ALL CHILDREN READING:
GRAND CHALLENGE FOR DEVELOPMENT

EVALUATION REPORT

Nos Enfants Apprennent à Lire
(Our Children Learn to Read)

Implemented by Réseau d'Acteurs Pour le Renouveau de l'Education in Mali

SEPTEMBER 2017

Prepared by
School-to-School International (STS)
For All Children Reading: A Grand Challenge for Development

USAID FROM THE AMERICAN PEOPLE
World Vision
Australian Aid
SCHOOL-TO-SCHOOL INTERNATIONAL

EVALUATION REPORT

Nos Enfants Apprennent à Lire (Our Children Learn to Read)

Implemented by Réseau d'Acteurs Pour le Renouveau de l'Education in Mali

Table of Contents

List of Acronyms

ACR GCD	All Children Reading: A Grand Challenge for Development
App	Application
CAP	Centre d'Animation Pédagogique
CEPROCIDE	Centre de Promotion de la Citoyenneté pour un Développement durable à la base
CI	Confidence Interval
CLSPM	Correct Letter Sounds per Minute
CNWPM	Correct Nonwords per Minute
CWPM	Correct Words per Minute
EDC	Education Development Center, Inc.
EGRA	Early Grade Reading Assessment
EOP	End-of-Project
FOI	Fidelity of Implementation
ICT	Information and Communications Technologies
IRB	Institutional Review Board
IRI	Interactive Radio Instruction
GoM	Government of Mali
M&E	Monitoring and Evaluation
MEN	Ministère de l'éducation nationale
NICRA	Negotiated Indirect Cost Recovery Agreement
NGO	Nongovernmental organization
OCLR	Our Children Learn to Read
OMAES	Œuvre Malienne d'Aide à l'Enfance
ORF	Oral Reading Fluency
PAJE-Nièta	Projet d'Appui aux Jeunes Entrepreneurs-Nièta
PHARE	Programme Harmonisé d'Appui au Renforcement de l'Education
RARE	Réseau d'Acteurs Pour le Renouveau de l'Education
RTI	RTI International
SD	Standard Deviation
SIRA	Selected Integrated Reading Activity
STS	School-to-School International
UNESCO	United Nations Educational, Scientific, and Cultural Organization
USAID	United States Agency for International Development

I. Executive Summary

All Children Reading: A Grand Challenge for Development (ACR GCD)—a partnership between the United States Agency for International Development (USAID), World Vision, and the Australian Government—is an ongoing series of grant and prize competitions that leverage science and technology to source, test, and disseminate scalable solutions to improve literacy skills of early grade learners in developing countries. Round 2 of ACR GCD, which started in 2014 and continues through 2017, supports technology-based innovations to improve early grade reading outcomes in developing countries.[1] These technology-based innovations concentrate on three focus areas:

1. Mother tongue instruction and reading materials
2. Family and community engagement
3. Children with disabilities

ACR GCD Round 2 increased its focus on the assessment of early grade reading skills to understand the ability of technology-based innovations to improve the literacy skills of early grade learners. To measure this, ACR GCD uses the Early Grade Reading Assessment (EGRA) to systematically assess reading skills across all Round 2 grantees. The EGRA is an oral assessment that measures students' most basic foundational literacy skills in the early grades—specifically, recognizing letters of the alphabet, reading simple words, understanding sentences and paragraphs, and listening with comprehension. The EGRA methodology was developed under EdData II and has been applied in more than 30 countries and 60 languages.[2] The EGRA instruments used by ACR GCD grantees were adapted to reflect the specific context of each grantee's project, including adaptations for students who have low vision or are blind and students who are deaf or hard of hearing.

Réseau d'Acteurs Pour le Renouveau de l'Education (RARE)—an ACR GCD Round 2 grantee—is a Malian nonprofit organization dedicated to improving education in the country. RARE implemented the Our Children Learn to Read (OCLR) project that began in February 2015 and concluded implementation in April 2017.[3] The project aimed to improve reading and writing instruction in Grades 1 and 2 by training teachers on the balanced literacy approach[4] to strengthen their competencies and skills in teaching literacy. By improving reading and writing instruction, the OCLR project sought to improve the early grade reading skills—specifically, pre-reading and foundational skills—of students taught by the OCLR-trained teachers. The project was funded under the mother tongue instruction and reading materials focus area; it promotes instruction in Bamanankan, a mother tongue language in the Sikasso region of Mali.

RARE promoted two key technological innovations through the OCLR project for use by the teachers. First, half of the teachers in the intervention group—teachers in intervention A— received tablets preloaded with the Stepping Stone application (app), a mobile lesson creation tool and mobile delivery platform developed by Education Development Center, Inc. (EDC), which was also an implementing partner on the project. Using the Stepping Stone app, RARE provided supporting materials, including videos of teachers demonstrating effective reading and writing instructional techniques. Teachers were instructed to use the tablet and Stepping Stone app to support their implementation of the balanced literacy approach. The other half of the teachers—those in intervention B—received the same in-person training but did not receive tablets. As a second component of the OCLR project, all teachers in both intervention groups had access to interactive radio instruction (IRI), which featured 30-minute

1 All Children Reading. (2017, June). About us. Retrieved from http://allchildrenreading.org/about-us/

2 EdData II was a contract mechanism funded by USAID from January 1, 2004, to December 31, 2013. Implemented by RTI International, the purpose of EdData II is to improve the accuracy, timeliness, accessibility, and use of data for education policy and program planning. See http://www.rti.org/sites/default/files/brochures/eddataii.pdf for additional details.

3 RARE's grant for the OCLR project ended on June 30, 2017.

4 In the balanced literacy approach, teachers weave together activities targeting language mechanics, connected reading, and knowledge implementation via writing. See Project Description for additional details.

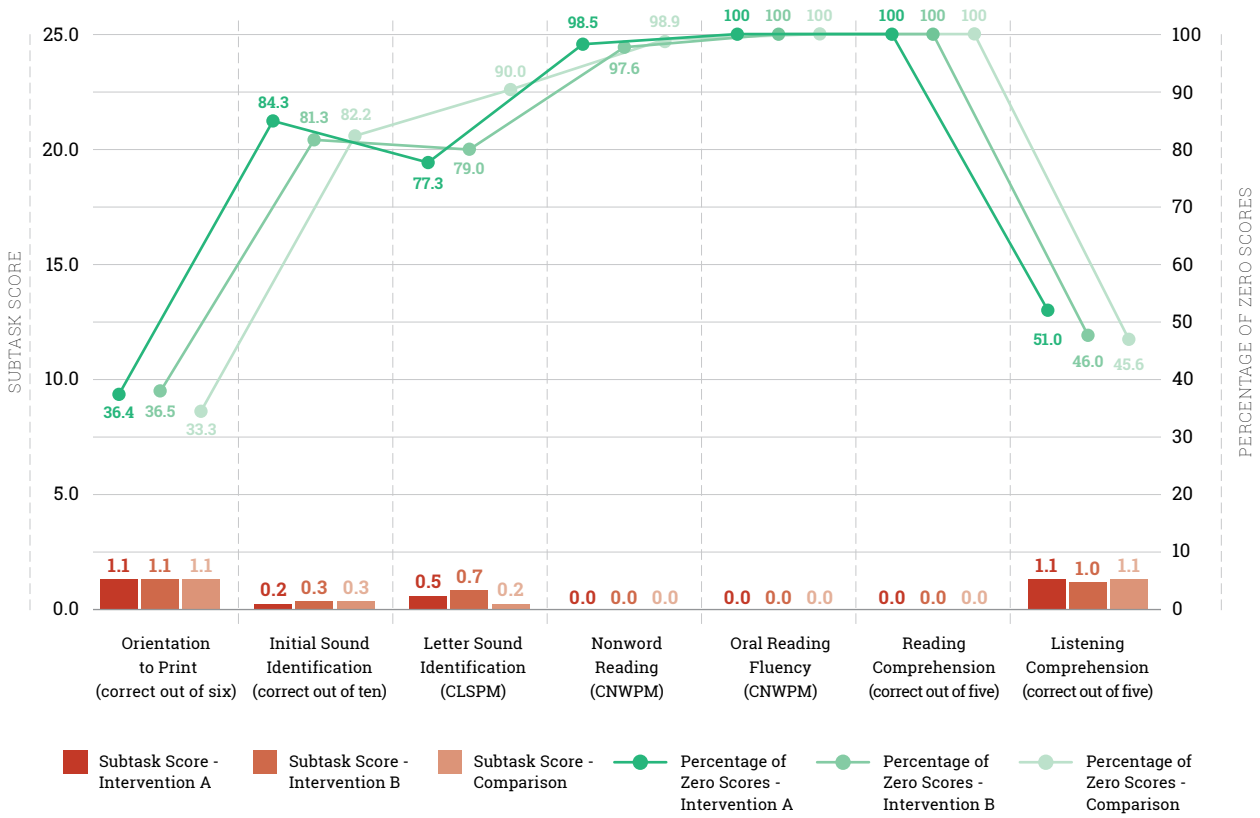
radio programs that modeled strategies and techniques to enhance student learning under the balanced literacy approach. RARE distributed the programs on flash drives.

To understand how the OCLR project impacted students' reading skills, School-to-School International (STS) and RARE conducted EGRAs twice during the project. Baseline data were collected in October 2015, and endline data were collected in May 2017. During and immediately following the endline data collection, STS also conducted semi-structured, end-of-project (EOP) interviews with the OCLR project management, pedagogical counselors, school directors, teachers, government representatives, and other stakeholders. The interviews were designed to explore any lessons learned from the project's implementation, elicit data that would lead to better understandings about how the project impacted students, and allow funders and researchers to assess the potential scalability of the OCLR project.

The following report presents a summary of lessons learned from project implementation, EGRA results, and scalability assessments.

