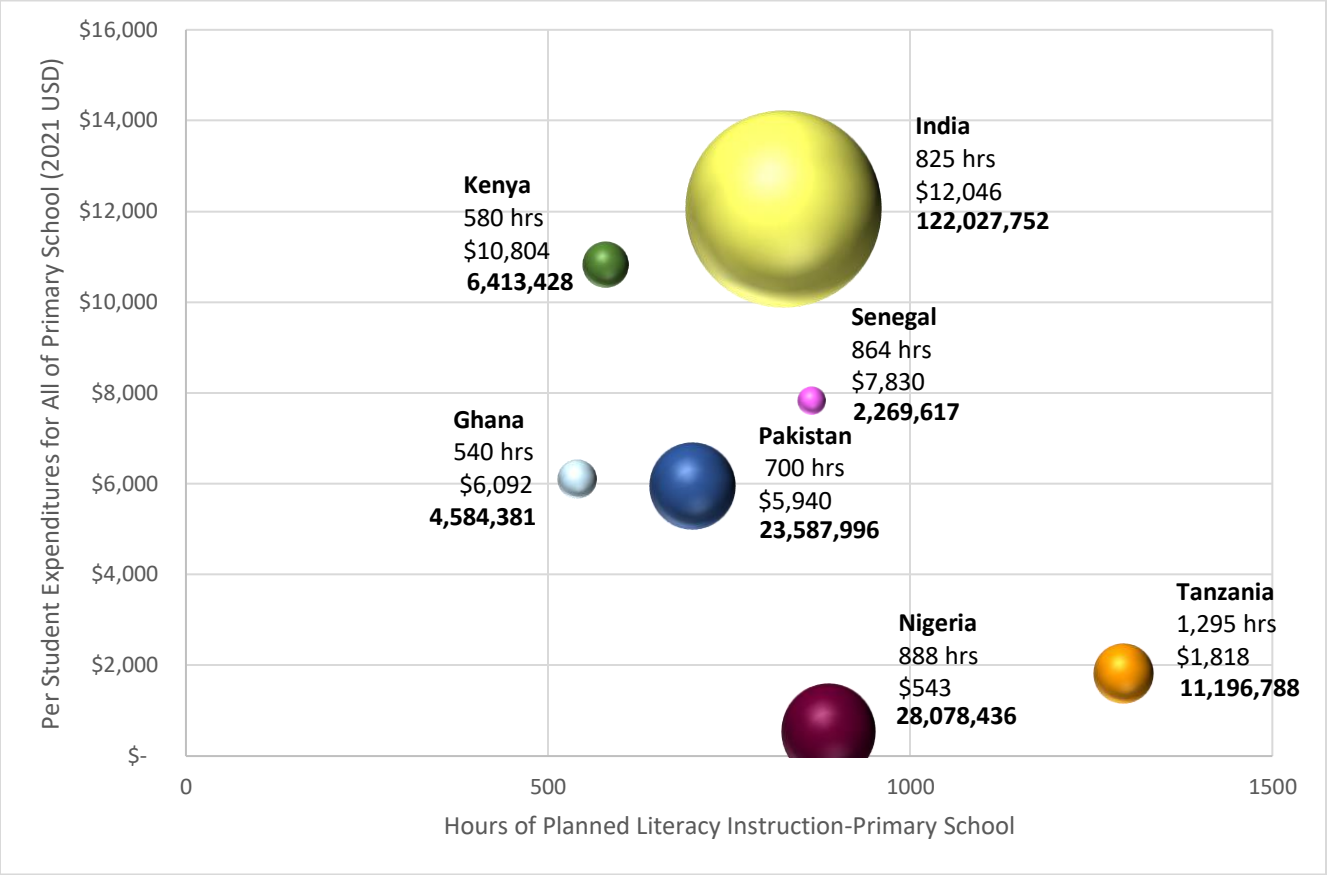
While recurrent student expenditures per country are available for some countries through UNESCO, not all are. Further, timely and reliable data on government budgets and primary school enrollment are often not available on ministry websites. Therefore, to avoid

generating estimates using different methods and sources, we chose to use a single method to increase comparability.[6]

The results of this analysis, shown in Figure 2, demonstrate the dramatic differences among countries' education systems. The most noticeable finding is that all of the systems included in this research have 500 hours or more of planned literacy instruction for the duration of primary school. It also shows that systems that have a similar overall structure can be very different in terms of their details. For example, Ghana, Kenya, Nigeria, and Senegal all dedicate six years to primary school education. However, during those six years, the Kenyan government allocates 580 hours for literacy instruction in primary school (for Kiswahili), Ghana allocates 540 hours, Senegal sets aside 864 hours of planned instruction. Tanzania assigns 1,295 hours of planned instruction, but has seven years of primary school, more than the other countries included in this study. However, this figure should be considered dynamic, as governments change the details of their school days regularly. Over the course of a multiyear contract, governments may adjust the number of academic days per year, literacy lessons per week, and minutes of instruction per lesson, thereby making the context in which the partners work fluid. The pedagogical design of a program may need to be adjusted due to shifts in government-mandated instruction times and frequencies. These changes can impact the costs incurred by partner organizations and development partners. Additionally, although time is government mandated, this does not mean that the mandate is adhered to in the classroom. In fact, the World Bank (2018) reports that the percentage of time teachers spend at school and teaching can drop to less than 40% of the planned instruction time. Furthermore, the time spent instructing does not address the content or quality of instruction (see also Beggs & Ogando Portela, 2019). At a fundamental level, the government-mandated literacy instruction time for primary school provides a window into understanding the time differences allocated for primary school students to achieve fluency, and an understanding that these estimates are high predictions for what actually occurs in the classroom.

[6] Thank you to Gustavo Arcia for his thoughtful comments on this methodology. Any errors remain my own.

Figure 2. Time, money, and magnitude by national education system (in 2021 USD)



Sources: World Bank (2015, 2018, 2021); World Bank Education Statistics (n.d.); UNESCO Institute for Statistics (2021); partner organizations' endline and midline reports and Learning at Scale data collection templates for each program referenced in Section 10 of this report

Figure 2 also shows an estimated per-student financial allocation for primary school by country. The expenditure figures shown here are not per year; instead, they represent the *estimated amount a government will invest in each primary school student for the duration of that student's tenure in primary school*. For instance, India spends an estimated $12,046 per student over the course of the student's five years in primary school. Compare this amount to Nigeria, which invests about $543 per primary school student over the course of their six years in government primary schools. These per-country estimates are inclusive of all primary school inputs (e.g., capital, teaching and learning materials, and in-service and pre-service training) and human capital (e.g., national and local government salaries, teachers' salaries, and so forth) for all academic subjects, not just literacy. Combining all costs, including human capital, into cost estimations is critical as more programs scale and integrate into existing government systems by utilizing the time of national and local

ministry officials, as well as head teachers and primary teachers. This research sought to estimate the cost of government officials' and teachers' time invested in each program to better understand the true cost of each program's implementation.

Figure 2 also shows the number of primary school-aged children receiving public school education.[7] In India, there are over 122 million such students. Senegal has a fraction of this primary school-aged population, at around 2.2 million. Taken together, these data points show some of the challenges facing ministries of education and those programs that seek to reach scale.

Using Ghana as an example, the findings in Figure 2 can be summarized as follows:

Ghana—There are an estimated 4.6 million primary school students in government schools. For the six years of government primary school, there are approximately 540 hours of planned literacy instruction scheduled. For the duration of primary school, in all, the Ghanaian government invests about $6,000 in each student. This amount includes all inputs and human capital, not just the amount invested per student for literacy instruction.

This information helps us understand some of the parameters to be considered when assessing if a program is sustainable and scalable.

These per-student expenditures and hours of planned literacy instruction can be rolled back to the end of grade 2 and displayed with the average grade 2 early grade reading assessment (EGRA) outcome of the number of correct words per minute (cwpm).[8, 9] We chose grade 2 since most of the programs included in this analysis conducted their endline impact estimations in this grade.[10]

These costs and time investments are relative to the expected achievement levels of students. For example, Table 1 shows that the Indian government invests about $4,800 per student through the end of grade 2 and plans for 330 hours of instruction to obtain 17 correct words per minute in Hindi. Conversely, Nigeria spends approximately $180 and nearly 300 hours for barely 3 cwpm. However, Ghana has the lowest results, with slightly more than 2 cwpm per student by the end of second grade after 180 hours of planned literacy instruction and over $2,000 per student. Tanzania fares better, with students

[7] These student enrollment counts do not include private school or nonformal education.
[8] The oral reading fluency results used in Table 1 were obtained from the programs' formative or summative assessments documenting the comparison group's midline or endline cwpm results. For those programs that did not have a comparison group, the results from the treatment group's baseline cwpm were used. This follows the statistical models used in the impact evaluations.
[9] This assumes that the costs for each grade are the same (i.e., grade 2 is one-third the cost for six years of primary).
[10] Donors, such as USAID Education, focus on grade 2 for impact assessments due to Sustainable Development Goal indicator 4.1.1, which is defined as the "proportion of children and young people (a) in Grades 2 /3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex."

reading on average 21 cwpm. The government of Tanzania allocates 370 hours of planned instruction and $500 per student for these results.

Table 1. Countries' planned literacy instruction time, per-student expenditures, and oral reading fluency outcomes through the end of grade 2

Country	Hours of planned literacy instruction	Government's per-student expenditure	Endline oral reading fluency for non-intervention group (cwpm)
Ghana (Ghanaian languages)	180	$2,031	2.35
India (Hindi)	330	$4,819	17.50
Kenya (Kiswahili)	193	$3,608	13.50
Nigeria (Hausa)	296	$181	3.00
Pakistan (Urdu)	280	$2,376	26.05
Senegal (Wolof)	288	$2,610	3.40
Tanzania (Kiswahili)	370	$519	21.33

Sources: World Bank (2018); World Bank Education Statistics (n.d.); UNESCO Institute for Statistics (2021); see also partner organizations' endline and midline reports and Learning at Scale data collection templates for each program referenced in Section 10 of this report

There are at least four takeaways from these results. First, countries are falling far short of fluency benchmarks for second grade despite investments in time and resources. Second, except for Nigeria, all of the countries included in this study are investing over 4% of their GDP on education, with the lion's share going toward primary education.[11] However, the sector lacks rigorous evidence to support the claim that an investment of 4–6% of a country's GDP is adequate across varied contexts. Third, implementing partners in the education sector could provide valuable evidence to governments demonstrating how to generate greater outcomes and higher impacts with restricted budgets that are context specific. In fact, spending could be more purposeful and customized to a particular system's prioritized needs, and this customization could be more explicitly described by the implementing partners.

Finally, and more broadly, we as a sector have a misalignment between the programs we implement, the dialogue we engage in, and the needs we are trying to address. Our programs typically focus on literacy gains at very specific grades, and our dialogue focuses

[11] UNESCO (2016) recommends that countries allocate at least 4–6% of their GDP to education.

on scaling solutions for those particular grades. However, the need is wider than simply achieving literacy rates in grade 2 or grade 3. The need, which as we have shown is different by context, is to achieve levels of literacy commensurate with what is needed to apply to all future learning. We argue that the sector should expand its focus on achieving foundational literacy goals throughout the duration of primary education. This is discussed further in Section 8 of this report.

5. LEARNING AT SCALE PROGRAMS: BRIEF RECAP

When assessing the success of these programs, also consider the activities that took place prior to the program included in the Learning at Scale research.

This section provides a brief recap of the programs included in this research. For each program, we provide descriptors and critical components for the specific time frame being analyzed. Readers can compare the programs side by side to see the important differences among them. For example, the programs differed in scale; in the languages assessed; in their literacy fluency thresholds; and in the grades measured in the endline or midline assessments. Lastly, we include important points about the unique aspects of each program. These points, which also serve as descriptors, demonstrate why each program was successful in its specific context. However, the success of many of the programs highlighted in this report is not due solely to the work conducted in the span of five or six years. Indeed, many of these programs spent years, if not a decade or more, of relationship building and government inclusion in decision-making processes, as well as refining implementation models. Referring back to the house analogy, some programs might have been doing renovations while others were in the process of initial construction, which requires larger capital outlays. Therefore, when considering the success of each program, we should also consider the work that occurred prior to the specific programs included in the Learning at Scale research. In some cases, the success of a program was incremental and built upon over time and through activities. When applicable, we note previous programs that might have served as building blocks to the program in question. Finally, some of the programs continued beyond the scope of this analysis and may reflect lower numbers of students served, schools reached, and teachers trained than they ultimately achieved. School, teacher, and student counts, as well as impact findings, are rounded to increase readability.