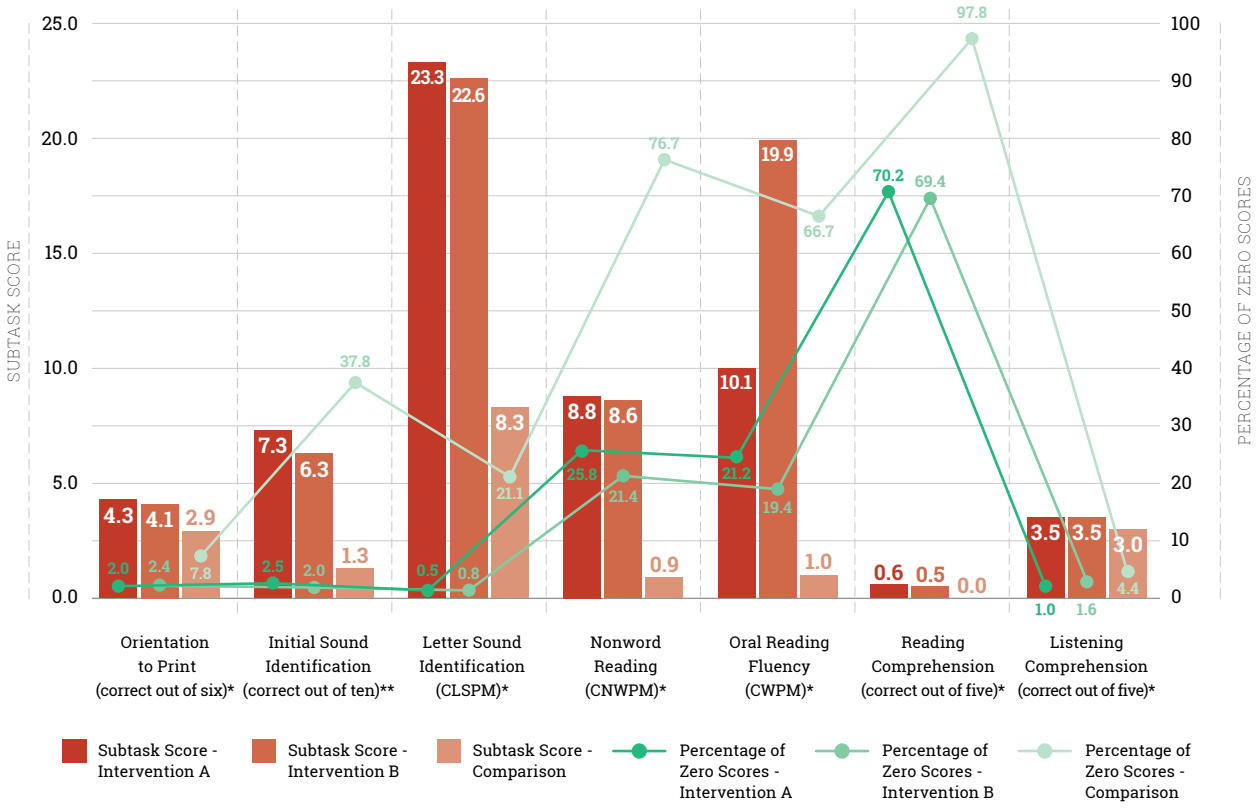
Key Findings

Figure 1: **Mean Results by EGRA Subtask on the EGRA Instrument 1 at Baseline**[5]



5 Mean score estimates are noted above each bar in the graph. The confidence interval (CI) is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: *N* All students=540; *N* Intervention A=198; *N* Intervention B=252; *N* Comparison=90.

Figure 2: **Mean Results by EGRA Subtask on the EGRA Instrument 2 at Endline**[6]



- **Students whose teachers were trained by the OCLR project outperformed their peers whose teachers did not receive support from the project.** At baseline, all students lacked the foundational reading and pre-reading skills measured by EGRA subtasks. At endline, students in interventions A and B achieved statistically significant higher average scores on all endline subtasks compared to the average score of students in the comparison group, whose teachers did not receive training through the OCLR project. Specifically, on the letter sound identification subtask, students in interventions A and B identified 23.3 and 22.6 correct letter sounds per minute (CLSPM), respectively, compared with 8.3 CLSPM in the comparison group. On the nonword reading subtask, students in intervention A identified 8.8 correct nonwords per minute (CNWPM), students in intervention B identified 8.6 CNWPM, and students in the comparison group identified only 0.9 CNWPM. On the oral reading fluency (ORF) subtask, students in interventions A and B read 10.1 and 9.9 correct words per minute (CWPM), respectively, as compared with 1.0 CWPM by students in the comparison group.
- **Across subtasks at endline, a lower proportion of students in both interventions A and B received zero scores than did those in the comparison group.** On the initial sound identification subtask, the proportion of comparison group students who were unable to identify a single initial sound was 37.8 percent, as compared with 2.5 percent and 2.0 percent of students in interventions A and B, respectively. On the letter sound identification subtask, less than one percent of students received zero scores in both interventions A and B, while 21.1 percent of comparison group students received zero scores. Additionally, 76.7 percent of

6 An asterisk (*) indicates the average score of both intervention groups were significantly higher than the comparison group at $p<0.05$, and no significant difference was detected between the mean scores of intervention A and intervention B groups. Two asterisks (**) indicate that the mean scores of both intervention groups were significantly higher than the comparison group and that intervention A's average score was significantly higher than both the comparison group and intervention B at $p<0.05$. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: *N*All students=540; *n*Intervention A=198; *n*Intervention B=252; *n*Comparison=90.

comparison group students were unable to decode a single item on the nonword reading subtask, but only 25.8 percent and 21.4 percent of students in interventions A and B received zero scores.

- **EGRA results indicate that students in intervention A—those whose teachers were trained by the OCLR project and had access to tablets with the Stepping Stone app for supplemental learning—did not significantly outperform their peers in intervention B, whose teachers received the same training but did not have access to tablets.** Students in interventions A and B outperformed their peers in the comparison group across all seven EGRA subtasks at endline. However, the only statistically significant difference in performance between the intervention groups was on the initial sound identification subtask; on this subtask, students in intervention A were able to correctly identify 1.0 initial sound more than students in intervention B.
- **Tablets for remote learning have high potential in the Malian context, as the lack of resources limit the ability of pedagogical counselors to provide enough in-person support to help teachers incorporate new pedagogy into their classrooms.** However, because the OCLR project did not offer substantial amounts of training content on the Stepping Stone app—only three videos—it is unclear if this was sufficient to provide an added value to teachers and students in intervention A.

II. Project Description

RARE is a Malian nonprofit organization dedicated to improving education in the country. It implemented the OCLR project to enhance literacy instruction in Grades 1 and 2 by training teachers on the balanced literacy approach, thus strengthening their skills and competencies. By improving reading and writing instruction, the OCLR project sought to improve the early grade reading skills—specifically, pre-reading and foundational skills—of students taught by the OCLR-trained teachers. Implemented in partnership with the Malian Ministry of Education (MEN) and with EDC as an implementing partner, the OCLR project promoted language instruction and reading materials in Bamanankan, a mother tongue language in the Sikasso region of Mali.[7]

RARE developed the OCLR project with the knowledge that while teachers often receive training through workshops, a key challenge in a resource-lean environment like Mali is providing sufficient ongoing support as teachers implement new strategies in their classroom.[8] To address this challenge, the OCLR project included two components:

1. Workshop-based teacher training and IRI, 30-minute radio programs that describe specific strategies and techniques, to enhance student learning under the balanced literacy approach. RARE also provided follow-up monitoring and support for teachers through visits from pedagogical counselors.
2. On-demand teaching resources through the Stepping Stone app, including three 90-minute videos with demonstrations of strong reading and writing instructional techniques, to provide teachers support in implementing the balanced literacy approach.

The first component of the OCLR project consisted of a seven-day teacher-training workshop with Grade 1 teachers and school directors from 36 schools[9] in three different centre d'animation pédagogique (CAP) areas in Mali.[10] In the workshop, RARE trained participants on the balanced literacy approach, which was developed under

7 In Mali, there are two curricula at the primary level: the classic curriculum, in which students are taught in French from Grade 1, and the bilingual curriculum, in which students are taught in a national language in Grade 1 and transition to receiving partial instruction in French in Grade 2.

8 Réseau d'Acteurs pour le Renouveau de l'Education. (2014). *Using Mobile Technology to Improve the Teaching-Learning of Reading-Writing in Bilingual Curriculum Schools of the Education Region of Bougouni, Mali*. (Unpublished proposal for funding)

9 RARE offered training and follow-up support to teachers in 36 schools, though the research presented in this report applies only to the 20 intervention and five comparison schools selected to be part of the research sample. See Sample for more details.

10 In Mali, education districts are called CAPs. CAPs refer to the geographic area and the schools under the auspices of an education district office.

USAID's Programme Harmonisé d'Appui au Renforcement de l'Education (PHARE) program with the collaboration of the MEN.[11] In the balanced literacy approach, teachers structure lessons and activities over the course of a week to reinforce students' skills across a maximum number of the competency areas described in the standards framework. Teachers weave together activities related to language mechanics (e.g., sound-symbol relationships, phonics, etc.) with activities targeting reading of connected and meaningful text and activities wherein students must use what they know about language to create their own writings.[12] The balanced literacy approach consists of seven teaching strategies:[13]

1. Games for learning language mechanics: Teachers are provided with a repertoire of songs and games to play with students in the classroom. The games target students' phonemic and phonetic awareness, helping them master sound-symbol relationships and providing the building blocks for becoming successful readers and writers.
2. The class news: Students generate a few sentences about real life experiences from the previous day that the teacher then writes on the blackboard with assistance from the students. Students eventually tell the teacher which letters and writing conventions to use. The sentences provide teachers with an engaging strategy for students to practice decoding, reading, fluidity, and comprehension.
3. Find what you know: The teacher provides students with a text, either written on the blackboard or distributed as a handout. Students are asked to spend time identifying what they already know in the text. They may identify letters, whole words, punctuation marks, or other elements of the text. For each element identified, the teacher takes the opportunity to review and discuss that element with the whole class, thus reinforcing everyone's understanding of that item.
4. Guided decoding: The teacher works on decoding a given text (written on the blackboard or distributed as handouts) with the class by isolating single letters, letter blends, or syllables and then asking the class to read the indicated element. Students' attempts to associate a sound with the given symbols are guided as needed. The teacher also ensures that students develop an understanding of the meaning of each word that is decoded. Once the entire text is decoded, the teacher invites the whole class or an individual student to read the text backward, read randomly isolated words, read similar texts, propose a different ending for the text, or other engaging activities.
5. Invented writing: The teacher asks students to write out words that are not written anywhere else in the classroom; students must rely only on what they know of the sound-symbol relationships or the letters of the alphabet of the language in which their class is working. Students are encouraged to say the words slowly to themselves, stretching them out to hear all the sounds, and then to write the letter or letters they believe correspond to those sounds. The teacher does not correct the students' individual work but guides the class as a whole by pinpointing spellings that are appropriate representations of the sounds.
6. Guided reading: The teacher reads a story designed for young children to the class. Students' comprehension skills are strengthened as they are asked to discuss various cues related to the text—such as illustrations or the story's title—and to make predictions about what the story is about or what might happen next. The teacher continues to read portions of the story, frequently stopping to ask students to react, either orally or in writing, to what they are hearing. In the end, students are asked to retell or act out the story, or develop another ending or alternative story with similar characters, plot, or outcomes.

11 USAID PHARE was a five-year program implemented nationally that supported the MEN's efforts to improve the quality of early grade education with a focus on literacy. See http://idd.edc.org/projects/mali-usaidphare-program-programme-harmonis%C3%A9-dappui-au-renforcement-de-leducation.