EQUIP-T: Education Quality Improvement Program in Tanzania

Tanzania
Cambridge Education[12]

Development partner: UK Department for International Development[13]
Years of the activity included in this research: 2014–2020
Supplementary to the government's official public school literacy instruction: No
Primary schools reached: 5,200
Students reached (grades 1–7): 4.9 million
Teachers reached: 54,000
Government system has primary and secondary language instruction in primary school: No
Government-recognized languages of instruction: 2 (Kiswahili and English)
Teacher training and support: Moved from a cascade model to reliance on communities of learning that are focused less on centrally developed modules and materials and more on peer learning and teacher-led identification of issues and discussion on how to teach different competencies.
Impact measured at grade: 3
Language(s) measured: Kiswahili
Planned instructional dosage: Not reporting[14]
Language fluency threshold: 50 cwpm at the end of grade 2 or grade 3
Endline impact: 14 percentage point increase from baseline in the number of students who reached the fluency threshold in Kiswahili
9 cwpm average increase in Kiswahili

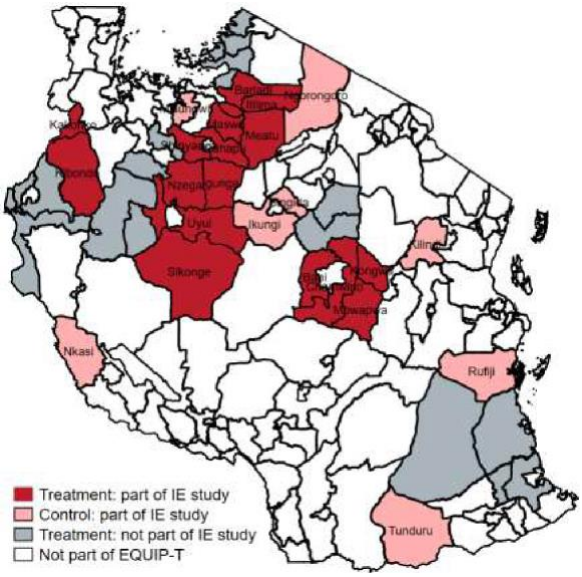
Unique Aspects of This Program

- Originally scheduled to run for four years but later expanded to six years, with activities continuing in a seventh year.
- The program included interventions at the systems level, such as budget support to the government.
- There was notable government involvement at the national and regional level.
- EQUIP-T gave money directly to the government structures it worked within. This created a higher degree of access to local government authorities and other Tanzanian government leaders.
- Each program's technical tasks had an institutional home within the government so that local-level implementation issues could be solved quickly.

[12] See Section 10 of this report.
[13] Now referred to as Foreign, Commonwealth & Development Office.
[14] We are not reporting the dosage, as I believe that the estimates are inaccurate.

- A cost analysis was conducted on this program in 2019.[15]

Implementation Map[16]



[15] All Cambridge Education reports are available at https://www.opml.co.uk/projects/assessing-equip-t.

[16] "This map applies to the EQUIP-T programme coverage prior to the extension in 2017. The programme extension covers districts in Singida region (including Ikungi DC and Singida DC which are control districts in the impact evaluation) and Katavi region" (Rawle et al., 2019a, p. 9).

USAID Partnership for Education: Ghana Learning

Ghana
FHI 360[17]

Development partner: USAID
Years of the activity included in this research: 2016–2019
Supplementary to the government's official public school literacy instruction: No
Primary schools reached: 7,400
Students reached (grades 1 and 2): 708,000
Teachers reached: 51,000
Government system has bilingual instruction in primary school: Yes (Ghanaian languages and English)
Government-recognized languages of instruction: 11 (Akuapem Twi, Asante Twi, Dagaare, Dagbani, Dangme, Ewe, Fanti, Ga, Gonja, Kasem, and Nzema)
Teacher training and support: Cascade model; in-school (i.e., head teachers or curriculum leads) and external (District Teacher Support Team) coaches. In the final year, teachers received certificates from the National Teaching Council, which were included in each teacher's record for career development.
Impact measured at grade: 2 Ghanaian languages
Language(s) measured: 11 Ghanaian languages
Planned instructional dosage: 300 hours
Language fluency threshold: 35 cwpm at the end of grade 2 or grade 3
Endline impact: 18 percentage point increase from baseline in the number of students who reached the fluency threshold
9 cwpm average increase across Ghanaian languages of instruction

Unique Aspects of This Program

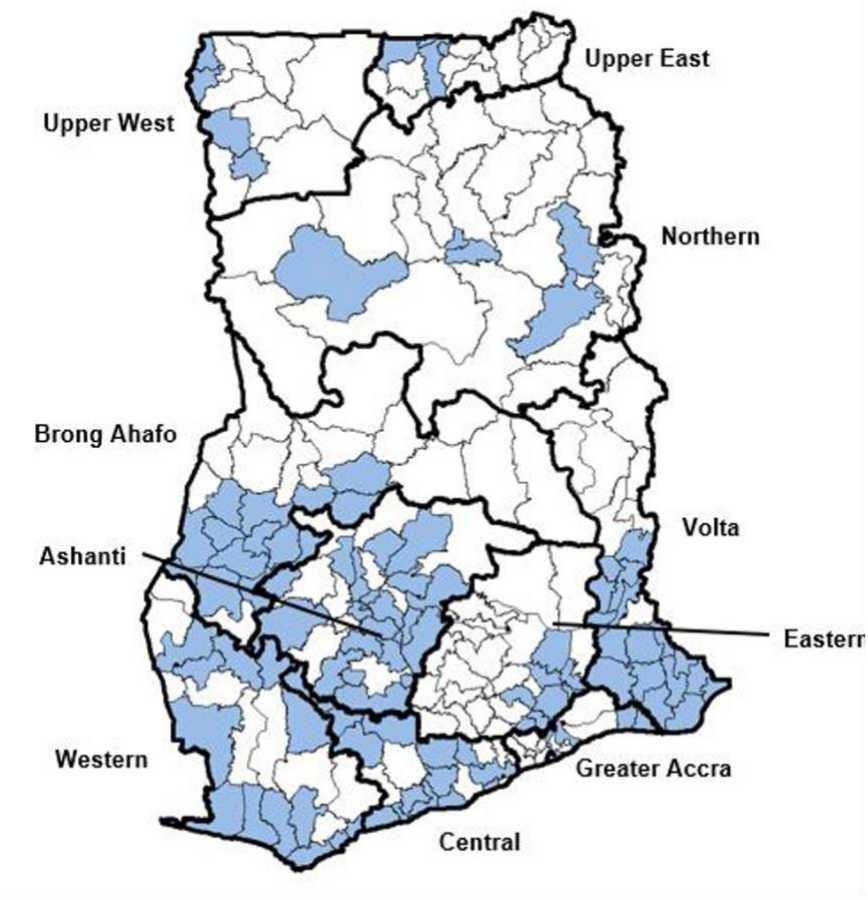
- This program was redesigned in 2016 due to a change in USAID's leadership in Ghana and subsequent ministry and project discussions. While Ghana has a bilingual primary school structure, English was dropped from the Ghana Learning program design in 2016, as were school management committees and math. Implementation for the final model ran from 2016 to 2019.
- The government recognizes 11 languages of instruction, meaning that teachers were sometimes instructing in a language that they were not confident teaching in, and more than 30% of the students were learning in a language that was not their mother tongue.
- The Ministry of Education and Ghana Education Services implemented the majority of the program's activities, with guidance from the Ghana Learning team. This required working within the government systems for training and support visits. Training

[17] See Section 10 of this report.

required intensive capacity building for reading instruction. Training activities for the program were funded primarily through grants to ministry bodies.

- Teaching and learning materials were sourced and distributed from Ghanaian private sector vendors instead of the government.

Implementation Map



Tusome Early Grade Reading Activity

Kenya
RTI International[18]

Development partner: USAID
Years of the activity included in this research: 2015–2021
Supplementary to the government's official public school literacy instruction: No
Primary schools reached: 22,000+ public schools, 5,000 private schools, 1,500 alternative schools
Students reached (grades 1–3): 5.8 million
Teachers reached: 77,000
Government system has bilingual instruction in primary school: Yes (Kiswahili and English)
Government-recognized languages of instruction: 42 (Kiswahili and English)[19]
Teacher training and support: Cascade model, with teacher training managed at the national level; trainings took place during school breaks and between terms and school years. Training manuals were considered essential; curriculum support officers provided coaching to teachers and observational data; the data were uploaded to a data visualization dashboard for the Ministry of Education and senior management to use in decision-making processes.
Impact measured at grade: 1 and 2
Language(s) measured: Kiswahili and English
Planned instructional dosage: 338 hours
Language fluency threshold: 45 cwpm Kiswahili at the end of grade 2 or grade 3
65 cwpm English at the end of grade 2 or grade 3
Endline impact: 8 percentage point increase from baseline in the number of grade 2 students who reached the fluency threshold in Kiswahili
6 cwpm average increase in Kiswahili
8 percentage point increase from baseline in the number of grade 2 students who reached the fluency threshold in English
12 cwpm average increase in English

Unique Aspects of This Program

- Tusome grew out of the Primary Mathematics and Reading Initiative, which ran from 2011 to 2014. Specific components of this initiative were incorporated into Tusome's design.

[18] See Section 10 of this report.
[19] There are 42 national languages that can be used for literacy instruction, depending on geographic location. For the first three years of primary school (grades 1–3), instruction is given in the predominant language spoken in the catchment or mother tongue, and English.

- Tusome was implemented through the existing government system. It was managed through a national technical team composed of education officers across the Ministry of Education directorates and key government bodies working on education. Tusome was structured to maximize the skills and abilities of staff from both the ministry and semi-autonomous government agencies.
- Copyright ownership of all instructional materials, as well as the digital infrastructure (such as the dashboard), were handed over to the government to strengthen government systems.
- The government of Kenya purchased and distributed revised versions of Tusome materials in English and Kiswahili for all grade 1–3 students. Changes in government procurements led to significant cost savings and 1:1 student- and teacher-to-book ratios for other subjects nationwide.

Implementation Map

Source: urbankenyans.com

Sénégal Lecture Pour Tous[20]

Senegal
Chemonics[21]

Development partner: USAID
Years of the activity included in this research: 2016–2021
Supplementary to the government's official public school literacy instruction: No
Primary schools reached: 3,900
Students reached (grades 1–3): 466,100
Teachers reached: 9,300
Government system has bilingual instruction in primary school: Yes (French and second language of instruction)[22]
Government-recognized languages of instruction: No official language of instruction policy, but 6 national languages officially codified for instruction, in addition to French, and others partially codified
Teacher training and support: Model evolved over time as the reform progressed; included a training cascade model strengthening teachers, directors, and inspectors; monthly in-school teacher learning circles run by school directors; quarterly cluster-based teacher learning circles run by senior directors and supported by inspectors; coaching sessions with classroom observations intended to happen twice monthly conducted by a combination of school directors and inspectors (shifted to once a month, complemented by more structured monthly school-based group coaching in teacher learning circles); and push SMS messages to teachers, directors, and inspectors with instructional tips, reminders, and ongoing motivational messages plus virtual communities of practice through WhatsApp groups. The program included a pre-service training component. It also conducted experimental supplemental coaching through nearby experienced school directors and via phone support by senior inspectors at a very small scale though a randomized control trial.
Impact measured at grade: Midline grade 1 and 2; endline grade 1 and 2
Language(s) measured: Pulaar, Seereer, and Wolof
Planned instructional dosage: Midline = 219 hours; endline = 294

---

[20] Lecture Pour Tous directly supported reform efforts led by the by the Ministry of Education to use evidence-based approaches, including national languages, for reading instruction in the early grades to increase foundational literacy. The ministry determined the critical aspects of the program, including first-language literacy instruction dosage (Chemonics, 2022; personal correspondence).
[21] See Section 10 of this report.
[22] At the time of Lecture Pour Tous, the core government system had not yet officially adopted bilingual instruction. There were other pilots before and during Lecture Pour Tus, but it wasn't until the end of Lecture Pour Tous that the government formally validated the reform plan to convert all public primary schools to a bilingual model (which is only just now slowly moving forward). Lecture Pour Tous was not officially a fully bilingual instruction program. Only first language was bilingual, while second language, using a different curriculum and reading approach, continued alongside.

Language fluency threshold: 25 cwpm for each language by end of grade 1 or grade 2

Impact: This program's endline findings were impacted by COVID-19. While some programs included in this study did not conduct an endline assessment once schools reopened following the pandemic, Lecture Pour Tous did conduct an assessment and, in doing so, provided strong anecdotal evidence on the dramatic negative impact of COVID-19 on learning outcomes. In this report, we present Chemonics' Lecture Pour Tous midline and endline findings.

Midline (Two Years of Implementation)

Pulaar

- 21 percentage point increase from baseline in the number of students who reached the fluency threshold
- 13 cwpm average increase in Pulaar

Seereer

- 16 percentage point increase from baseline in the number of students who reached the fluency threshold
- 16 cwpm average increase in Seereer

Wolof

- 36 percentage point increase from baseline in the number of students who reached the fluency threshold
- 18 cwpm average increase in Wolof

Endline (Post-COVID-19)

Pulaar

- 17 percentage point increase from baseline in the number of students who reached the fluency threshold
- cwpm average increase in Pulaar was not provided at endline

Seereer

- 9 percentage point increase from baseline in the number of students who reached the fluency threshold
- cwpm average increase in Seereer was not provided at endline

Wolof

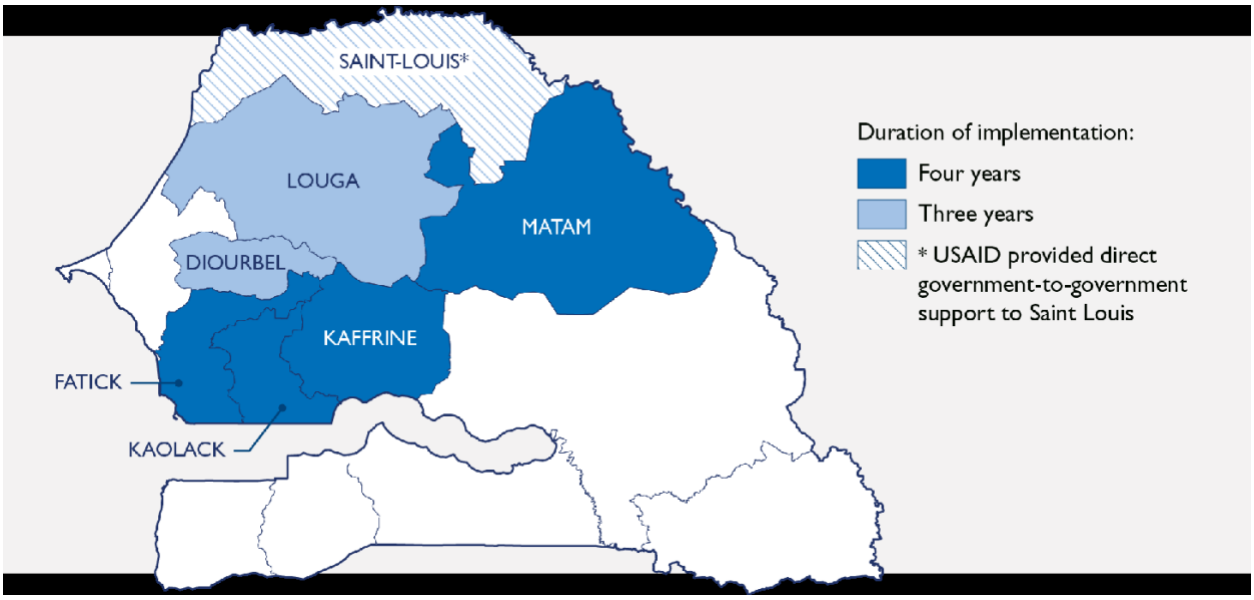
- 20 percentage point increase from baseline in the number of students who reached the fluency threshold
- cwpm average increase in Wolof was not provided at endline

Unique Aspects of This Program

- The program worked at all levels of the Ministry of Education—national, regional, and department. Government actors served as trainers and coaches.
- Incorporated a *faisons ensemble* approach, or doing it together, which was difficult but critical to Lecture Pour Tous's success.