12 Ministère de l'Éducation de l'Alphabétisation et des Langues Nationales. (2009, October). *Guide de l'enseignant à l'Approche Équilibrée: Niveau I*. Retrieved from http://pdf.usaid.gov/pdf_docs/PA00J3HG.pdf

13 Réseau d'Acteurs pour le Renouveau de l'Education. (2014). *Using Mobile Technology to Improve the Teaching-Learning of Reading-Writing in Bilingual Curriculum Schools of the Education Region of Bougouni, Mali*. (Unpublished proposal for funding)

7. Guided writing: Students and their teacher approach guided writing as a series of four phases: pre-writing, first drafts, revision, and final publication. In pre-writing, the teacher provides a focus for the writing activity by suggesting a theme or writing pattern; the teacher helps students enrich their vocabulary around that focus. Students then write their first drafts, concentrating on putting their ideas to paper rather than on perfect spelling or punctuation. In the revision stage, a student shares with the teacher or other students what he or she has produced, revising as needed for greater clarity or to correct mistakes. In the publication stage, the students post their work in the classroom or share by reading their stories aloud.

During the teacher-training workshop, RARE distributed teaching kits that included balanced literacy guides as well as songs with written lyrics, games, and sample texts. RARE also gave teachers flash drives containing IRI, 30-minute radio programs that describe specific strategies and techniques to enhance student learning under the balanced literacy approach. Following the training, RARE pedagogical counselors conducted periodic visits to teachers to monitor their use of the balanced literacy approach and to provide broad support for their literacy instruction.

The second component of the OCLR project—and the key technological innovation being tested—was the provision of teaching resources and videos accessible through the Stepping Stone app on tablets that allowed teachers to access on-demand support for implementing the balanced literacy approach. The OCLR project team uploaded three 90-minute videos with demonstrations of strong reading and writing instructional techniques to the Stepping Stone app. These videos were filmed in real Bamako classrooms and allowed teachers to review good practices after the project workshops were completed. Half of the teachers who participated in the OCLR project's workshop received tablets with the Stepping Stone app and its teaching resources.

Because the OCLR project spanned two academic years, teachers who were trained in the balanced literacy approach advanced with their students, teaching Grade 1 the first year and Grade 2 the second year. In total, the OCLR project trained 36 teachers who taught roughly 1,800 students in the Sikasso region of Mali.[14]

III. Research Purpose and Design

The goal of the OCLR project was to improve reading and writing instruction in Grades 1 and 2 in Mali by training teachers on the balanced literacy approach, thus strengthening their competencies and skills in teaching literacy. By improving reading and writing instruction, the OCLR project sought to improve the early grade reading skills—specifically, pre-literacy and foundational skills—of students taught by the OCLR-trained teachers. The project also incorporated two key technological innovations: IRI programs that were accessible to all participating teachers, and tablets with the Stepping Stone app that were provided to half of the participating teachers. The research conducted by STS and RARE sought to answer the following questions specific to the OCLR project:

1. What is the effect of the balanced literacy approach, including IRI, on students' reading outcomes?
2. What value is added by providing tablets with the Stepping Stone app to teachers for continued use after the in-person workshop training?

In addition, EOP research was conducted to answer the following supplemental questions common to all ACR GCD grantees:

1. How successful was the rollout of the project?
2. How did the project influence or impact adults' (teachers, parents, community members) knowledge, skills, or attitude regarding their role in helping children read?

14 Twenty intervention and five comparison schools were part of the research design of the OCLR project. Although the project reached about 1,800 students, findings presented in this report should not be generalized outside of the research population.

3. How did the project influence certain subsets of the student population more than others based on identifiable contextual factors?
4. How much did the development, implementation, and management aspects of the project cost?
5. Are this project and technology suitable for scaling?

To answer these research questions, STS and RARE collected EGRA data twice during the project. Baseline data were collected in October 2015, and endline data were collected in May 2017. Qualitative and cost data were also collected to answer ACR GCD's supplemental questions.

The research design for the OCLR project included two intervention groups and a comparison group to answer each of the project-specific research questions and isolate the impact of the Stepping Stone app on students' reading gains.

Sample

RARE identified 36 schools in the Sikasso region to implement the OCLR project, specifically in three CAPs. To be eligible for selection into the RARE sample, schools had to meet the following initial criteria:

- Be a public, community, Bamanankan-language bilingual curriculum school located less than 50 kilometers away from their CAP office
- Have Grade 1 students enrolled for the 2015-16 academic year
- Have a school director who was previously trained in the bilingual curriculum

STS determined that 25 schools—20 intervention and five comparison—would be sufficient to measure the impacts of the OCLR project in its pilot implementation period. Thus, RARE randomly selected 25 schools from the total of 36 eligible and randomly assigned ten schools to intervention A, ten schools to intervention B, and five schools to the comparison group.

The school director and one teacher from each intervention school participated in the project. In intervention A, teachers and school directors attended the in-person teacher training workshop and received follow-up visits from pedagogical counselors; teachers also received tablets with the Stepping Stone app for use after the workshop. In intervention B, teachers and school directors attended the in-person workshop and received follow-up visits from pedagogical counselors; however, teachers did not receive tablets. Teachers and school directors at comparison schools did not receive any support from the OCLR project.

Selection of the EGRA student sample took place during the baseline data collection. RARE intended to randomly select 25 students in Grade 1 at each school for a total of 250 students in intervention A, 250 students in intervention B, and 125 students in the comparison group. In instances in which there were fewer than 25 students in a classroom, the OCLR project team oversampled from another school in the same intervention group to reach the total quota. The team was unable to meet the 25 students per school quota in both intervention A and comparison group schools, and thus oversampled in intervention B schools. At baseline, 637 students were assessed—232 students in intervention A, 312 students in intervention B, and 93 students in the comparison group.

At endline, RARE attempted to assess the same students who were assessed at baseline. However, several factors impeded their ability to do so. First, a teachers' strike occurred on the first day of endline data collection, which hindered the assessors' ability to locate students from baseline (see Considerations). Second, because the OCLR intervention spanned two academic years, Grade 1 students who were assessed at baseline and did not graduate to Grade 2 were excluded from the sample. Third, some of the students assessed at baseline dropped out or migrated. Finally, one school in intervention B changed from offering a bilingual curriculum, a requirement of participation in the project, to offering a classic curriculum with French-only instruction; all students at this school were dropped from the OCLR project sample.

Table 1: **Research Design of Project Groups at Baseline**[15]

Intervention Group	Number of Schools/Teachers
Intervention A: Teacher training and tablets	10
Intervention B: Teacher training	10
Comparison	5

Because of these challenges, almost one-third of students assessed at baseline were not assessed at endline; instead, replacements were selected from students with similar profiles. Replacement students were randomly selected from the same classroom as the students they replaced, and assessors attempted to replace students from baseline with students of the same gender at endline.[16] In total, 540 students were assessed at endline: 372 students from baseline and 168 replacement students who were assessed at endline only. Table 2 shows the breakdown of the endline EGRA student sample by gender and group.

Table 2: **EGRA Sample Characteristics at Endline**

Gender	Intervention A: Training and Tablet	Intervention B: Training	Comparison	Total: All Students
Boys	103	129	54	286
Girls	95	123	36	254
Total	**198**	**252**	**90**	**540**

STS, with support from a team of Bamako-based Malian consultants and a representative from World Vision Mali, conducted EOP interviews between May 1 and June 5, 2017, during and following the endline EGRA data collection (see End-of-Project Interviews). EOP interview details are provided in Table 3.

15 Because one school from intervention B became a classic curriculum school, there were only nine schools and teachers in intervention B at endline.

16 Among replacement students, 20.8 percent were of a different gender than the students they replaced from baseline.

Table 3: **EOP Interview Sample**

Type of Interview	N	Description
Project management	2	One RARE project manager and one EDC staff member
Pedagogical counsellor	3	One pedagogical counselor from each CAP
School director	8	Four directors from intervention A schools and four from intervention B schools
Teacher	8	Four teachers from intervention A schools and four from intervention B schools
Stakeholder	6	Five MEN representatives and one USAID/Mali staff member
Total	**27**	

Project management interviews were conducted with RARE and EDC staff members. A pedagogical counselor from each of the three CAPs, as well as school directors and teachers from eight schools—four from intervention A and four from intervention B—were interviewed.[17] A team member of World Vision Mali conducted interviews with MEN representatives from the intervention area. STS interviewed RARE project management and a staff member from USAID/Mali by phone.

IV. Fieldwork Preparation and Data Collection

EGRA Instrument

The EGRA used for the OCLR project measured reading skills in Bamanankan. Although French is the national language of Mali, Bamanankan is the most widely spoken language in Mali and is the dominant mother tongue language in the implementation areas of the OCLR project.

Two previous EGRAs have been conducted in Mali. In 2009, RTI International (RTI) and the Centre de Promotion de la Citoyenneté pour un Développement durable à la base (CEPROCIDE), a Malian nongovernmental organization (NGO) and research firm, conducted an EGRA in Bamanankan, Bomu, Fulfulde, and Songho—all mother tongue languages in Mali. From 2014 to 2015, RTI conducted an EGRA baseline for students who had completed second grade. However, since teachers trained as part of the OCLR project were to work with students in Grades 1 and 2, STS adapted RTI's EGRA instrument specifically for students in these grades.

STS conducted a six-day adaptation workshop in October 2015. Both ACR GCD grantees in Mali—RARE and Œuvre Malienne d'Aide à l'Enfance (OMAES)—participated in the workshop and used the same Bamanankan EGRA instruments. In addition to STS, RARE, and OMAES, representatives from Direction Nationale de la Pédagogie, EDC, Direction Nationale de l'Enseignement Normal, Direction Nationale de l'Enseignement Fondamental, USAID/Mali, and World Vision also attended the workshop and participated in the adaptation of the subtasks. During the workshop, new content was developed for three subtasks on the Grades 1 and 2 EGRA: initial sound identification, letter sound identification, and nonword reading. Additionally, workshop participants developed reading passages and corresponding questions for the ORF and reading comprehension subtasks, as well as the passages and corresponding questions for the listening comprehension subtask.

The EGRA was pretested at a rural school on the outskirts of Bamako where conditions resembled those where RARE and OMAES would implement their interventions. The results from the pre-test showed a high number of zero scores for all students on all subtasks, except listening comprehension. In light of these results, the workshop

17 STS, in collaboration with RARE, purposefully selected the schools in which EOP interviews were conducted based on the endline data collection schedule.

participants modified the subtasks and piloted three new versions of the instrument, which resulted in a marginal decrease in the number of zero scores. Upon review of the data, workshop participants noted these marginal decreases and finalized the instruments with approval from MEN.

The final EGRA instruments included the following seven subtasks: orientation to print, initial sound identification, letter sound identification, nonword reading, ORF, reading comprehension, and listening comprehension. Two versions of the EGRA instrument were developed: the EGRA instrument 1, to be administered at baseline, and the EGRA instrument 2, to be administered at endline. Four of the subtasks—orientation to print, initial sound identification, letter sound identification, and nonword reading—were identical across instruments. The passages used in the ORF and reading comprehension subtasks were different on the two instruments: the EGRA instrument 1 contained a 50-word passage plus five corresponding comprehension questions, while the EGRA instrument 2 contained a 52-word passage plus five corresponding comprehension questions. The passages for the listening comprehension subtask were also different on EGRA instruments 1 and 2, although both contained five comprehension questions.[18]

Institutional Review Boards

Institutional review boards (IRBs) are responsible for ascertaining the acceptability of proposed research regarding institutional commitments and regulations, applicable laws, standards of professional conduct and practice, and ethical and societal norms. IRBs examine subject recruitment procedures, proposed remuneration, and the informed consent process. IRBs also evaluate the potential risks and benefits to participants, as outlined in each protocol.

The ACR GCD team, in consultation with in-country partners, determined that there was no appropriate local IRB process for Mali. To address this, RARE provided the MEN with details about the OCLR project's research design and obtained a letter of approval to proceed.

Baseline EGRA

The baseline EGRA assessor training, which included trainees for both the RARE and OMAES data collections, took place from October 12 to 16, 2015. RARE recruited the assessors. All candidates had previous survey and assessment experience, including working with an Annual Status of Education Report, a widely used international literacy test to determine the reading level of early-primary school students. Many candidates also previously served as EGRA assessors for other projects. STS, with support from RARE and OMAES, trained assessors on how to administer the Bamanankan EGRA, on both paper and on tablets using electronic data collection software Tangerine.[19]

The training included a variety of simulation methods and a half-day spent practicing data collection with students in rural schools near Bamako. In addition to student reading assessments, a brief student questionnaire was developed and piloted during the assessor training. The questionnaire was used to gather data on the contextual factors that could affect reading proficiency, such as availability of Bamanankan reading materials and access to an adult at home who can read.

18 The ORF, reading comprehension, and listening comprehension passages from instruments 1 and 2 were not equated as there was insufficient pilot data available from the adaptation workshop. As a result, findings from these three subtasks are presented as baseline scores and endline scores rather than gains and should not be directly compared. See EGRA Data Analysis and Considerations.

19 Tangerine® (http://tangerinecentral.org/) is an electronic data collection software designed for use on mobile computers, including netbooks, tablet computers, and smartphones. Its primary use is to enable recording of children's responses in oral early grade reading and mathematics skills assessments—specifically EGRA and Early Grade Mathematics Assessment—and of interview responses from children, teachers, and principals on home and school context information.

As part of their training, assessors participated in assessor accuracy testing.[20] Assessor accuracy testing is conducted to ensure consistency in scoring among assessors and to measure the degree to which assessors agree in their assessment decisions. At least 90.0 percent consistency is the minimum requirement; this means that at least 90.0 percent of assessors' ratings must be consistent with the list of acceptable responses. During the assessor accuracy testing on the final day of training, two candidates were unable to meet this threshold and did not participate in the baseline operational data collection.

Table 4: **Fieldwork Preparation and Data Collection Timeline**

Task	Dates
EGRA instrument adaptation workshop	October 5–10, 2015
Baseline assessor training, including pilot test and assessor agreement	October 12–16, 2015
Baseline EGRA operational data collection	October 19–28, 2015
Endline EGRA refresher training	April 24–28, 2017
Endline EGRA operational data collection	May 1–12, 2017
EOP interviews	May 1–June 5, 2017

Following assessor training, assessors collected operational baseline EGRA data between October 19 and 28, 2015.

Endline EGRA

Before operational data collection, RARE and OMAES identified assessors to conduct endline EGRAs. CEPROCIDE conducted a refresher training session for assessors from both organizations from April 24 to 28, 2017.[21] The training included review sessions on the EGRA endline instrument and administration protocols. Although assessors practiced administering the EGRA during the refresher training, CEPROCIDE did not conduct assessor accuracy testing using commonly accepted protocols, and no agreement between assessors was calculated during the training (see Considerations). Although ten assessors were trained to conduct endline data collection for RARE, one assessor was not able to participate in the data collection.

The endline EGRA was conducted from May 1 to 12, 2017, in 19 intervention schools and five comparison schools. A teachers' strike started on the first day of the data collection, which led to the establishment of a contingency plan in collaboration with MEN. MEN representatives directly contacted teachers and students to arrange appointments to collect EGRA data in a neutral location. The strike was resolved after one-day, and the remainder of the endline EGRA data collection was held at schools, as originally planned. It is possible that the strike may have caused residual disturbances, such as increased teacher or student absences, which may have interfered with the data collection (see Considerations).

20 Assessor accuracy testing is similar to interrater reliability testing. According to the EGRA Toolkit (2nd Edition), assessor accuracy refers to the testing conducted during training, while interrater reliability is conducted during operational data collection.

21 World Vision directly contracted CEPROCIDE to conduct the assessor training due to their experience training assessors and administering EGRAs in Mali.

End-of-Project Interviews

STS coordinated the administration of EOP interviews between May 1 and June 5, 2017, during and following the endline EGRA data collection.[22] The purpose of the interviews was to explore the contextual factors that may have impacted the variations in implementation and results among schools and students. They also explored the potential scalability of the project. EOP interviews were conducted with five groups of project participants: project management, pedagogical counselors, school directors, teachers, and stakeholders.

STS conducted project management interviews by phone with key staff from RARE and from EDC. Project management interviews consisted of open-ended questions related to general project information, the intervention timeline, characteristics of the implementing organizations, perceptions of project design and implementation quality, and considerations for scalability.

STS hired a team of five local consultants—one supervisor and four data collectors—with prior experience in administering qualitative interviews in the education sector to conduct interviews with pedagogical counselors as well as school directors and teachers at intervention group schools. Pedagogical counselors were asked 16 open-ended questions about their role in the project, implementation challenges, and recommendations for improvement. School directors and teachers at intervention schools were asked 12 and 18 open-ended questions, respectively. Questions related to their use of technologies in the classroom, challenges faced in implementing the project with fidelity, training and support they received through the project, and, depending on their intervention group, usage of the Stepping Stone app.

A local representative of World Vision Mali conducted interviews with MEN representatives from the intervention area. MEN representatives were asked nine questions related to the national and political context surrounding education and technology, relevance of the OCLR project to education policy priorities, relative advantage of the project in comparison with existing policies or programs, and potential for scalability.

Finally, STS conducted a phone interview with a representative of USAID/Mali to discuss agency priorities for education programming, perspectives on the use of technology in educational contexts, and potential for scalability.

V. Project Implementation

The OCLR project commenced in February 2015, conducted its baseline assessment in October 2015, and began implementation in schools in January 2016. Implementation ended in April 2017, and endline data collection occurred from May 1 to 12, 2017.

This section presents implementation challenges, solutions, and successes that help answer the ACR GCD research question: *How successful was the rollout of the intervention?*

Development

For the first component of the OCLR project—teacher training in the balanced literacy approach coupled with IRI—RARE adopted an existing literacy approach, including training materials, to train the teachers in both intervention groups. EDC, in collaboration with the MEN, had previously implemented the balanced literacy approach in Mali through USAID PHARE; thus, the OCLR project did not have to dedicate considerable time or resources to the development of the literacy approach. However, the teacher-training component did necessitate the production of materials to distribute to schools, including pedagogical booklets and instructional materials, picture cards, and teacher observation checklists. A pedagogical adviser from MEN provided guidance on the production of these materials.

22 Due to security concerns in Mali, STS staff were not present during data collection.

For the second component of the OCLR project—teaching resources and videos accessible through the Stepping Stone app on tablets—RARE relied heavily on EDC, both for technical expertise and to produce content for the platform. EDC originally developed the Stepping Stone app for the USAID Projet d'Appui aux Jeunes Entrepreneurs (PAJE-Nièta),[23] and, accordingly, significant technological resources were not needed for its adaptation to the OCLR project. EDC also managed the production of the teacher training videos that would be available through the app. A pedagogical adviser helped develop scripts for the videos, and EDC hired a videographer in Bamako to film the videos; EDC staff edited and formatted the videos for the tablets. The three videos were filmed at three different bilingual-curriculum schools; two Grade 2 teachers and one Grade 1 teacher were filmed. The videos illustrate teaching practices using three exercises within the balanced literacy approach: news of the classroom, guided reading, and guided writing. EDC staff provided technical assistance by loading the Stepping Stone app onto the tablets.

Overall, the development phase of the project did not present major challenges or delays because RARE relied on the strong technical expertise of EDC and utilized existing training concepts, materials, and technologies.

Implementation

The implementation phase of the project faced some challenges—a delay in the training, irregular support visits from pedagogical counselors, and teacher strikes—that may have impacted RARE's desire to deliver the components of the project effectively.

The OCLR intervention phase spanned two school years and featured a single cohort of teachers who followed their students from Grade 1 to Grade 2. Although the project initially intended to conduct teacher training near the beginning of the academic year in September 2015, delays in finalizing the research design—including the school selection—necessitated a change in the start of implementation. Instead, RARE conducted the seven-day training from December 23 to 30, 2015. This delay meant that teachers were trained at the start of the second term of the first year of intervention. According to feedback provided in EOP interviews, the OCLR project team felt that the intervention would have had a more important impact if it had been implemented for two full school years; they cited a belief that, in the Malian education system, students and teachers are often most engaged during the first term. RARE held a follow-up, in-person training in the second year of implementation from August 22 to 29, 2016. Additionally, the USAID Selected Integrated Reading Activity (SIRA) project[24] also began offering training on the balanced literacy approach for Grade 1 teachers in all of the OCLR project's schools during the 2016-17 academic year. Although this did not necessarily impact previously trained teachers in the intervention schools, who had since moved up to Grade 2 with their students, it is possible that the training USAID SIRA provided to comparison group schools had a contamination effect: Grade 1 teachers in comparison schools who received balanced literacy approach training from USAID SIRA may have shared knowledge with Grade 2 teachers in the same school, thus changing their teaching practices.

Following RARE's teacher training workshops, and to ensure implementation of balanced literacy approach techniques, pedagogical counselors visited teachers throughout the implementation period to provide additional support. These visits were also an opportunity to observe teachers' classroom practices and students' reading and writing skills. In EOP interviews, project management expressed that they would have liked more participation from the pedagogical counselors; although RARE tasked them with providing regular coaching and monitoring to teachers, counselors were overburdened and unable to visit teachers regularly. Finally, a nationwide teacher strike during the second year of the OCLR project also impeded follow-up visits—all three pedagogical counselors who were interviewed cited the teacher strike as a major challenge for implementation.

23 See http://sstone.edc.org/en for more details.

24 USAID SIRA is a five-year project that began in 2016 and aims to improve reading skills of primary school students in the Koulikoro, Ségou, and Sikasso regions of Mali.