- Chemonics worked with the Ministry of Education to develop reading standards and later to update these standards using the new data generated and in accordance with the Global Proficiency Framework. Monitoring was integrated into the education system's existing standardized quarterly assessments.
- Student-level data were shared widely, including in community dissemination events.
- Understanding the political economy of bilingual reading reform in Senegal was critical at the time of implementation and will remain so in the future.

Implementation Map



Northern Education Initiative Plus (NEI Plus)

Nigeria
Creative Associates[23]

Development partner: USAID
Years of the activity included in this research: 2015–2021
Supplementary to the government's official public school literacy instruction: No
Primary schools reached (grades 1–3): 2,700
Students reached: 568,000
Teachers reached: 7,900
Government system has bilingual instruction in primary school: No (although reading instruction switches from Hausa, which is taught in grade 1, to a transition-to-English program starting in grade 2)
Government-recognized languages of instruction: 4 (Hausa, Igbo, Yoruba, and English)
Teacher training and support: Cascade model; teachers received two to three rounds of training each year; pre-service early grade reading courses were also provided; education officials served as school support officers, who provide external coaching and observations; principal quality assurance officers, who were employees of the State Universal Education Board, supported school support officers; teacher learning circles and cluster learning circles were also held.
Impact measured at grade: 2
Language(s) measured: Hausa
Planned instructional dosage: 360 hours in Bauchi (midline)
443 hours in Sokoto (midline)
Language fluency threshold: 20 cwpm at the end of grade 2
Midline impact: This program did not conduct an endline assessment due to the COVID-19 pandemic. Therefore, we use the midline findings in this report.

Bauchi

- 5 percentage point increase from baseline in the number of students who reached the fluency threshold in Bauchi
- 4 cwpm average increase in Bauchi

Sokoto

- 3 percentage point increase from baseline in the number of students who reached the fluency threshold in Sokoto
- 2 cwpm average increase in Sokoto

[23] See Section 10 of this report.

Unique Aspects of This Program

- NEI Plus built on the past success of NEI, also led by Creative Associates. NEI, a four-and-a-half-year project funded by USAID, strengthened basic education systems in Bauchi and Sokoto by increasing access to quality basic education and social services (Creative Associates, n.d.).
- The program focused on reading outcomes for school-aged children and increased access to basic education for out-of-school children. Learning at Scale focused on formal education results.
- The program concentrated its efforts in Bauchi and Sokoto.
- The program was built on two previous USAID programs: Literacy Enhancement Assistance Project (2001–2004) and the Nigeria Reading and Research Access Activity (2014–2015).
- Government counterparts included the State Universal Basic Education Board, the State Ministry of Education, and local government education authorities. The National Commission for Colleges of Education provided pre-service teacher training, and the Nigerian Education Research and Development Council provided training on curriculum and materials development.
- The program made increased time available for reading for all schools in both states.
- A memorandum of understanding with the government called for joint financing of some activities (e.g., for the state to use the government budget to print textbooks, train teachers outside of NEI Plus schools, and conduct EGRAs and local education monitoring approach assessments).

Implementation Map

[Not available]

Pakistan Reading Project

Pakistan
International Rescue Committee (IRC)[24]

Development partner: USAID
Years of the activity included in this research: 2013–2019
Supplementary to the government's official public school literacy instruction: No
Primary schools reached (grades 1–2): 1,800
Students reached: 123,500
Teachers reached: 27,100
Government system has bilingual instruction in primary school: No
Government-recognized languages of instruction: 6 (Punjabi, Pashto, Sindhi, Saraiki, Urdu, and Balochi; some of these languages have multiple dialects)
Teacher training and support: Cascade model; collaboration with the Higher Education Commission and faculty from teacher training institutions to develop five reading integration courses and five reading specialization courses, as well as orientation sessions for teacher training institute faculty on these courses; support primary focused on school visits from external coaches and mentors (who were government staff, senior teachers, and supervisors) and teacher inquiry groups. Teachers also had access to virtual mentoring videos and audio lessons through program-purchased tablets.
Impact measured at grade: 2 (cohort 3)
Language(s) measured: Urdu and Sindhi
Planned instructional dosage: 210
Language fluency threshold: 60 cwpm Urdu at end of grade 3
Impact: 18 percentage point increase from baseline in the number of students who reached the fluency threshold in Urdu.

13 cwpm average increase in Urdu

Unique Aspects of This Program

- The objective was to support provincial and regional departments of education in Pakistan.
- A needs analysis was conducted to identify key government policies to focus on.
- The program incorporated government-led teacher trainings, teacher inquiry groups, and school support visits.
- The original cost analysis for this program conducted by the IRC estimated the recurring costs for the government of Pakistan to sustain the program.[25]

[24] See Section 10 of this report.
[25] For more on the IRC's cost analysis of the Pakistan Reading Project, see Airbel Impact Lab (2020).

- Pakistan is decentralized at the provincial level. Each province has its own curriculum, textbooks, and teacher training institutes. Despite the Pakistan Reading Project working nearly at scale, this decentralization is the equivalent of managing individualized programs in each province.
- Realized oral reading fluency gains varied by language and province from a low of 3 additional cwpm to 27.

Implementation Map

[Not available]

Read India Program

India
Pratham[26]

Development partner: Multiple
Years of the activity included in this research: 2016–2020[27]
Supplementary to the government's official public school literacy instruction: Yes
Primary schools reached: 22,100
Students reached: 876,400
Teachers reached: 28,000
Government system has bilingual instruction in primary school: No
Government-recognized languages of instruction: 24 (Hindi, English, Assamese, Bengali, Bodo, Dogri, Gujarati, Kannada, Kashmiri, Konkani, Maithili, Malayalam, Marathi, Metei, Nepali, Odia, Punjabi, Sanskrit, Santali, Sindhi, Tamil, Telugu, and Urdu)
Teacher training and support: Top-down train-the-trainer cascade model provided by Pratham to District Resource Group, then to block and cluster resource coordinators, and then to teachers; practice classes were seen as being critical and unique to Pratham's approach. Trainings were "lean" and delivered by one person; cluster resource coordinators, or coaches, supported teachers over the 60-day program cycle. Coaches were required to have firsthand experience implementing Teaching at the Right Level (TaRL).
Impact measured at grade: 4 and 5
Language(s) measured: Hindi
Planned instructional dosage: 45 hours
Language fluency threshold: 45 cwpm at the end of grade 2
Impact: Grade 4 = 33 percentage point increase from baseline in the number of students who reached the fluency threshold for Hindi
Grade 5 = 31 percentage point increase from baseline in the number of students who reached the fluency threshold for Hindi
cwpm was not measured

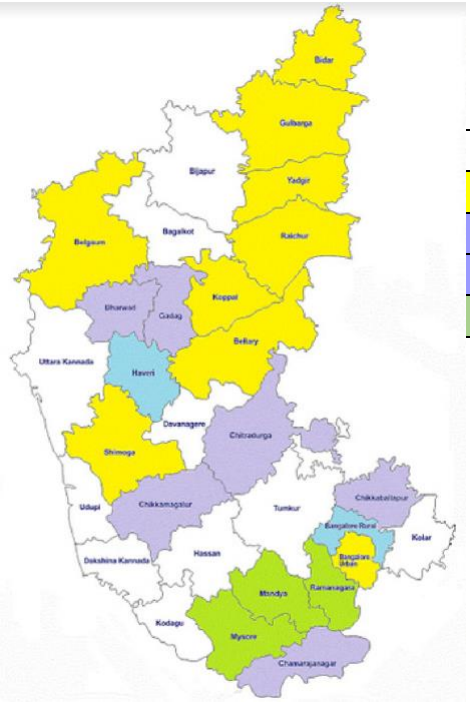
Unique Aspects of This Program

- Read India was a program supplemental to the government's official public school instruction.
- The impact of Read India may not be easily replicable due to its deep history with the Indian government. This program was first implemented in 2016 and evolved and matured over time. The impact discussed in this report is the result of years of work and the program progressively improving the activity.

[26] See Section 10 of this report.
[27] Although the iteration of the program included in this report was from 2016 to 2020, the Read India model continues to be used in various forms.

- Teachers not only had Pratham Read India instruction but also may have had government-led teacher training. There was an acknowledgment by the government that teachers needed to be trained differently and that funding needed to be provided for this.
- Pratham's hallmark pedagogical approach is TaRL.
- The government needed to have a "shift" in thinking, and the materials used in this activity were developed and printed specifically for this program.
- The program was delivered through Pratham-trained instructors and through a partnership with state and district governments and government teachers.
- Midline assessments occurred 30 days after baseline to allow for quick modifications and faster student learning.
- Coaches also conducted administrative work for the government, including monitoring sanitation and the midday meal program, and collecting administrative data.
- The objective of Read India was not to change central government policy. Instead, the program was implemented at the state or substate level with the goal of strengthening practice, which could in turn influence national and state policy.

Implementation Map



Legend	
Districts of 2017-18 (where program ran for 2 year)	
Districts of 2018-19 (where program ran for 1 year)	
Districts of 2019-20	
Pilot districts (2016-17)	

Scaling-Up Early Reading Intervention (SERI)

India
Room to Read[28]

Development partner: Multiple[29]

Years of the activity included in this research: 2015–2020

Supplementary to the government's official public school literacy instruction: Yes

Primary schools reached: 2,000

Students reached: 388,700

Teachers reached: 10,800

Government system has bilingual instruction in primary school: No

Government-recognized languages of instruction: 24 (Hindi, English, Assamese, Bengali, Bodo, Dogri, Gujarati, Kannada, Kashmiri, Konkani, Maithili, Malayalam, Marathi, Metei, Nepali, Odia, Punjabi, Sanskrit, Santali, Sindhi, Tamil, Telugu, and Urdu)

Teacher training and support:

- The instruction point teachers' training was aligned with the government's in-service teachers' training. Of the ten days of in-service teachers' training, four days were given to Room to Read for language-instruction training for partnership schools.
- Four days of teacher training were conducted each academic year in two phases (two days in each phase).
- Two days of training were organized at the block level and two days of refresher training were organized at the cluster level in each academic year.
- In the first academic year, teachers' trainings were imparted by Room to Read staff, and from the second year onward they were imparted jointly by master trainers from the government and by Room to Read's internal staff.
- Sessions on the various components of reading instruction—namely, phonological awareness, phonics, vocabulary, fluency, and comprehension—were held, and each component was followed by a demonstration by literacy facilitators and practice by participants. Two other components (independent reading time and writing) were also discussed.
- In demonstration schools, eight days of teachers' training were imparted each academic year by Room to Read staff in two phases (four days in each phase).

Impact measured at grade: 2

Language(s) measured: Hindi

Planned instructional dosage: 288

Language fluency threshold: 45 cwpm at the end of grade 2

Impact:

Chhattisgarh and Uttarakhand-Partnership

- 20 percentage point increase from baseline in the number of students who reached the fluency threshold in Hindi

[28] See Section 10 of this report.

[29] Room to Read has worked with many development partners over the years. USAID funding should be viewed as a "boost" to the overall work.

- 18 cwpm average increase in Hindi

Madhya Pradesh and Uttar Pradesh-Partnership

- 9 percentage point increase from baseline in the number of students who reached the fluency threshold in Hindi
- 6 cwpm average increase in Hindi

Unique Aspects of This Program

- Room to Read has a rich history of work in India, beginning with the rollout of library programs in 2003. The SERI program had a more than ten-year ramp-up and was deeply integrated into the country's system. The gains realized in SERI are a result of this history and the sole product of the work being discussed in this research.
- One of the objectives was to scale up early grade reading. However, SERI recognized the limitations of having a nongovernmental organization expand its implementation or of handing over the program to the government.
- Gains achieved in the demonstration phase were replicated in the partnership phase despite a decrease in Room to Read's involvement in the program.
- Because SERI historically had access to unrestricted funding, caution should be used when comparing SERI to other programs.

Implementation Map

[Not available]

6. PROGRAM LIFE CYCLES

We, as a sector, need to better understand and be transparent about local challenges so implementers can design their programs to fit within the local context and allocate funds appropriately to cover necessary start-up and implementation activities. Development partners should expect to invest different amounts for start-up costs contingent on a host of factors.