Despite the litany of challenges faced by teachers due to a nationwide strike and the irregular follow-up of overburdened pedagogical counselors, project management noted that a key strength of the OCLR project's design was that intervention A teachers could access videos on the tablets at any time. This on-demand access to additional training was thought to have mitigated the inability of the Government of Mali (GoM) to provide subsequent in-person training. Furthermore, fidelity of implementation (FOI) data indicate that most teachers across both intervention groups were practicing the balanced literacy approach daily (see Fidelity of Implementation).

Management

The project benefited from the support of multiple organizations with which RARE had preexisting working relationships. Members of RARE's team have longstanding relationships with peers at the MEN and received ongoing support from them. MEN's national, regional, and district education officials were closely involved in project implementation, and RARE project managers were very satisfied with the MEN partnership on the OCLR project. Representatives from the MEN at the national level conducted monthly visits to RARE schools, and RARE held joint supervisions every semester with regional and local stakeholders from the MEN.

EDC also provided significant technical support, both locally and from the United States home office, on the Stepping Stone app and in analyzing its log data. Because of its experience implementing the balanced literacy approach in Mali through USAID PHARE and USAID SIRA, EDC also provided RARE with technical and implementation guidance. Finally, the OCLR project was also supported through RARE's collaboration with fellow Malian organization and ACR GCD grantee OMAES. The two organizations benefited from sharing venues in enumerator training and from cooperating on the evaluation and some implementation aspects of their respective programs.

No specific management challenges were identified by project managers, although they did express that the irregular monitoring visits by pedagogical counselors were a roadblock that additional management oversight would not necessarily solve. Instead, they suggested that future project management should better engage local community members and find alternative solutions to mitigate pedagogical counselors' busy schedules.

Technology

The main technological component of the OCLR project was tablets loaded with teaching resources and videos accessible through the Stepping Stone app. EDC staff, both in Mali and in the United States, provided technical assistance in loading the Stepping Stone software and videos onto tablets, analyzing the app's usage data, and troubleshooting technical issues throughout implementation.

In EOP interviews, project management noted that the three videos available to teachers via the Stepping Stone app were too few for the nearly two years of implementation. Instead of three videos, they suggested that teachers have access to ten or more videos. Additionally, they recommended improving the quality of videos by including a greater diversity of content. Project management expressed that diverse content might attract users and cover a larger range of competencies for evolving teacher practices. For instance, it would be an improvement to include examples of various levels of teachers—including low-, average-, and high-performing—so that viewers could observe the difference.

Furthermore, project management and pedagogical counselors observed that, in some schools, principals took the tablets and used them more than the teachers. The RARE team noted that, in the future, they would strongly emphasize to school directors that the teacher should be the primary user of the tablet; alternatively, RARE suggested that the project could offer two tablets per school so that both the school director and the teacher could

have access. Nevertheless, all stakeholders, from the teachers themselves to pedagogical counselors monitoring classroom activities, noted that the teachers very frequently used the Stepping Stone app.

Additional challenges were noted regarding the usage of log data recorded by Stepping Stone. Ideally, the app would provide accurate usage data by user—for example, login time, content views, and log-out time for each session—to help program managers understand what materials were most utilized by teachers and to provide usage data to help better understand EGRA results. The Stepping Stone app automatically records when users turn the app on and off, materials viewed, and time spent viewing the materials; these data are automatically recorded and can be manually retrieved by transferring tablet memory cards to a computer. However, the format of the logs—only downloadable in text files—made interpreting and analyzing the logs for monitoring and evaluation (M&E) purposes difficult, as data had to be manually transferred from text files into Excel. Furthermore, the Stepping Stone app did not have login accounts for users, meaning that there was no way to verify who had used the app during a recorded session—a consideration when multiple people were using the same tablet over the course of the project. There were also issues with incorrect time stamps and improper recording of usage times. For example, if a teacher left the app running in the background while using the tablet for other purposes, the Stepping Stone log recorded this as time spent on the app. Because of this unreliability, user data could not be included in this report.

EDC staff responsible for the development of the Stepping Stone app explained that because it is designed to run fully offline, they rely on a manual export procedure to download app logs; this sometimes results in duplicate, missing, or mislabeled logs. They are also dependent upon the time and date-stamp settings of the host device rather than relying on a standardized external source like an internet server, which means that log time stamps are not standardized. EDC is currently working on a version 3.0 of the platform, which should retain the offline mobile learning features—the major advantages of the platform—while also introducing an optional online data synchronization feature. EDC is also considering adding user accounts, though they expressed concerns that this function would introduce complex issues of privacy, data security, and increased training and support costs, in addition to further design and coding work. Finally, EDC hopes to develop analysis tools within the Stepping Stone app that would allow project staff and users to view usage data on-demand.

Despite these issues, there was evidence that access to the tablets helped solve some of the difficulties faced in previous attempts to implement the balanced literacy approach in Mali—specifically, lack of adequate follow-up and support to teachers, and lack of sufficient quantities of the supporting materials. Anecdotal evidence cited by RARE project managers and pedagogical counselors in their EOP interviews indicate that teachers in intervention A with access to tablets had a better mastery of teaching and were more comfortable with balanced literacy techniques than their counterparts in intervention B who had no access to teaching videos or tablets. RARE staff also said that, in general, teachers used the games aspect of the balanced literacy approach very infrequently but that teachers with access to tablets appeared the most likely to employ games in their classrooms. Project staff did not observe significant issues with teachers' use of the technology or experience technical issues with tablets, and pedagogical counselors and teachers themselves confirmed this in EOP interviews.

Fidelity of Implementation

By definition, FOI is the accurate and consistent application of an agreed-upon procedure. FOI research is used to assess the degree to which a project is implemented as intended. Measuring FOI helps implementers and researchers understand and differentiate between what was supposed to happen and what actually happened during the life of a project. When FOI is high and an intervention group experiences gains, then it is possible to associate gains with the intervention; this, in turn, makes it possible to recommend scaling the intervention. FOI also makes it possible to identify which components of an intervention are most strongly associated with

outcomes. When FOI is low and gains are low, it is impossible to know whether the reason for low gains is poor design or poor implementation. FOI research can also be coupled with M&E to provide feedback to implementers during the project cycle to improve adherence to project design in the case of low FOI.[25]

As part of their projects, all ACR GCD Round 2 grantees conducted FOI research during the implementation period. The primary objectives of FOI for grantees were to

1. Understand what FOI is and why it is important throughout the life of the project
2. Identify essential components, activities, and questions for each phase of project implementation
3. Create relevant, project-specific FOI tools to monitor participant adherence to the intervention plan

STS held a series of FOI meetings with each ACR GCD Round 2 grantee to develop project-specific FOI tools and an implementation plan for FOI research. After finishing the FOI sessions, ACR GCD grantees were expected to pilot test their FOI tools and collect data. Grantees were advised to collect a minimum of one round of FOI data; two or more rounds of data collection were ideal. The data collected served several purposes:

1. To indicate where revisions in data collection tools were necessary
2. To highlight where improvements in implementation were needed
3. To attribute impact when combined with assessment data

RARE staff participated in a series of FOI calls with STS, developed FOI tools, and collected paper-based FOI data in January 2017. The FOI tools included interviews with school directors, teachers, CPs, and the project coordinator. Results from FOI interviews with five intervention school teachers, as summarized by RARE in its April 2017 quarterly report, are presented in Table 5.

Table 5: **FOI Results from January 2017**

Question	Yes (%)	No (%)
Have you been trained by the OCLR project?	100.0	0.0
Has the school received an IRI kit?	100.0	0.0
Have you received a Balanced Approach to Literacy Instruction kit?	100.0	0.0
Are the kits kept in the school director's office?	100.0	0.0
Do you practice classroom news every day?	100.0	0.0
Do you practice guided reading every day?	33.3	66.7
Do you practice guided writing every day?	44.4	55.6
Do you practice IRI every day?	44.4	55.6

FOI results indicate that teachers trained by the OCLR project received supplementary materials but that their application of the balanced literacy approach in the classroom may be inconsistent. In future iterations of the project, more regular and detailed classroom observations would help the OCLR project team better understand how the balanced literacy approach is adopted and implemented by teachers.

25 Creative Associates International, Inc. (2015). Fidelity of Implementation (FOI) how-to guide (unpublished). Washington, D.C.: USAID.

VI. EGRA Data Analysis

EGRA data were analyzed using Microsoft Excel and IBM SPSS Statistics. Subtask fluencies and scores were reported with means and standard deviations (SD) relevant to those mean values for the EGRA instrument 1 at baseline and the EGRA instrument 2 at endline.[26] Zero scores were also calculated for all subtasks.[27]

Students who had data at both baseline and endline, as well as replacement students at endline, were used in this analysis. Due to the high rate of replacement at endline, additional analysis was conducted to establish comparability of the student samples at baseline and endline, including all students—total endline sample—and excluding the replacement students—matched sample (see Considerations). Due to the desire to maximize statistical power through a larger sample size, and given the comparability of observable traits and the similarity of performance on the EGRA across subtasks between the total endline sample and matched sample, the total endline sample is used in this report.[28]

For each subtask, decision rules were applied to assess whether outliers would need to be removed. For example, if the time remaining for a timed subtask resulted in a fluency rate that was outside a reasonable range, then that student's fluency rate was not included in the analyses. Reasonable ranges for the time remaining were based on multiple factors, including the rate at which letters or words in the language tested are typically read, the distribution—or relative performance—of students in the sample, and the mean fluency rates with and without the outlier data point(s). After consideration of the reasonable ranges in the data, no outliers were removed.

Table 6 provides details on the EGRA subtasks, including how results were calculated.

Table 6: **EGRA Subtask and Data Analysis Method**

Subtask	Type	Analysis
Orientation to print	Untimed	Orientation to print is measured as the number of questions a student can correctly answer regarding text direction, the concept of a word, or basic knowledge of printed material. Students had the opportunity to answer six questions on this subtask.
Initial sound identification	Untimed	Initial sound identification is measured as the number of correct initial sounds identified out of ten. Initial sound identification is a measure of phonological awareness. Each student had the opportunity to identify ten beginning phonemes that are different from two others in a series of words.
Letter sound identification	Timed	Letter sound identification is measured as the number of correct letter sounds read in one minute (CLSPM). Letter sound identification is a measure of alphabet knowledge. Each student had the opportunity to read up to 100 upper-and lower-case letters.
Nonword reading	Timed	Nonword reading is measured as the number of correct "nonwords" read in one minute (CNWPM). Nonword reading measures decoding. Each student had the opportunity to read up to 50 one- or two-syllable nonwords.

26 SD describes how much observed values vary from the mean. A smaller SD indicates that the majority of values are close to the mean; a larger SD indicates that values are further from the mean.

27 Students receive a zero score if they are unable to correctly identify a single item on a subtask. In this report, zero scores are shown as the number of students or as the percentage of the total students unable to correctly identify a single item on a subtask.

28 These traits included gender, intervention group, and school. Mean scores, proportion of zero scores, and significant differences were compared between the total endline sample and matched sample, and these remained consistent across all subtasks within both samples of students. The total sample was also found to be desirable due to the maximization of statistical power. For more details, see Annex H.

Table 6: **EGRA Subtask and Data Analysis Method** (continued)

Subtask	Type	Analysis
Oral reading fluency	Timed	ORF is measured as correct words read in one minute (CWPM). ORF is a decoding and reading fluency measure. Each student had the opportunity to read 50 words in one minute on the EGRA instrument 1 at baseline and 52 words in one minute on the EGRA instrument 2 at endline. The ORF passage formed the textual basis for the reading comprehension subtask.
Reading comprehension	Untimed	Reading comprehension is measured as the number of correct answers verbally delivered to the assessor based on questions asked about the passage read as part of the ORF subtask. Each student had the opportunity to answer up to five questions on the EGRA instruments 1 and 2.
Listening comprehension	Untimed	Listening comprehension is measured as the number of correct answers verbally delivered to the assessor. Listening comprehension is a measure of vocabulary. Each student had the opportunity to answer five questions on the EGRA instruments 1 and 2 based on a passage read aloud to them by the assessor.

Considerations

Nonequated Passages on the Oral Reading Fluency, Reading Comprehension, and Listening Comprehension Subtasks

The passages used for the ORF and reading comprehension subtasks on the EGRA instrument 1 at baseline and on the EGRA instrument 2 endline were different, as were the passages used for the listening comprehension subtask. The passages for both instruments 1 and 2 for these three subtasks were developed and piloted during the adaptation workshop in October 2015. At that time, the adaptation team selected two passages for the ORF and reading comprehension subtasks and two for the listening comprehension subtask based on observations that the passages were similar in difficulty. However, not enough data were collected during the pilot process to fully equate the two sets of passages. STS's psychometricians attempted to conduct an external anchor equating exercise but, ultimately, were unable to conclusively equate the passages from the EGRA instrument 1 and the EGRA instrument 2.

Student results from the EGRA instrument 1 at baseline and the EGRA instrument 2 at endline are presented separately to emphasize that, although the passages were similar in difficulty, they are not equated. As a result, scores from baseline and endline on these subtasks should not be directly compared. The orientation to print, letter sound identification, and nonword reading subtasks were identical on the baseline and endline EGRA instruments, and results for these subtasks can be compared.

Lack of Assessor Accuracy Testing During Endline EGRA

CEPROCIDE conducted a refresher training session for the endline assessors. Although assessors practiced administering the EGRA during the refresher training, CEPROCIDE did not conduct assessor accuracy testing using commonly accepted protocols, as per the EGRA Toolkit (2nd Edition), and no agreement among assessors was calculated during the training. It is not possible to conclusively determine whether assessors were consistent in their scoring decisions. As a result, the reliability and consistency of the endline data should be considered a significant limitation.

Teacher Strike

The endline EGRA was conducted from May 1 to 12, 2017. A teachers' strike started on the first day of the data collection, which led to the establishment of a contingency plan in collaboration with the MEN. The MEN representatives and RARE staff directly contacted students and family members to arrange appointments to collect EGRA data in a neutral location. The strike was resolved after one day, and the remainder of the endline EGRA data collection was held in schools as originally planned. Given the high rate of replacement students at endline, it is possible that the strike may have discouraged certain students from returning to school after its resolution. This may have introduced bias into the endline student sample, though this is difficult to determine as the project did not collect data on students who did not return after the strike.

Replacement Students

At endline, RARE attempted to assess the same students from baseline. However, due to several challenges (see Sample), almost one-third of students were assessed only at endline; these students were randomly selected to replace students assessed at baseline who were not present at endline. At endline, 372 students were assessed at both time points, and 168 students were replacements assessed at endline only. Assessors attempted to replace students from baseline with a student from the same classroom and of the same gender at endline. In some instances, the replacement met the classroom but not the gender criteria, leading to some discrepancies in the proportion of girls and boys at baseline and endline assessment. Despite the sizable number of replacement students at endline, the overall demographic proportions were largely maintained (see Annex Table H.1 for the sample breakdown by gender and group).

After consideration of the large number of replacements at endline, the comparisons of demographics, and average endline results between the total and the matched samples, STS's psychometricians decided to present EGRA results for the total sample—including replacement students assessed at endline—and analyze by time point and intervention group. Gain scores are not computed between baseline and endline for the total sample; however, differences in mean scores between intervention A, B, and the comparison group and between time points were tested for significance using analysis of variance. In analysis involving more than two variable sub-categories, analysis of variance was used with Bonferroni corrections for multiple testing to examine the significance of differences in scores between each pair of sub-categories (e.g., intervention A, B, and comparison at endline).[29] Differences in the proportion of zero scores were tested for significance using a chi-square test.[30] Furthermore, gain scores for the matched sample are noted in footnotes for the orientation to print, letter sound identification, and nonword reading subtasks.

VII. EGRA Results

This section presents EGRA results to answer the key research questions posed by the OCLR project: W*hat is the effect of the balanced literacy approach, including IRI, on students' reading outcomes?* and *What is the value added of providing tablets with the Stepping Stone app to teachers for use after the in-person workshop training?*

The following section contains findings by intervention group across EGRA subtasks. Results by gender are also presented for all subtasks.

Figures 3 and 4 present subtask results by group at baseline and endline. On the EGRA instrument 1 at baseline, the data revealed that students in all three groups—intervention A, intervention B, and comparison—lacked the foundational reading and pre-reading skills as measured by the EGRA. **Students in interventions A and B achieved significantly higher average scores on all subtasks in the EGRA instrument 2 at endline, in contrast to students in the comparison group.**

29 The Bonferroni correction is a conservative way to conduct statistical significance testing across multiple variable subcategories. It calculates a new pairwise in order to keep the familywise value at 0.05.

30 The chi-square test is a statistical test comparing proportion of students with zero scores that were observed in the data against what was expected.

Figure 3: **Mean Results by EGRA Subtask and Group on the EGRA Instrument 1 at Baseline**[31]

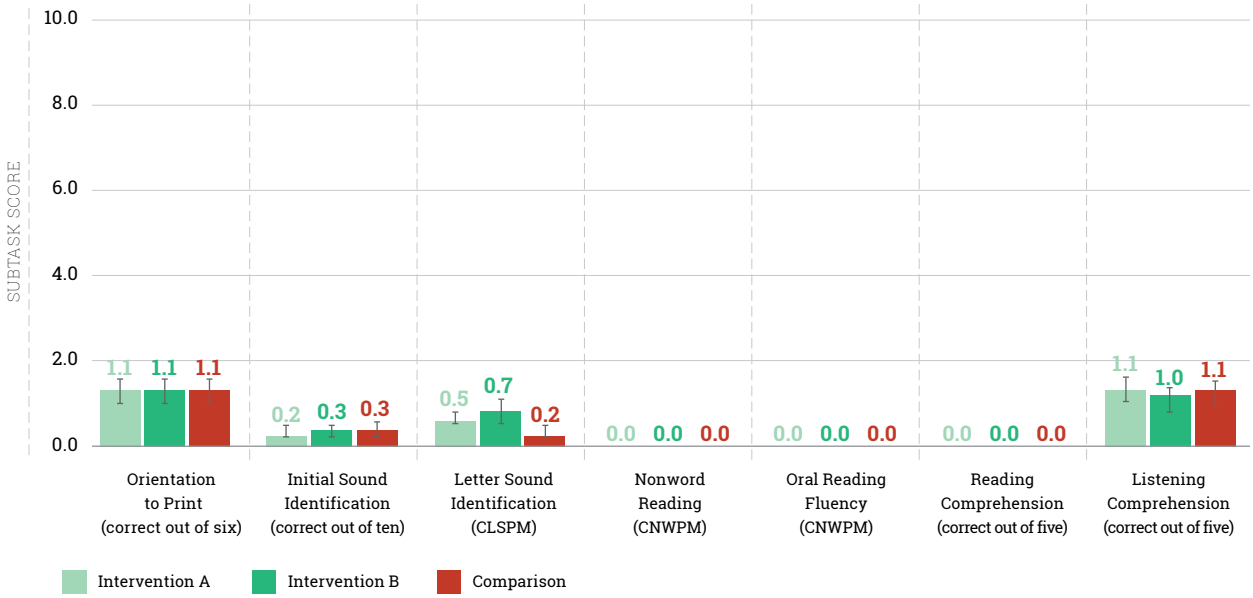
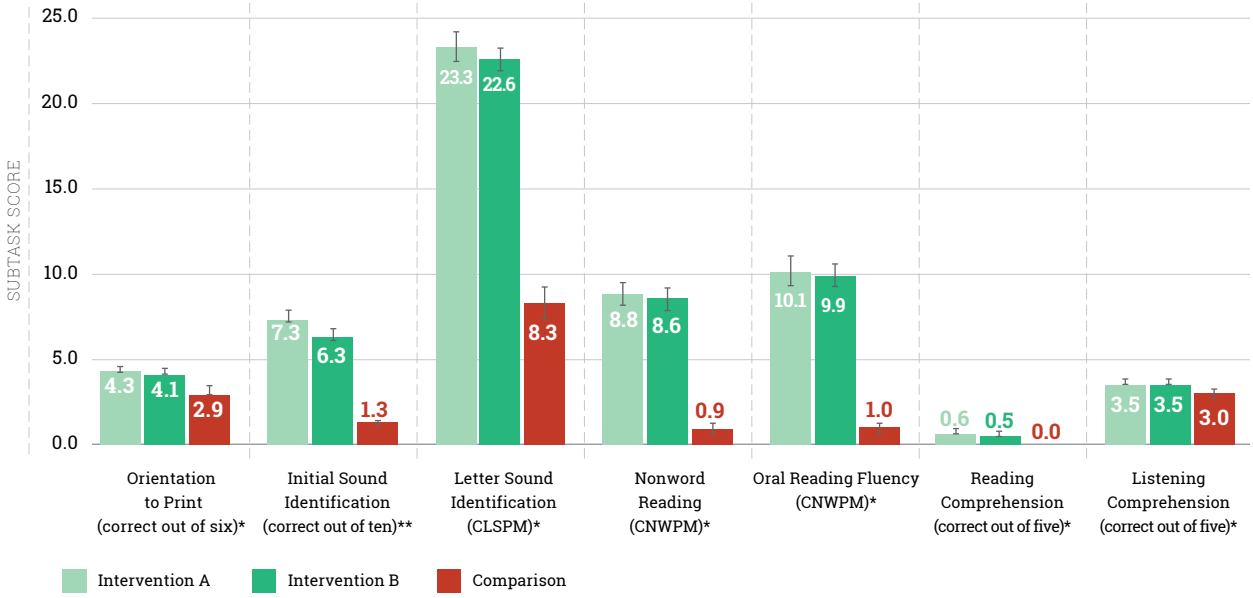


Figure 4: **Mean Results by EGRA Subtask and Group on the EGRA Instrument 2 at Endline**[32]



31 An asterisk (*) indicates that the mean scores of intervention A and intervention B was significantly higher than the comparison group at p<0.05, though the difference between intervention A and intervention B was not statistically significant. Two asterisks (**) indicate that the average score of intervention A was significantly higher than those of both the comparison group and intervention B at p<0.05. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}$=540; $n_{\text{Intervention A}}$=198; $n_{\text{Intervention B}}$=252; $n_{\text{Comparison}}$=90.