A program life cycle provides a visualization of the different phases of implementation and helps frame the data needed for a cost analysis. For example, for this analysis, we attempted to capture and calculate the costs associated with implementing the impact assessment portion of each program. Implementation time is reflected in the green cells for each program shown in Figure 3. Mapping the programs' life cycles demonstrated that the time period from the month that implementation began to the month that the commitment closed ranged from 47 months to over 90 months. Where applicable, the figure shows which months were impacted by the COVID-19 pandemic.[30] The program with the lowest number of commitment months was Room to Read's SERI–Madhya Pradesh and Uttar Pradesh (47 months), and the program with the highest number of commitment months was IRC's Pakistan Reading Project (90 months).[31]

Looking at the variation in start-up time among the programs was interesting. For this analysis, start-up time was defined as the period after the commitment was signed but before the baseline was conducted. This start-up time includes more than just setting up an office and hiring staff. It can also include the development of teacher training and support models and of teaching and learning materials. Interestingly, the program with the longest start-up time was Room to Read's SERI–Chhattisgarh and Uttarakhand (nine months), and the program with the shortest start-up time was Room to Read's SERI–Madhya Pradesh and Uttar Pradesh (one month). Room to Read had been working on foundational literacy in India for more than ten years prior to the USAID-funded SERI project. It was able to dramatically shorten the start-up time for Madhya Pradesh and Uttar Pradesh because it had

[30] As previously mentioned, several of the programs were impacted by the COVID-19 pandemic. Of the eight programs studied here, COVID-19 impacted six; the other two programs ended before the pandemic. For some organizations, the pandemic disrupted implementation and endline data analysis (n = 3). For others (n = 3), program close-out was affected. For example, Creative's NEI Plus lost three months of implementation time, Chemonics' Lecture Pour Tous lost five months of implementation time, and RTI International's Tusome lost nine months for the data analysis and close-out process because of the pandemic. These time estimates should be considered the minimum amount of time lost. Once schools reopened, in-person training for teachers may have been delayed by additional months.

[31] No start-up time is included for the IRC's Pakistan Reading Project, as it commenced in 2013 and start-up activities were integrated with the other arms of the activity.

already developed the instructional materials, training packages for teachers, academic monitoring, and books for the library in developing Chhattisgarh and Uttarakhand.

Additionally, this research looked at how long each program was implemented in the field, including breaks and holidays. The green bars in Figure 3 denote how long the activity was implemented. As the figure shows, RTI International's Tusome had the longest implementation time (71 months), and Room to Read's SERI–Chhattisgarh and Uttarakhand had the shortest implementation time (19 months). Understanding implementation time may help us understand the differences in context. For example, early on in EQUIP-T's implementation, Cambridge Education experienced high levels of teacher absenteeism, late teacher arrivals, student absenteeism, and inefficient use of instruction time (Rawle et al., 2019a). Therefore, Cambridge Education may have used the early months of the program's implementation to motivate teachers and change their behavior rather than use the time for actual instruction. In fact, from baseline to midline, students in grades 1 and 2 received, on average, only 58 minutes of instruction time per day, but by endline, these same students were receiving an additional 30 minutes of instruction time per day (Rawle et al., 2019a).

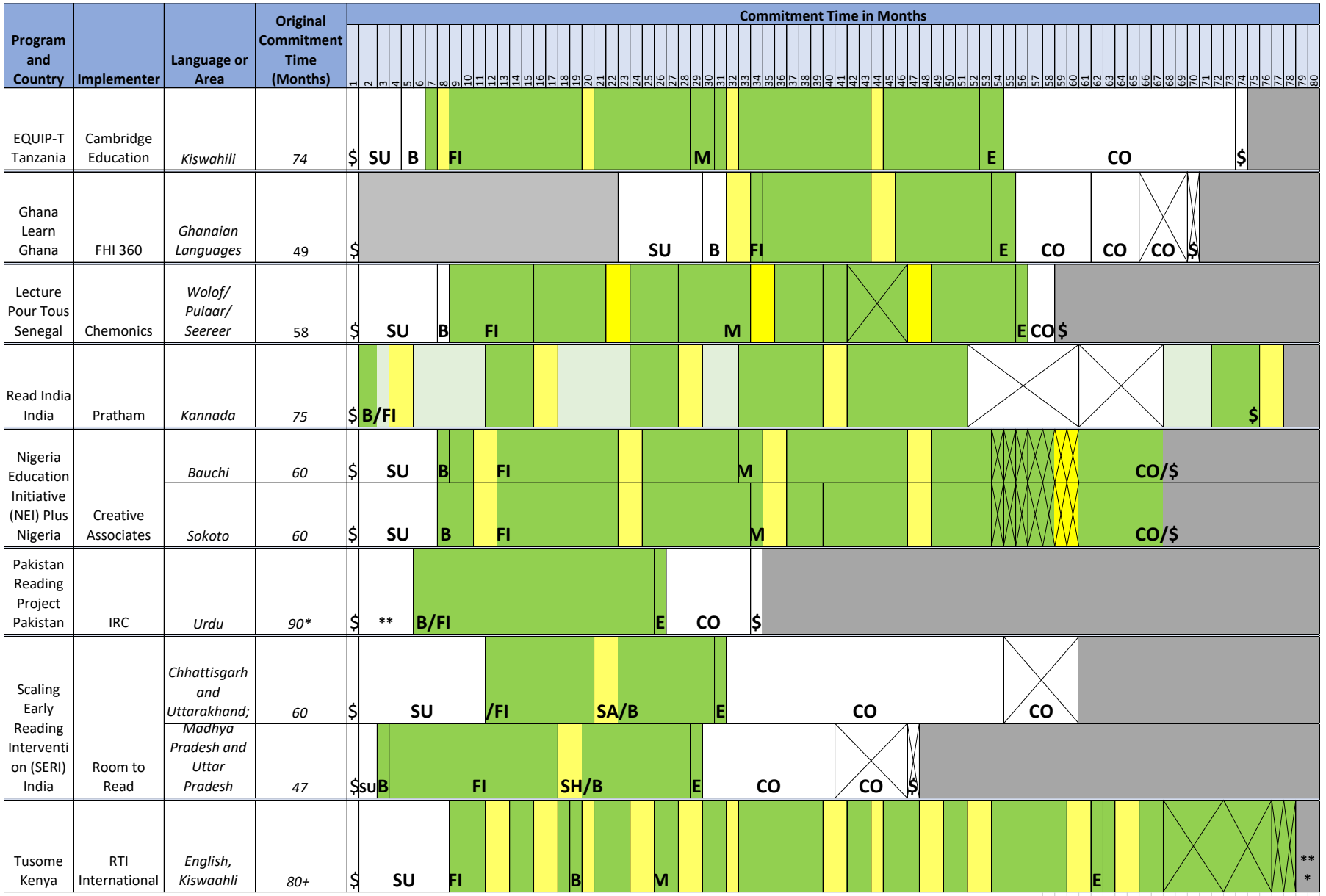
Partner organizations were also asked to provide their best estimate of when their program's implementation actually began. A program was considered to be in implementation when those inputs listed in its theory of change had been realized. This might include when books had been distributed and teacher training had occurred. Three of the programs (Pratham's Read India, IRC's Pakistan Reading Project, and Room to Read's SERI–Chhattisgarh and Uttarakhand) were considered to have commenced implementation during the same month that baseline was taken. This could be due to the previous work conducted by these partners in these countries.

Because unexpected delays are a reality in developing and implementing these types of programs—such as delays in the arrival of books, delays in the timing of teacher trainings, government delays due to policy changes or approvals, or delays in the hiring of key staff—it is not uncommon for a program to not be considered fully underway at baseline. The delay between baseline and full implementation is of interest because it means that valuable instruction time is lost and that costs are still incurred. Finally, it is worth noting that only one program (Room to Read's SERI) implemented literacy activities during school breaks. Given the high learning loss experienced during the COVID-19 pandemic, activities during school breaks may now be desirable in some contexts.

This visual display of programs' life cycles shown in Figure 3 demonstrates that we as a sector need to better understand and be transparent about local challenges so that implementers can design their programs to fit within the local context and allocate funds appropriately to cover necessary start-up and implementation activities. Development partners need to anticipate different start-up costs contingent on a host of factors when considering the duration of a program and how long it may take to achieve the desired

outcomes. These factors include, but are not limited to, political complexities surrounding the number and type of instruction languages, disabilities, and locations.

Figure 3. Program life cycles: Impact assessment portion



*No start up time as estimated for IRC PRP because contract was signed in 2013 and start up was in 2014 and 2015 and integrated with the other arms of the activity.

**The contract was signed and start up took place approximately 60 months earlier (2013)

***Close out is 34 month later

Legend	
$	Contract signed or ended
SU	Start-up
B	Baseline
M	Midline
E	Endline
FI	Fully implemented
SA	Summer activities, if applicable
CO	Close-out
	COVID impact, if applicable
Regular school implementation	
School holiday/break	
Implementation	
Not applicable	

Source: Learning at Scale data collection templates for each program referenced in Section 10 of this report

7. PROGRAM DOSAGE AND IMPACT

Success looks different in different contexts.

To understand the cross-country multilanguage findings, we must first understand how oral reading fluency is measured and differences in this threshold by country and language. The percentage of students who meet the oral reading fluency threshold is measured by the number of cwpm a learner is able to read aloud. EGRA data from 35 language-specific benchmarks in 20 countries show that the majority of benchmarks set are in the range of 40–50 cwpm (RTI International, 2017). However, similarly to other literacy benchmarks, such as comprehension, gains in a student's cwpm are heavily influenced by the context. For example, Table 2 shows the countries included in this study, the language of instruction, and the minimum cwpm reading fluency threshold for the grades included in this study. The number of correct words read aloud in one minute varies from 20 words per minute to 65 words per minute. Kiswahili, which is the language of instruction in Kenya and Tanzania, has different fluency thresholds. Kenya's benchmark for Kiswahili fluency is 45, whereas Tanzania's benchmark for Kiswahili fluency in the same grade is 50.[32]

These differences can be explained by a range of reasons, including some governments setting lower fluency benchmarks when faced with severely low literacy rates. Literacy, in addition to the language of instruction itself, is also influenced by home and environmental circumstances, as well as other factors, such as whether the language is nonalphabetic, the depth of a language's orthography, and whether the instruction is provided in a student's first language. Understanding these fundamental differences can help us appreciate the barriers to making gains, especially in more challenging contexts with more complicated languages.

[32] A 5 cwpm difference does not represent a substantive difference in literacy skills.

Table 2. Reading fluency by country and language (cwpm)

Country	Language	Fluent reader benchmark (minimum cwpm)[33]	Grade
Nigeria	*Hausa*	20	Grade 2
Senegal	*Pulaar*	25	Grades 1 or 2
	Seereer	25	
	Wolof	25	
Nigeria	*Hausa*	30	Grade 3
India	*Kannada*	35	Grades 4 or 5
Ghana	*Ghanaian languages*	35	Grades 2 or 3
India	*Hindi*	45	Grade 2
Kenya	*Kiswahili*	45	Grades 2 or 3
Tanzania	*Kiswahili*	50	Grades 2 or 3
Pakistan	*Urdu*	60	Grades 2 or 3
Kenya	*English*	65	Grades 2 or 3

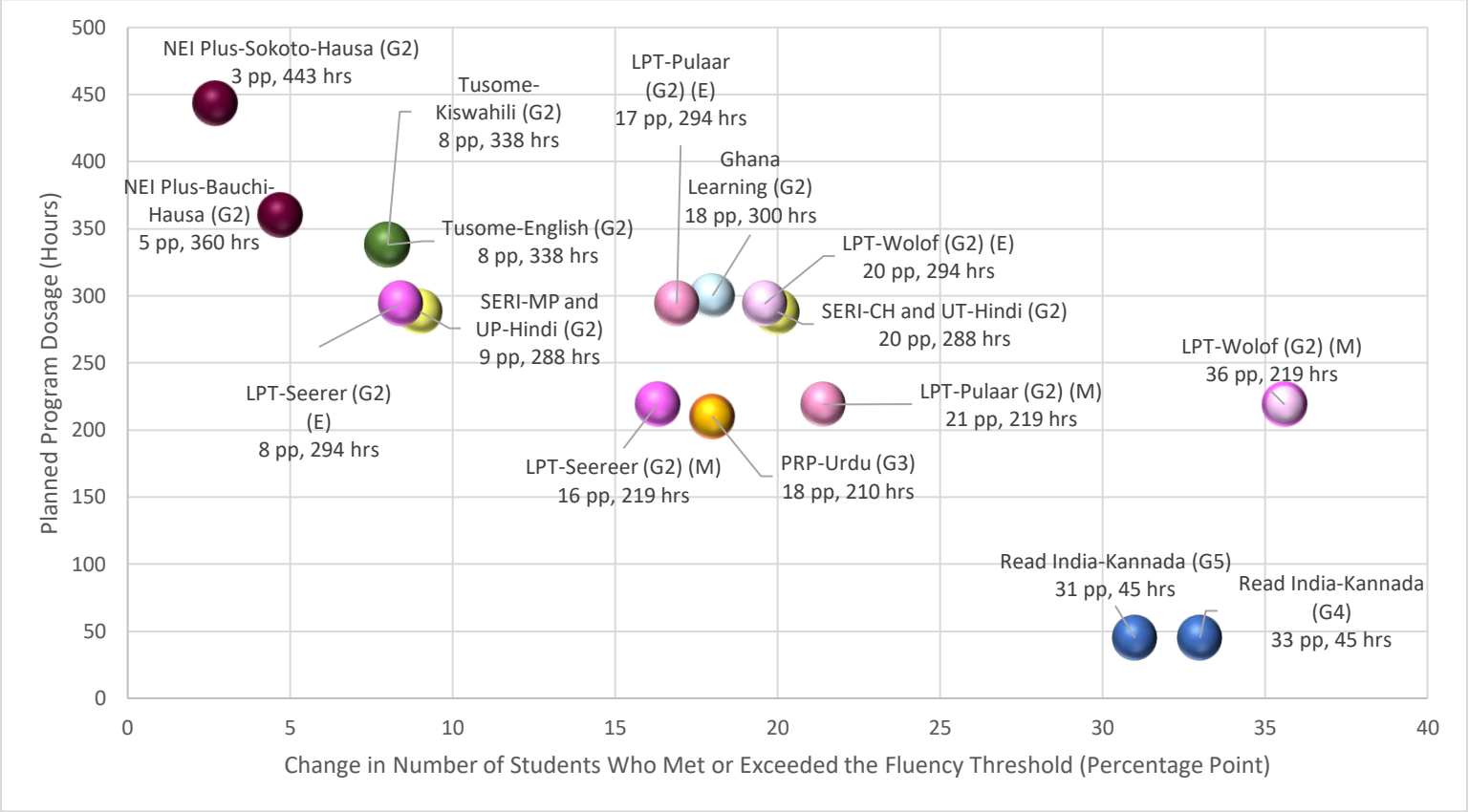
Sources: Partner organizations' endline and midline reports and Learning at Scale data collection templates for each program referenced in Section 10 of this report

The programs in the Learning at Scale research were included due to their success in increasing literacy skills within the primary school populations in their respective countries. Figure 4 shows how success can look very different across and within the different contexts. In this section, we look at each program's planned instructional dosage time and the associated impact. To calculate the dosage time, the instruction time for each literacy lesson was multiplied by the number of lessons provided that contributed to the realized endline impact estimations. This number attempts to capture how many hours of instruction the average student in the program's treatment group received to generate the impact. Impact estimations were measured using the increase in the percentage of students who reached the fluency threshold for the language of instruction. For example, in Room to Read's SERI–Madhya Pradesh and Uttar Pradesh project, 26% of the children in the project schools met or exceeded the oral reading fluency benchmark of 45 cwpm at endline, compared to 17% of the children in the comparison school. This difference of 9 percentage points, or 9 pp, is documented in Figure 4. The same calculation was applied for each program in this research. It should be noted that some of the programs showed larger impacts at midline than endline. This is because the COVID-19 pandemic set some of these programs back and prevented their endline gains from being as high as expected. For transparency, this

[33] Given how frequency government's change literacy thresholds, these literacy benchmarks may not reflect the current standards.

research used each program's endline impact results—and if the program was affected by the pandemic, their midline results are also shown.

Figure 4. Program dosage and impact: Success looks different in different areas



Note: Cambridge Education's EQUIP-T is not included in Figure 4, as verifying the estimated dosage hours was not possible.

Sources: Partner organizations' endline and midline reports and Learning at Scale data collection templates for each program referenced in Section 10 of this report

Figure 4 shows interesting differences among the programs. First, it shows an inverse relationship between dosage and percent point increases in fluency. Given the wide range of contexts and languages included in this research, we are uncertain what conclusions to draw from this, but the trend is worth noting.

It also shows that the programs worked in different grades. Most of the programs included in this research measured the impact in or at the end of grade 2. These include six of the eight programs—Cambridge's EQUIP-T, Chemonics' Lecture Pour Tous, Creative Associates' NEI Plus, FHI 360's Ghana Learning, Room to Read's SERI, and RTI International's Tusome. Please note that Creative Associates' NEI Plus did not conduct an endline study due to the COVID-19 pandemic, and therefore just its midline results are presented here. IRC's Pakistan Reading Project conducted its endline study when students were in grade 3, and

Pratham's Read India conducted its endline study when students were in grades 4 and 5. This mixed-grade approach is consistent with Pratham's TaRL pedagogical approach.[34]

Planned instructional dosage varied greatly across the programs. Creative Associates' NEI Plus–Sokoto had the highest number of instructional hours provided to the treatment group, at 443 hours. Conversely, Pratham's Read India had the lowest number of instructional hours, at 45. This should not be interpreted as Read India being more "time efficient" than the other programs. Instead, Pratham Read India was a supplementary program to the government's official public school literacy instruction.

When reading these instructional hours, please do so with caution. They should not be interpreted as fidelity of implementation findings, which include important aspects such as teacher absenteeism, student absenteeism, and instructional time on task. Such elements impacting actual dosage were not factored into the program dosage estimations due to the lack of detailed information reported. Therefore, the instructional dosage estimations used here should be read as likely higher than actual.

Taking impact estimations and the estimated instructional dosage times together allows us to consider how much time was invested to generate the impact. Figure 4 shows how contextualized program implementations are. For example, Room to Read, after earlier direct implementation, scaled SERI in Chhattisgarh and Uttarakhand and in Madhya Pradesh and Uttar Pradesh using a partnership approach. Both programs measured the impact in grade 2 and provided literacy instruction in Hindi. In Chhattisgarh and Uttarakhand at baseline, 12% of the students in the comparison school could meet the fluency threshold at the end of grade 2, but approximately 32% of the children in the treatment schools met the same threshold, for a difference of 20 percentage points. The difference between the students in the project schools and the comparison schools in terms of cwpm was 18, with 288 hours of dosage. In Madhya Pradesh and Uttar Pradesh, the difference between the treatment and comparison school students who met the fluency threshold was about 9 percentage points and about 6.4 cwpm with the same amount of dosage.[35] When reading these results, consider that students in the comparison schools may have received the same amount of instructional dosage but with inferior instruction. We would need substantive fidelity of implementation assessments for each program's comparison schools to confirm or deny this.

[34] Some of the programs had multiple impact arms. For example, they assessed if their intervention showed an impact after one year and two years of implementation, which is also referred to as instructional dosage or dosage. The partner organization then reported its findings for grade 1 students (who had one year of dosage) and grade 2 students (who had two years of dosage). In this report, we are referring to the impact arm with the highest dosage, which in this example is grade 2 students. The balance of this report will continue to reference each program's impact arm with the highest amount of dosage.

[35] Room to Read reported two possible suggestions for these impact differences. See page 99 of the *Learning at Scale: Interim Report*.

Consider also Chemonics' Lecture Pour Tous, which implemented literacy instruction in Wolof, Pulaar, and Seereer. This program showed remarkable gains. At baseline, only 0.3% of second graders met the fluency threshold in national languages. At midline, after the average student received about 219 hours of instruction, the proportion of students who reached the fluency threshold in Wolof increased by 36 percentage points, students who reached the fluency threshold in Pulaar increased by 21 percentage points, and students who reached fluency in Seereer increased by 16 percentage points. This program was dramatically impacted by the COVID-19 pandemic. However, Chemonics still conducted its endline assessment and, in doing so, documented strong evidence of learning loss most likely caused by school closures and missed teacher training and coaching due to the pandemic. At endline, the proportion of students who reached the fluency threshold in Wolof language instruction dropped by 16 percentage points. Students who reached the fluency threshold for Pulaar dropped by 5 percentage points, and students who reached fluency in Seereer dropped by almost 8 percentage points.

FHI 360's Ghana Learning was implemented in 11 languages, the most of any of the programs included in this research. At baseline, all 11 languages had a negligible level of oral reading fluency. By endline, the comparison group showed a gain of only 2 percentage points, whereas the comparison group reached a gain of 20%, for a difference of 18%. Impressively, Ghana Learning also had a low level of instructional dosage, at about 300 hours, and increases of about 9 cwpm on average across all of the Ghanaian languages included in the program.

Pratham's Read India, which is known for its successful TaRL approach, had the lowest dosage hours, at about 45, and showed gains of 33 percentage points for grade 4 and 31 percentage points for grade 5 in Kannada. No cwpm were reported for Read India because its assessment did not measure reading fluency. IRC's Pakistan Reading Project, which essentially ran six different programs due to Pakistan being decentralized at the provincial level, demonstrated gains of 18 percentage points, with 210 hours of dosage and an incredible additional 13 cwpm in Urdu.

RTI International's Tusome, which worked at scale across all primary schools in Kenya, included literacy instruction in English and Kiswahili. We know that instruction in two languages likely requires greater time to achieve fluency in both languages. The average student had about 338 hours of instruction in each language, and the gains in each were about 8 percentage points. Forty-one percent of grade 2 students reached the fluency threshold in English, and at endline this percentage increased to 49%. When testing in Kiswahili, the percentage of grade 2 students who met the fluency threshold was 20% at baseline but increased to 28% at endline. Given that the fluency threshold for English in Kenya is 65 cwpm, whereas the threshold for Kiswahili is 45 cwpm, we may expect the baseline and endline differences based solely on the different thresholds.

Creative Associates' NEI Plus had the highest dosage in Sokoto, at about 443 hours, and gains of 3 percentage points and about 2 cwpm in Hausa. The program in Bauchi had a higher impact and lower dosage. With approximately 360 hours of instructional dosage, the gains were 5 percentage points and about 4 cwpm in Hausa. This program was also impacted by the COVID-19 pandemic, and an endline analysis was not conducted.

At a minimum, Figure 4 shows the different grades the programs work within, the variations in the programs' instructional dosage, and the impact of each program. Discussing issues related to dosage and impact generates interesting questions. For example, using these programs as guides, what is a reasonable amount of time required by the average nonreading student to obtain a minimum level of fluency in a language in a specific country? In heavily multilingual contexts such as Ghana, should more or less instructional time be expected to achieve fluency? How does bilingual instruction affect the time estimations for literacy gains? In summary, given the rich knowledge possessed by these partners, what information can be shared to help institutional-based knowledge evolve toward sector-level knowledge?

8. COST

Resources cost different amounts in different contexts.

The purpose of this section is to consider the differences in costs across the programs, inclusive of the largest type of costs, and to compare which costs were borne by the development partner versus the government. While we acknowledge that this information is simplified and likely suffers from some level of measurement error, it nonetheless sheds light on the potential strengths and challenges for the sustainability and scalability of the programs.

Program costs were organized into four categories: teacher training, ongoing teacher support, implementation, and grants. Grants were separated into their own category when the costs were too general to be allocated to one of the other three categories. Ongoing teacher support was further differentiated by support provided to teachers from someone external to the school and school-based teacher support. In addition, the implementation category was divided into program activities that took place during the school day and program activities that took place outside of the school day. Partner organizations were then asked to document expenditures for common resources, such as personnel, materials and supplies, equipment and technology, space, vehicles and transportation, and other categories. They were also asked to monetize, to the best of their ability, contributions

made by the government to their program. The objective was to show the joint financial effort required to achieve the successful gains made by each program. All costs were adjusted for a common currency and inflation. Costs are shown in 2021 US dollars. Because of the confidential nature of these data, all cost findings are anonymized. Each program was assigned a number, and the numbering remains consistent throughout the cost findings.

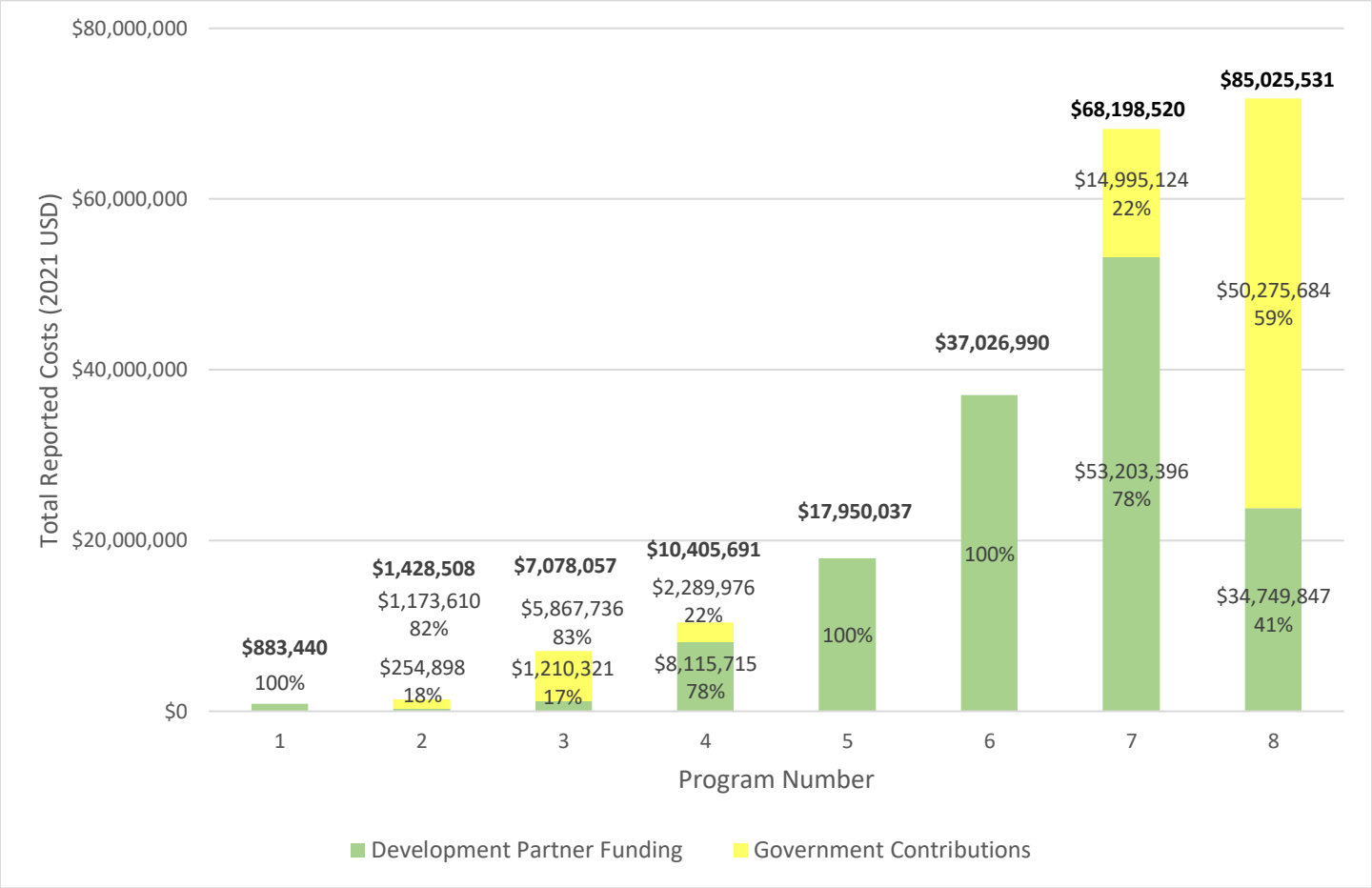
8.1 Program Costs

The programs included in this research were vastly different in terms of funding amounts and subsequent total costs reported. Figure 5 shows expenditures and monetized contributions as reported by the programs. Expenditures are resources purchased using development partner program funds. These figures are often found in an organization's accounting system. (Please refer back to the limitations of this study.) Monetized contributions are program resources provided by the government. These resources could include individuals' time, spaces, and physical inputs such as books. Recall that these findings are narrowly defined by our research. They do not include the costs to develop the programs and overhead (to the extent possible). Therefore, they do not represent the total cost of the programs, which would go beyond the scope of this research.