32 An asterisk (*) indicates that the mean scores of intervention A and intervention B was significantly higher than the comparison group at p<0.05, though the difference between intervention A and intervention B was not statistically significant. Two asterisks (**) indicate that the average score of intervention A was significantly higher than those of both the comparison group and intervention B at p<0.05. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}$=540; $n_{\text{Intervention A}}$=198; $n_{\text{Intervention B}}$=252; $n_{\text{Comparison}}$=90.

Figures 5 and 6 show the percentage of students who received zero scores at baseline and endline. Across subtasks, interventions A and B had lower proportions of students receiving zero scores at endline than did the comparison group.

Figure 5: **Percentage of Students Receiving Zero Scores by EGRA Subtask and Group on the EGRA Instrument 1 at Baseline (%)**[33]

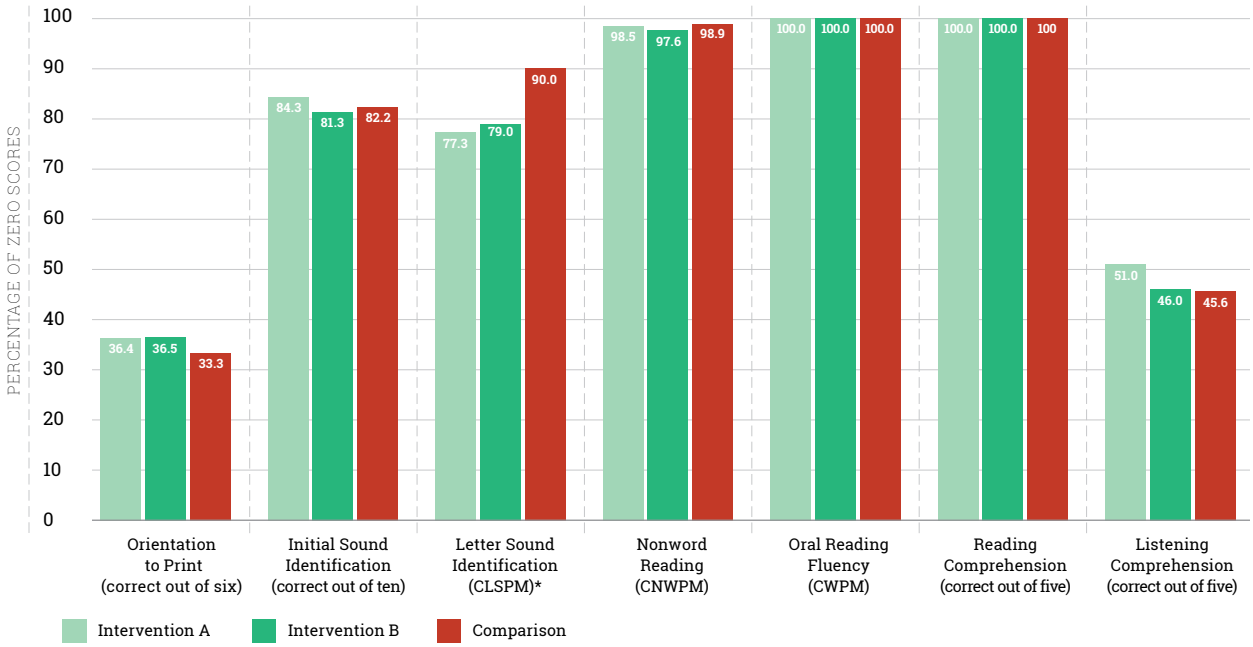
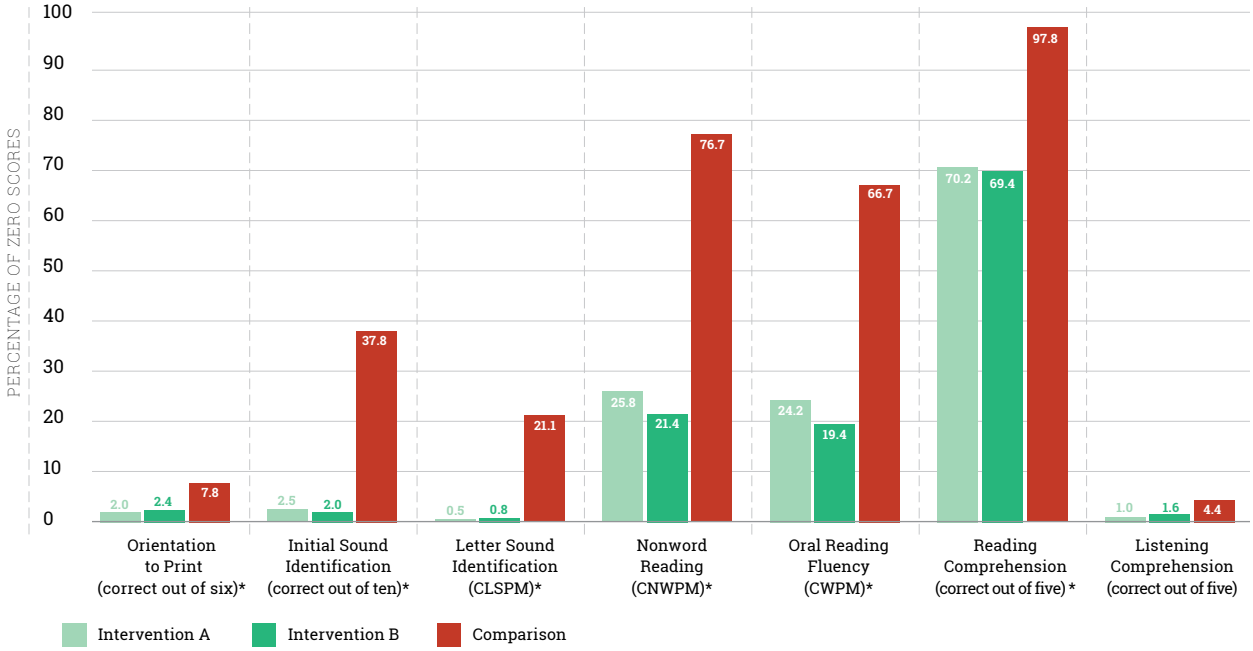


Figure 6: **Percentage of Students Receiving Zero Scores by EGRA Subtask and Group on the EGRA Instrument 2 at Endline (%)**[34]



33 An asterisk (*) indicates a significant difference in the proportion of zero scores among the groups at $p<0.05$. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

34 An asterisk (*) indicates a statistically significant difference in the proportion of zero scores among the groups at $p<0.05$. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

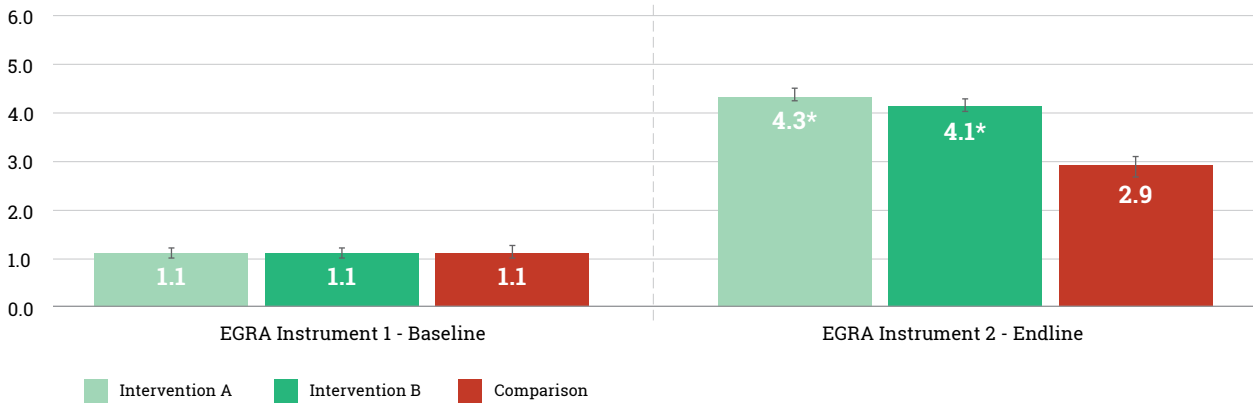
EGRA Results by Subtask

Orientation to Print[35]

The orientation to print subtask measures students' knowledge of how words are organized on a page, the direction of print, and how print materials are organized. In this subtask, students were given a text and asked a series of questions that measured their understanding of how words on a page were organized and read. Students indicated their response to the six questions asked by pointing to the correct part of the page or by indicating the correct direction of reading. This is an untimed subtask, and results are presented as the average number of questions answered correctly out of six.

Mean results for the orientation to print subtask are presented in Figure 7. Overall, the mean scores on the EGRA instrument 2 at endline were statistically significantly higher than the mean scores on the EGRA instrument 1 at baseline.

Figure 7: **Mean Results by Group at Baseline and Endline—Orientation to Print (Correct out of Six)[36]**



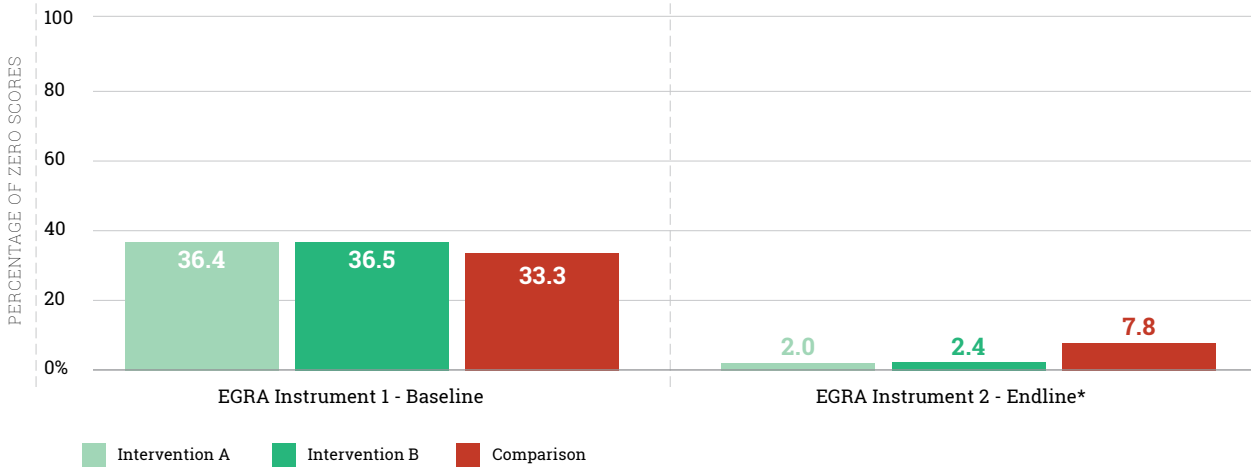
Results varied significantly across time points and groups. **At baseline, all three groups had comparable scores; at endline, students in both intervention groups had statistically significantly higher scores on the average orientation to print subtask than did students in the comparison group.** On average, students in interventions A and B outperformed students in the comparison group by 1.4 and 1.2 correct answers, respectively. There was no significant difference between the average scores for interventions A and B at endline.