As shown in Figure 5, the partner organizations reported total costs ranging from $883,440 to $85,025,531 for the specific window of time being evaluated for each program. Five of the eight programs reported monetized government contributions. This does not imply that the three programs that didn't report contributions did not receive any contributions from the government. Nor does this imply that the five programs that did report contributions documented all of the government resources they received. Until there is a systemized method for capturing, monetizing, and assessing the accuracy of contributions in a standardized manner across development partners and their funded programs, gaps will exist.[36] For the programs that reported contributions, the percentage of cost incurred by the development partner versus the government is shown. These programs reported that government contributions ranged from about 20% to 80% of the total costs, as framed by this research. Interestingly, two of the programs that had the lowest reported total costs also had the highest percentage of government contributions. We'll look at this more closely as we consider the largest type of costs for each program.

[36] For more information on efforts to capture contributions in a standardized manner, see EducationLinks (n.d.).

Figure 5. Total costs by program (in 2021 USD)



Source: Learning at Scale data collection templates for each program referenced in Section 10 of this report

8.2 Per-Student Costs

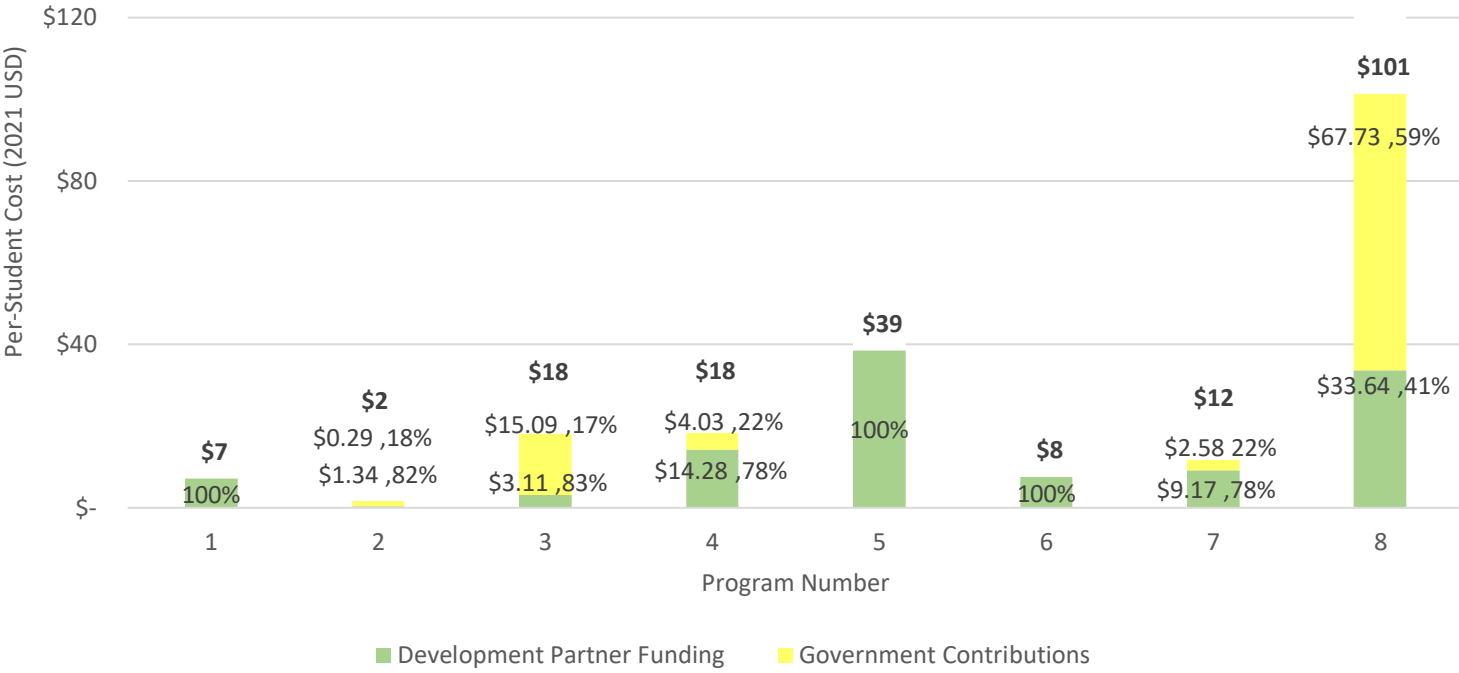
The per-student costs for each program were estimated to be from about $2 to $101. Per-student cost is the total cost divided by the number of students who enjoyed those resources, such as teachers trained or teaching and learning materials produced and distributed. Often, but not always, language influences costs. For example, contexts with multilingual instruction or a high number of languages of instruction will have increased production costs for teaching and learning materials given that they are run in smaller batches and will not benefit from economies of scale. These higher costs, as compared to non-bilingual instruction contexts with a low number of languages of instruction, should be planned and expected. Had development costs been included in this analysis, these costs would also likely be higher in language-rich contexts as well.

Additionally, some of the programs stated that they served "the most disadvantaged populations," whereas others concentrated on standard public school populations. There is

an underlying belief in our field that the most disadvantaged populations are also the populations with the greatest need and the highest costs. These students are sometimes referred to as the "last mile" students. This research cannot dispute or confirm that the most disadvantaged populations are the most expensive to serve. Instead, we urge for clearer distinctions in how the "most difficult to reach and serve" is defined. Reach may be defined as a location, such as remote areas or conflict zones. Serve may mean to instruct, for example, small populations of students with learning differences instructed in a variety of languages in remote areas. Clearly defining these terms can help with the per-student cost estimations by openly discussing cost assumptions and then measuring if those assumptions are true. Evening the scales so that education systems are more equitable is a critical aspect of the work. However, more research is required to determine under what circumstances the assumption that the "most difficult to reach and serve" are the most expensive is true and, if so, why.[37] Discovering that this assumption may not be true for all students grouped into this category would be positive news for governments and partner organizations serving "last mile" students.

[37] IRC's Airbel Impact Lab has also been pushing for greater clarity in defining "last mile students" and the cost to serve them. For more information, see Airbel Impact Lab (n.d.).

Figure 6. Per-student costs (in 2021 USD)



Source: Learning at Scale data collection templates for each program referenced in Section 10 of this report

8.3 Per-Teacher Training and Support Costs

To assess the costs to train and support teachers, this analysis looked at the first two Learning at Scale cost categories: teacher training and ongoing teacher support. Teacher training includes the costs incurred for all personnel and activities related to teacher training, such as master training, training of trainers, and associated materials, in addition to the teacher training per diems, food, and other expenses. These costs represent both orientation and refresher sessions. For ongoing teacher support per teacher, costs include the cost of coach training, support structures, and transportation reimbursement (using a wide variety of methods). Both teacher training and support costs include monetized government contributions if reported by the partner organization.

As shown in Figure 7, the costs for teacher training and ongoing teacher support ranged from about $16 per teacher to nearly $1,687 per teacher. For most programs, the cost to train each teacher was a much larger percentage of the total training and support costs. The per-teacher training cost ranged from about $9 to $1,666. It is unclear which of these programs worked in contexts where the teacher training per diems were set by the

government. Government-set per diems can quickly drive up the cost of teacher trainings and the cost incurred by partner organizations and their development partners.

Most of the programs designed their teacher trainings using a three-level cascade model with an initial training session and consistent refreshers. As discussed in the early sections of this report, teachers viewed these trainings as substantially different from the ones they had participated in previously. Taken together, the implications might suggest that implementers should consider investing less in long, expensive training programs and instead in shorter, quicker trainings focused on particular instructional skills and well-structured ongoing support. At a minimum, programs can assess if a training increased teacher instructional or content capacity and determine if the teacher training was considered "successful" by some measure.

The cost to support each teacher was much lower than the cost of training. These per-teacher support costs were between $3 and $354. For one program, the cost to support teachers was only 1% of the total per-teacher training and support costs. For two programs, the percentage allocated to per-teacher support was over 60%. While each program offered some level of teacher support, the structure of that support varied widely. Some programs provided external coaching and classroom observations, some structured teacher learning circles, and others trained head teachers to provide in-school support. The frequency and duration of these support structures were different across the programs. An interesting takeaway from these findings is to consider whether implementing partners could design rigorous evaluations that could determine the relative contribution to improved outcomes that is attributable to training as compared to coaching.

Five of the eight programs reported government contributions. This does not imply that governments did not provide support to the three programs that did not report contributions. It simply means that these resources were not reported to this research. For the five programs that reported government contributions, the level of support ranged from 16% to 93% of the total costs for teacher training and support. Program 3, for example, invested about $462 on average per teacher for teacher training and coaching. Of this amount, the government provided more than $429 in contributions, and the development partner invested $33. The primary contributions documented by this program include teachers', head teachers', and coaches' time during training, support, and implementation. Some of the other programs documented this type of contribution in addition to other types, such as student learning materials and teacher training materials. We did not ask partners to monetize school space, as we do not have a strong theory of change about how space influences learning outcomes. It is our hope that this important research is forthcoming.

The implications of focusing on teacher training and support relate to the wider challenges embedded in the system in which each program operates. For example, having a top-notch in-service teacher training program is of little value if the teachers attend the trainings

because of the attraction of the per diems (which can be a substantial bonus to a teacher's monthly salary) but do not implement the training content once they return to their schools. Understanding the context and the likelihood of teacher uptake, or how much of the teacher training pedagogical content is actually used by the teacher in classroom instruction, is critical before substantial funds are invested into a multiday training program. We believe that greater dialogue should take place on important but often overlooked topics such as teacher turnover and in-class pedagogical uptake.[38]

Per-teacher training costs can be compared to the amount spent by the government for in-service training. In some countries, budgets are set for in-service teacher training, but the trainings are not implemented. In these cases, the comparator is zero since there are no realized teacher training expenditures. However, in other countries, in-service training programs do occur even if they are inconsistently provided and serve as the comparator.

Additionally, the per-teacher costs can be compared to a teacher's salary. Because the programs are anonymized, this is a purely illustrative example. Let's say that the monthly salary of a primary school teacher in a given country is $500 per month, or $24,000 over the course of four years. The per-teacher training and support costs can be compared to a teacher's salary over the same period of time. For example, say that Program 4 conducted its trainings and support for four years in that hypothetical country. On average, the per-teacher training and support costs for those four years is $1,078. As a benchmark, we can consider if it is reasonable for the program's teacher training and support to cost about 4% of a teacher's salary over those four years.[39] However, this percentage could be lower if there is evidence that teachers used the pedagogical methods beyond the four years of the project and thus the costs could be spread out over perhaps five years. Indeed, 4% may be far too little to move the needle on teacher pedagogical practices and learner outcomes. We suggest that the sector needs more evidence about the reasonable costs, effect on classroom pedagogy, and effect on learner outcomes, customized to context, to increase the capacity of teaching staff.[40]

---

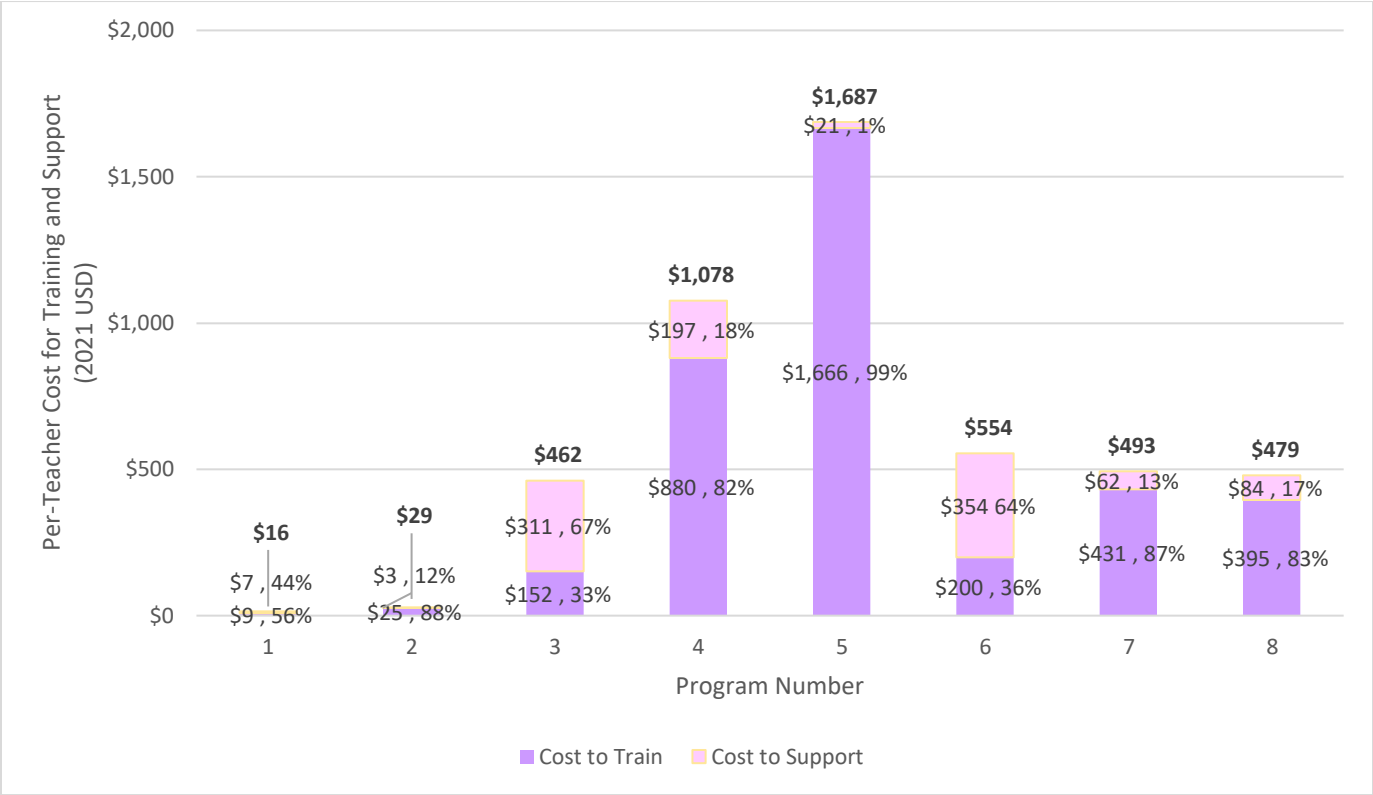
[38] Caitlin Tulloch of the IRC once said that it can be difficult to measure per-teacher training costs because teacher turnover in some contexts is so high. This is a valid observation often overlooked in our work.

[39] A similar benchmark is provided in Walls, Tulloch, & Harris-Van Keuren (2021).

[40] When making these calculations, we should be careful not to generalize about whether teachers are overpaid or underpaid. As we have argued throughout this report, context matters. In some countries, teachers' salaries can be extremely low relative to a country's economic development, whereas in other countries, teachers are fairly compensated, especially when benefits are included in the analysis. Therefore, when using teachers' salaries as a benchmark, the explanation could be that the per-teacher training costs are high or that teachers' salaries are very low. An analysis of teachers' salaries relative to those of comparable workers reveals different findings if one looks at monthly or hourly compensation. Monthly earnings may show that teaches are underpaid, but hourly estimations may show that they are overpaid or that they are at least fairly compensated. An analysis of teachers' salaries is also highly affected by the inclusion of benefits, which are often not enjoyed by workers in comparable industries. See Evans, Yuan, & Filmer (2022).

When analyzing costs, the wider perspective should also be considered. In the absence of these interventions and the associated teacher training and support programs, teachers may continue to deliver pedagogy that produces poor outcomes. The investment in teacher training and support can be seen as a means of levering the substantial investment that governments already make in teacher pay. The investment is worthwhile not only if the cost is low but also if the impact generated is greater than what is currently produced.

Figure 7. Per-teacher training and support costs (in 2021 USD)



Source: Learning at Scale data collection templates for each program referenced in Section 10 of this report

8.4 Largest Type of Cost

Not all government contributions are created equal.

In this section, we discuss the largest types of costs for each program.[41] We begin by looking at the percentage that the largest type of cost constitutes of the total program

[41] We avoid using the term "cost drivers" in this section and throughout the report, as cost drivers can be more of a technical term used to describe those resources that when the price or quantity is

costs, then turn to assess the type of resources, and finally look at whether this resource was provided through development partner funding or government contributions.

Figure 8 shows the largest cost for each program as documented by partner organizations. This figure is color coded, with green cells denoting resources purchased through development partner funding and yellow cells denoting resources contributed by the government. The largest resources for the programs included in this research ranged from 19% to 46% of a program's total reported cost. Because percentages can belie the size of the cost, we also include the cost as well. The reported cost for the largest resources came in at a low of under $250,000 and a high of over $22 million. This again demonstrates the vast difference in the funding size of these programs and the fact that highly effective programs may spend money differently from one another.

Let's consider programs that had similar resources—for example, Programs 2 and 4. Each program noted student materials and supplies as its largest single cost.[42] However, Program 2 documented this resource as accounting for 40% of its total costs, at $573,000, and Program 4 documented this as representing 19% of its total costs, at $2 million. There are a host of reasons why these costs, despite being for the same type of resource, are different. These reasons could include differences in the types of books and supplies ordered; differences in specifications regarding size, paper, and color; differences in the number of books and supplies needed due to different numbers of students; differences in the cost of shipping; and differences in where the books were printed.[43] This is not a surprise. What is interesting is that Program 2's student materials and supplies were provided by the government, whereas Program 4's student materials and supplies were paid for by the development partner. A similar trend is shown with Programs 3 and 6 regarding external teacher support. In Program 6, the costs were borne by the development partner, but Program 3's external support costs were borne by the government. These findings open a door to discuss the circumstances under which the government can or should pay for the cost of certain program inputs (whether in full or in part), especially if the program is working at scale. This information also allows us to explore whether a country has an in-country print mandate, which can impact the cost of teaching and learning materials for a variety of reasons.

Let's now consider programs whose largest costs were in different categories. Program 5 documented that master trainers was its largest line-item cost. An analysis of Program 7's

adjusted can yield large changes in a program's total costs. The data used in this Learning at Scale analysis are not disaggregated to a level that would allow for a meaningful exploration of programmatic cost drivers. Instead, we discuss the largest types of cost. For more information on cost drivers, see Walls, Tulloch, & Harris-Van Keuren (2021).

[42] Student materials and supplies include items such as workbooks and leveled readers. Student equipment and technology are items like tablets.

[43] For more information on the cost differences of student learning materials, see the USAID-funded research by EnCompass (2021).

costs showed that costs related to teacher training and transportation, especially teacher per diems, was its highest cost. As mentioned earlier, high teacher per diems (i.e., those much greater than the cost to cover food, lodging, and incidentals) can be of concern because they may motivate a teacher to attend trainings but provide no lingering incentives to implement once the teacher is back in the classroom. Also, high teacher per diems may be set by the government and may be unsustainable at any level of scale. Program 1's highest cost was related to student equipment.

Program 8 reported that its largest cost was the monetized time of teachers and para teachers. This is one of the key points of this research: teachers' salaries are the highest line-item resource in a government budget, and literacy instruction time is one of the most precious resources that governments possess. Development programs can provide a more collaborative partnership by focusing not just on impacts made but also on the time invested to generate the literacy gains.

Investigating the largest line-item costs of an activity allows us to explore two lines of thought. The first is to consider whether we can decrease these largest costs without sacrificing the quality of the program. Unfortunately, as a sector, we do not have clear evidence on which resources positively contribute to increased student learning outcomes. For example, did investing nearly $16 million in transportation related to teacher training or over $8 million into master trainers positively contribute to increased teacher pedagogy and student outcomes commiserate to the cost? The point is that we don't know, but we should.

The second line of thought focuses on government contributions and the following question: Does the amount and type of government contribution provided to a program offer a signal of the government's commitment and the program's potential ability to scale and sustain? We believe that not all government contributions are created equal and that some might be more important than others. The questions then become these: Should certain government contributions be prioritized, according to the context, in order to increase the chances of the program's sustainability? If so, how will we identify the "sustainability levers" in a context? We believe that these topics warrant wider sector discussions and more research.

Figure 8. Largest type of cost (in 2021 USD)

	Program 1	Program 2	Program 3	Program 4	Program 5	Program 6	Program 7	Program 8
Teacher training: master trainers					$8.3m 37%			
Teacher training: travel and transportation							$15.8m 23%	
Implementation: student equipment	$238k 27%							
Implementation: student materials and supplies		$573k 40%		$2.0m 19%				
Implementation: teachers and para teachers								$22.5m 31%
Ongoing teacher support; external support for teachers			$2.1m 30%			$11.6m 31%		

Green cells = development partner funding; yellow cells = government contributions
Source: Learning at Scale data collection templates for each program referenced in Section 10 of this report

9. SCALE AND SUSTAINABILITY IMPLICATIONS FOR GOVERNMENTS

We need to shine a light on resources provided by the government and program costs driven by the context.

This research highlights several important cost implications for governments.

1. We, as a sector, should flip our focus by first highlighting and more actively discussing what resources and systems are in place in a context and then describing how a development project supports or complements those structures. As programs move to scale and integrate with existing government structures, we need to shine a light on the costs borne by the governments, including (perhaps especially) human capital. Teachers' salaries are the highest line-item resource in a government budget and a government's most valuable resource in improving student literacy rates. Using teacher time, and that of other critical government support staff, as efficiently and effectively as possible to maximize the government's investment is of paramount importance.

2. When conducting a cost analysis, we must be clear on not only *which resources are borne by the government* but *which program costs are driven by the context*. This includes program costs such as government-set teacher training per diems and in-country printing mandates. These cost decisions can be beyond the control of the partner organizations but still drive up program costs.

3. System-level reforms must be driven by evidence about which resources positively contribute to student learning gains in a given context.

4. In the movement toward scale and sustainability, development partners may require a realignment in funding and focus. Programs designed to increase student literacy rates by increasing the capacity of pre-service teacher training require longer time periods than short in-service teacher training schedules. Additionally, development partners may need to focus on system-level indicators such as measuring increased teacher instructional capacity after trainings and demonstrated use in classroom instruction when analyzing the cost and impact of support. These system-level indicators will require that development partner expenditures better align with government contributions to allow for a more targeted response to a specific system-level challenge.

5. Practitioners must be able to communicate impact findings to policy makers in easy-to-understand terms such as "additional students who reached a fluency benchmark."

6. As a sector, we may consider expanding our focus beyond literacy gains and scaling solutions only at very specific grades. The need is wider than simply achieving literacy rates in grade 2 or grade 3. The need—which, as we have shown, is different by context—is to achieve levels of literacy commensurate with what is needed to apply to all future learning. The sector should expand its focus on achieving foundational literacy goals throughout the duration of primary education (not just on a single grade or two).

7. Finally, as stated at the beginning of this report, our objective was not to declare a winner or a loser among the successful programs included in this research. Instead, we tried to demonstrate how the cost of success can be very different depending on the context. As the work on costs progresses, partner organizations will play an increasingly vital role in providing cost and cost-related data, especially since the passing of USAID ADS Chapter 201. ADS 201 mandates that for USAID-funded activities that include an impact evaluation, the assessment must include a cost analysis of the intervention(s) being studied.[44] Making the results public and

[44] See USAID ADS Chapter 201.3.6.4.

discussing the findings[45] will foster deeper conversations about the costs borne and driven by different stakeholders and will allow for more nuanced spending and decision-making. But in addition to costs, we ought to also focus on time and program durations in different contexts to achieve desired outcomes. These conversations about time and money will contribute to governments' ability to sustain and scale cost-effective programs, development partners' ability to spend funds more strategically, and partner organizations' ability to design more targeted responses with the objective of increasing student learning outcomes.

[45] See the IRC's publicly available cost data at https://airbel.rescue.org/.

10. PROGRAM DATA, REPORTS, AND CORRESPONDENCE[46]

Education Quality Improvement Program (EQUIP-T)—Tanzania

Cambridge Education. (June 2020; December 2022). *Learning at scale cost capture template.* Learning at Scale.

Cambridge Education. (October 2020). *Data quality checklist.* Learning at Scale.

Mbiti, I., & Schipper, Y. (2020). *Teacher and parental perceptions of performance pay in education: Evidence from Tanzania.* Research on Improving Systems of Education.

Rawle, G., Binci, M., Pettersson Gelander, G., Harb, J., Jasper, J., Khan, S., . . . Ruddle, N. (2019a). *EQUIP-Tanzania impact evaluation: Endline quantitative technical report, volume I.* USAID.

Rawle, G., Binci, M., Pettersson Gelander, G., Harb, J., Jasper, P., Khan, S., . . . Ruddle, N. (2019b). *EQUIP-Tanzania impact evaluation: Endline quantitative technical report, volume II.* Oxford Policy Management.

Ruddle, N. & Elte, G. (2020). *EQUIP-Tanzania impact evaluation: Endline cost study*. Oxford Policy Management.

SALT Analytics (2019–2023) correspondence with Cambridge Education.

Lecture Pour Tous—Senegal

Chemonics. (July 2020; December 2022; May 2023). *Learning at Scale cost capture template.* Learning at Scale.

Chemonics. (October 2020). *Data quality checklist.* Learning at Scale.

Chemonics. (2022). *Lecture Pour Tous: A Revolution to get all children reading in Senegal; final performance report*. USAID.

Mount-Cors, M., Rousseau, M., & de Galbert, P. (2019). *Lecture Pour Tous/All Children Reading: Senegal early grade reading assessment (EGRA) midline report: Second draft.* USAID.

SALT Analytics (2019–2023) correspondence with Chemonics.

Northern Education Initiative Plus (NEI Plus)—Nigeria

Creative Associates. (n.d.). *Nigeria: Northern Education Initiative Plus*. creativeassociatesinternational.com/africa/nigeria-northern-education-initiative-plus/

Creative Associates. (June 2020; December 2022). *Learning at scale cost capture template.* Learning at Scale.

Creative Associates. (November 2020). *Data quality checklist.* Learning at Scale.

Creative International. (2018). *Northern Education Initiative Plus early grade reading assessment midline report.* USAID.

Creative International. (2021). *Northern Education Initiative Plus: Early grade reading assessment endline evaluation report.* USAID.

SALT Analytics (2019–2023) correspondence with Creative Associates.

Pakistan Reading Project—Pakistan

Airbel Impact Lab. (2020). *Pakistan Reading Project: Cost-effectiveness of a literacy intervention for early grades.* https://airbel.rescue.org/studies/pakistan-reading-project-cost-effectiveness-of-a-literacy-intervention-for-early-grades/

IRC. (2017). *Early grade reading assessment: Urdu endline study.* USAID.

IRC. (2017). *Early grade reading assessment: Khyber Pakhtunkhwa.* USAID.

IRC. (n.d.). *Cost analysis approach: Cohort 3 cost effective analysis (PRP).*

[46] All program cost capture data and data-quality checklists are unavailable to the public due to confidentiality.

IRC. (July 2017). *Cost of the PRP model.*
IRC. (June 2020). *C3 analysis results for RTI.*
IRC. (March 2020). *Cohort 1&2 cost-effective analysis: Analysis workbook for RTI.*
IRC. (March 2020). *Export-PRP materials unit costs.*
IRC. (November 2020; December 2022). *Learning at Scale cost capture template.* Learning at Scale.
IRC. (November 2020). *Data quality checklist.* Learning at Scale.
Keaveney, E., Fierros, C., Rigaux, A., Pepitone, P., Mittelberg, T. & Menende, A. (2021). *Pakistan Reading Project (PRP) 2020 early grade reading endline: Khyber Pakhtunkhwa*. USAID.
SALT Analytics (2019–2023) correspondence with the IRC.

Partnership for Education: Learning (Ghana Learning)—Ghana

Darko Osei, R., Adobea Owusu, G., Asem, F., & Afutu-Kotey, R. (2009). *Effects of capitation grant on education outcomes in Ghana.* Global Development Network 1999–2009.
FHI 360. (September 2020; December 2022). *Learning at Scale cost capture template.* Learning at Scale.
FHI 360. (October 2020). *Data quality checklist.* Learning at Scale.
Government of Ghana. (2007). *Teaching syllabus for Ghanaian languages and culture*. Ministry of Education Science and Sport.
SALT Analytics (2019–2023) correspondence with FHI 360.
Social Impact. (2019). *Impact evaluation: Ghana early grade reading program.* USAID.

Read India—India

ASER. (2012—2018). *Annual status of education report 2012–2018.* ASER.
Government of India. (2021). *National Initiative for Proficiency in Reading with Understanding and Numeracy.* Department of School Education and Literacy, Ministry of Education.
Pratham. (2020). *Odu Karnataka annual report, 2019–2020.* Pratham.
Pratham. (September 2020; December 2022). *Learning at Scale cost capture template.* Learning at Scale.
Pratham. (November 2020). *Data quality checklist.* Learning at Scale.
SALT Analytics (2019–2023) correspondence with Pratham.

Scaling-Up Early Reading Intervention (SERI)—India

Joddar, P. (2018). *Impact evaluation of the literacy program-partnership approach under Scaling Up Early Reading Intervention (SERI) funded by USAID.* USAID.
Joddar, P. (2019). *Impact evaluation of Room to Read's literacy program under partnership approach in Madhya Pradesh and Uttar Pradesh, India.* USAID.
Room to Read. (May 2020; December 2022). *Learning at Scale cost capture template.* Learning at Scale.
Room to Read. (October 2020). *Data quality checklist.* Learning at Scale.
SALT Analytics (2019–2023) correspondence with Room to Read.

Tusome—Kenya

Freudenberger, E., & Davis, J. (2017). *Tusome external evaluation: Midline report.* USAID.
Keaveney, E., Fierros, C., Rigaux, A., Menendez, A., with Dayaratna, V., & Munene, C. (2020). *Tusome external evaluation: Endline report*. USAID.
RTI International. (August 2020; December 2022). *Learning at Scale cost capture template.* Learning at Scale.
RTI International. (November 2020). *Data quality checklist.* Learning at Scale.
SALT Analytics (2019–2023) correspondence with RTI International.

11. REFERENCES

Abdul Latif Jameel Poverty Action Lab. (n.d.). https://www.povertyactionlab.org/

Abdul Latif Jameel Poverty Action Lab. (2016). *J-PAL costing guidelines.*

Abdul Latif Jameel Poverty Action Lab. (2019, September 23). *Innovation in Government Initiative (IGI)*. https://www.povertyactionlab.org/innovation-government-initiative

Airbel Impact Lab. (n.d.). *Audio-class system.* https://airbel.rescue.org/projects/audio-class-system/

Asante Sana for Education. (n.d.). *Tanzania education system*. https://www.asantesanaforeducation.com/tanzania-education-system-

Beggs, C., & Ogando Portela, M.J. (2019). *A Study on Kinyarwanda instructional time in lower primary*. Soma Umenye/USAID Rwanda.

Business Insider. (2020, July 30). *Understanding the new 5+3+3+4 education model in India and what it changes in the education system*. https://www.businessinsider.in/education/news/what-is-new-education-system-5334-in-india-and-how-it-is-different-from-old-system-102/articleshow/77261056.cms

EducationLinks. (n.d.). *USAID Education cost measurement tools*. https://www.edu-links.org/resources/usaid-cost-measurement

EnCompass. (2021). *Report: Cost of teaching and learning materials.* USAID. https://www.edu-links.org/resources/report-cost-teaching-and-learning-materials

Evans, D. K., Yuan F., & Filmer, D. (2022). Teacher pay in Africa: Evidence from 15 countries. *World Development*, *155*. doi: 10.1016/j.worlddev.2022.105893.

Kenyayote. (2022, May 16). *Breakdown of Kenya's new 2-6-6-3 education curriculum framework: Subject taught.* https://kenyayote.com/breakdown-kenyas-new-2-6-6-3-education-curriculum-framework-subject-taught/

Levin, H. M. (1975). Cost-effectiveness analysis in evaluation research. In M. Guttentag & E. Struening (Eds.), *Handbook of evaluation research* (Vol. 2). Sage.

Levin, H., McEwan, P., Belfield, C., Bowden, A., & Shand, R. (2018). *Economic evaluation in education cost-effectiveness and benefit-cost analysis.* Sage.

McEwan, P. (2012, June). Cost-effectiveness analysis of education and health interventions in developing countries. *Journal of Development Effectiveness, 4*(2), 189–213. http://academics.wellesley.edu/Economics/mcewan/PDF/cea.pdf

McEwan, P. J. (2015, September). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research, 85*(3), 353–394. http://academics.wellesley.edu/Economics/mcewan/PDF/meta.pdf

UNESCO. (2016). *Education 2030: Incheon declaration and framework for action for the implementation of sustainable development goal 4*. https://unesdoc.unesco.org/ark:/48223/pf0000245656

UNESCO Institute for Statistics. (2021). 2021 data. http://uis.unesco.org/

Walls, E. (2021). *Cost reporting guidance for USAID-funded education activities.* USAID.

Walls, E., Tulloch, C., & Harris-Van Keuren, C. (2020). *Cost analysis guidance for USAID-funded education activities* (2nd ed.). USAID. https://www.edu-links.org/sites/default/files/media/file/USAID%20Cost%20Analysis%20Guidance_Final%20Feb20_0.pdf

World Bank. (2015). *Governance and finance analysis of the basic education sector in Nigeria.* https://openknowledge.worldbank.org/server/api/core/bitstreams/c1788ffb-90e4-567b-93fa-59c6b2aa1c1f/content

World Bank. (2018). *The world development report 2018: Learning to realize education's promise*. https://www.worldbank.org/en/publication/wdr2018

World Bank. (2021). *Tanzania education sector institutional and governance assessment*. https://openknowledge.worldbank.org/server/api/core/bitstreams/8695777b-8971-5f22-bde9-c75041d9bc9b/content
World Bank Education Statistics. (n.d.). https://datatopics.worldbank.org/education/

APPENDIX A. LEARNING AT SCALE COST ANALYSIS METHODOLOGY

Learning at Scale is a research effort designed to estimate the implementation costs of eight literacy programs conducted in low- and middle-income countries (see Table 3). Each program was assessed for its total cost, its average per-teacher cost for training and support, and considerations for the government.

Table 3. Programs included in the Learning at Scale cost analysis

Program	Country	Implementer
Education Quality Improvement Program in Tanzania (EQUIP-T)	Tanzania	Cambridge Education
Ghana Learning	Ghana	FHI 360
Lecture Pour Tous	Senegal	Chemonics
Northern Education Initiative Plus (NEI Plus)	Nigeria	Creative Associates
Pakistan Reading Project	Pakistan	IRC
Read India	India	Pratham
Scaling-Up Early Reading Intervention (SERI)	India	Room to Read
Tusome	Kenya	RTI International

METHODOLOGY

The Learning at Scale team followed an adapted version of USAID's *Cost Reporting Guidance* (Walls, 2018) and *Cost Analysis Guidance* (Walls, Tulloch, & Harris-Van Keuren, 2021). The guidance provides a systematic framework for USAID evaluation partners, implementing partners, and USAID missions to conduct cost-analysis studies. This guidance was selected as the Learning at Scale cost analysis methodology because of its rigor, specificity for large international education projects, and step-by-step transparency. Additionally, five of the eight Learning at Scale programs were funded by USAID. As a result, there was a higher likelihood that the programs would be familiar with their funder's cost methodology. This familiarity would potentially minimize the program staff's learning curve and time burden, while increasing the accuracy of the submitted data.

Alignment with USAID Cost Categories

The Learning at Scale team reduced the number of cost categories included in the analysis to minimize the time burden on the participating organizations and to potentially increase the accuracy of the data submitted. The four cost categories included in the Learning at Scale data collection tool correspond to USAID's cost categories (see Table 4). The categories of "teacher training" and "ongoing teacher support" correspond to USAID's "higher education/pre-service teacher training" and "in-service teacher training." "Implementation" aligns with USAD's "teaching and learning materials." "Grants and scholarships" aligns with UASID's "grants, scholarships, and cash transfers to individuals/families" and "grants to organizations." All of the other USAID cost categories

were omitted due to their irrelevance to the Learning at Scale research questions and program structures.

Table 4. Learning at Scale alignment with USAID cost categories (2018)

Learning at Scale cost categories	Learning at Scale sub-cost categories (if applicable)	Definition	Alignment with USAID cost categories
Teacher training	—	This category includes the costs for all resources used to train the teachers who implemented the program.	Category 3. Higher education/pre-service teacher training (implementation only) Category 4. In-service teacher training (implementation only)
Ongoing teacher support	External support during the academic year	This category includes the costs for all resources used for external teacher support (e.g., coaching) and for extra assessments that directly influence teacher instruction.	
	School-based support during the academic year	This category also includes the cost for all resources used for school-based support teacher support (e.g., teacher learning circles, communities of practice, etc.) during the school year.	
Implement-ation	During the school day	This category includes the costs for all of the resources used to implement the program during school hours.	Category 5. Teaching and learning materials
	Outside of the school day	This category includes the costs for all of the resources used to implement the program outside of school hours (e.g., after school reading programs and summer camps).	
Grants and scholarships	—	This category includes the fully loaded cost of the grants and scholarships manager, and grants and scholarships amounts.	Category 10. Grants, scholarships, and cash transfers to individuals/families Category 11. Grants to organizations
			Unused USAID cost categories
			Category 1. General operations, management, and reporting
			Category 2. Assessments and evaluations
			Category 6. Systems strengthening
			Category 7. Private sector engagement
			Category 8. Parents & community engagement
			Category 9. Safe schools and infrastructure
			Category 12. Other

APPENDIX B. COUNTRY EDUCATION SYSTEMS

Table 5 shows the current education structure of government-run primary schools. Given the evidence of the benefits of early childhood education, many countries are moving to formally incorporate preschool education into their country's education systems. For example, in 2022 the Ministry of Education in Kenya announced that it was adjusting its 8+4+4 structure (i.e., eight years of primary school, four years of lower secondary school, and four years of upper secondary school) to a 2+6+6 system. This structure includes two years of pre-primary education, six years of primary education (i.e., three years of lower primary and three years of upper primary), and six years of secondary education (i.e., three years of lower secondary and three years of upper secondary). In 2020, India made a similar move and adjusted its education system to five years of foundation education (i.e., three years of pre-primary education and grades 1 and 2), three years of primary school (i.e., grades 3, 4, and 5), three years of middle school education (i.e., grades 6, 7, and 8), and four years of upper secondary education (grades 9, 10, 11, and 12).

Table 5. Country education systems

Country	Pre-primary	Primary	Lower secondary	Upper secondary	Total years
Ghana	2	6	3	4	15
Kenya[1]	2	6	3	3	14
Nigeria	1	6	3	3	13
Pakistan	2	5	3	4	14
Senegal	3	6	4	3	16
Tanzania[2]	2	7	4	2	15
Country	**Foundation**	**Preparatory**	**Middle**	**Secondary**	**Total years**
India[3]	5	3	3	4	15

Sources: UNESCO Institute for Statistics (2021); Kenyayote (2022); Asante Sana for Education (n.d.); Business Insider (2020)

APPENDIX C. GOVERNMENT-SET PRIMARY SCHOOL LITERACY DOSAGE BY COUNTRY

Country	Instruction language	Years of primary school education[1]	Weeks of instruction per year[2]	Lessons per week[3]	Hours of instruction per lesson[4]	Total hours of literacy instruction
Ghana	Ghanaian languages	6	30	6	0.50	540
India: Haryana	Hindi	5	44	5	0.75	825
India: Karnataka	Kannada	5	44	5	0.75	825
India: Uttarakhand, Uttar Pradesh, Madhya Pradesh, and Chhattisgarh	Hindi	5	44	5	0.75	825
Kenya	Kiswahili	6	36	5	0.58	580
	English	6	36	4	0.58	504
Nigeria: Bauchi and Sokoto	Hausa	6	37	4	1.00	888
Pakistan	Urdu	5	28	5	1.00	700
Senegal	Wolof, Pulaar, and Seereer	6	36	4	1.00	864
Tanzania	Kiswahili	7	37	5	1.00	1,295

Note: For India's calculations, we assumed two years of foundational learning and three years of preparatory education as "primary school."
Source: Learning at Scale data collection templates for each program referenced in Section 10 of this report