Figure 8 shows the percentage of students who received zero scores at baseline and endline on the orientation to print subtask. At baseline, more than one-third of students in each group received zero scores; there was no significant difference in the proportion of zero scores among the groups. **At endline, there was a statistically significant difference in the proportion of zero scores among the groups: comparison group students had the highest proportion of zero scores (7.8 percent), followed by intervention B (2.4 percent) and intervention A (2.0 percent).**

35 The analysis was extended to review gains on the orientation to print subtask for the 372 students for the matched sample. Mean gain scores for orientation to print were as follows: on average, intervention A students increased by 3.0 correct answers from baseline to endline, followed by intervention B students with an increase of 2.8 correct answers, and then by comparison group students with an increase of 1.8 answers. Both intervention gain scores were found to be significantly higher than those of the comparison group, though there was no statistically significant difference found between intervention groups.

36 An asterisk (*) indicates the average score was significantly higher than the comparison group at $p<0.05$. Two asterisks (**) indicate that intervention A's average score was significantly higher than that of both the comparison group and intervention B at $p<0.05$. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

Figure 8: **Percentage of Students Receiving Zero Scores by Group at Baseline and Endline—Orientation to Print (%)**[37]

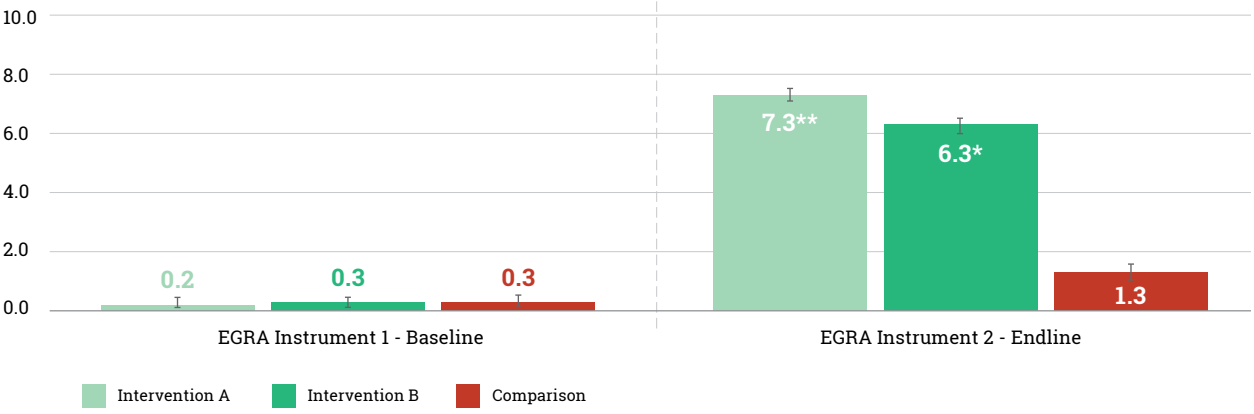


Initial Sound Identification[38]

The initial sound identification subtask measures students' ability to identify the initial sounds of words. The ability to identify isolated sounds within a word is a test of phonemic awareness and indicates that a student understands that words are made up of sounds—an understanding he or she can then use to associate sounds with letters, which is a building block of decoding. In this subtask, the assessor read ten words, and students were asked to verbally indicate the initial sound—or phoneme—of each word. This is an untimed subtask, and results are presented as the average number of questions answered correctly out of ten.

Mean results for the initial sound identification reading subtask are presented in Figure 9. Overall, the mean scores on the EGRA instrument 2 at endline were significantly higher than the mean scores on the EGRA instrument 1 at baseline.

Figure 9: **Mean Results by Group at Baseline and Endline—Initial Sound Identification (Correct out of Ten)**[39]



37 An asterisk (*) indicates a significant difference in the proportion of zero scores across groups at $p<0.05$. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

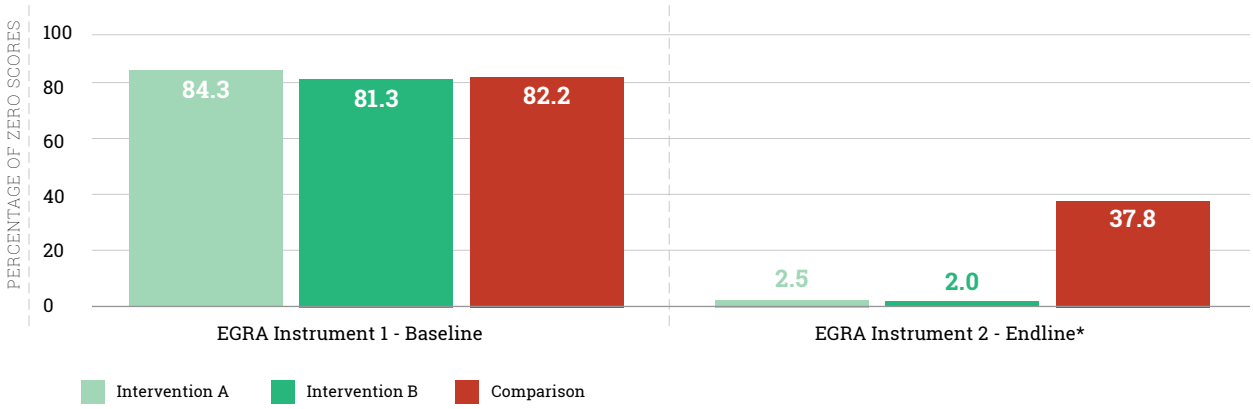
38 The analysis was extended to review gains on the initial sound identification subtask for the 372 students for the matched sample. Mean gain scores for initial sound identification were as follows: on average, intervention A students increased by 7.2 correct answers from baseline to endline, followed by intervention B students with an increase of 6.1 correct answers, and then by comparison students with an increase of 1.0 answers. Both intervention gain scores were found to be significantly higher than those of the comparison group. Moreover, intervention A gains were found to be significantly higher than intervention B gains.

39 An asterisk (*) indicates the average score was significantly higher than the comparison group at $p<0.05$. Two asterisks (**) indicate that intervention A's average score was significantly higher than that of both the comparison group and intervention B at $p<0.05$. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

Results on this subtask varied by time point and by group. On the EGRA instrument 1 at baseline, scores on the initial sound identification subtask were low across all three groups; there were no statistically significant differences between these groups at baseline. **On the EGRA instrument 2 at endline, both intervention groups averaged significantly higher initial sound identification scores than the comparison group; intervention A students had significantly higher average scores than intervention B.** In other words, students in intervention A outperformed students in intervention B by 1.0 correct initial sounds and students in the comparison group by 6.0 initial sounds. Furthermore, students in intervention B outperformed students in the comparison group by 5.0 correct initial sounds.

The percentage of students receiving zero scores at baseline and endline are presented in Figure 10. At baseline, more than 80 percent of students in the three groups received zero scores; there was no statistically significant difference in the proportion of zero scores across groups. At endline, there was a significant difference in the proportion of zero scores among the three groups. **Both intervention groups had comparatively low proportions of students receiving zero scores (2.5 percent and 2.0 percent for intervention A and intervention B, respectively).** By contrast, more than one-third of students in the comparison group (37.8 percent) received zero scores on the initial sound identification subtask.

Figure 10: **Percentage of Students Receiving Zero Scores by Group at Baseline and Endline—Initial Sound Identification (%)**[40]



Letter Sound Identification[41]

The letter sound identification subtask measures students' understanding of the alphabetic principle, which states that each letter of the alphabet corresponds to a specific sound. To demonstrate letter sound identification, students must identify the appropriate sounds for each letter. The ability to match letters with correct sounds is critical to reading fluency and comprehension. For this subtask, students were presented with a stimulus of 100 letters and asked to read aloud as many of the sounds as they could in one minute.[42] Results for this subtask are reported as a fluency rate of CLSPM.

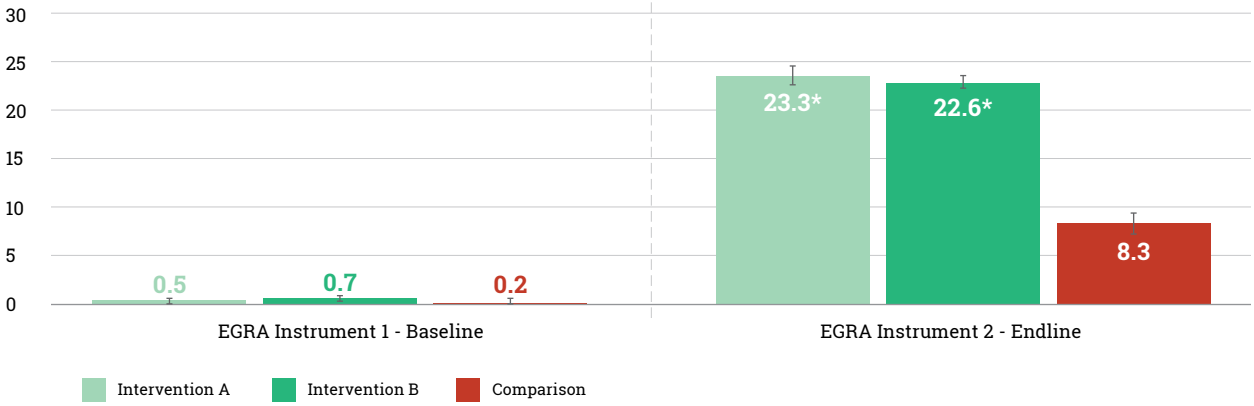
Mean results for the letter sound identification subtask are presented in Figure 11. Overall, the mean fluency rates of the three groups at endline were statistically significantly higher than the mean fluency rates of the three groups at baseline.

40 An asterisk (*) indicates a significant difference in the proportion of zero scores across groups at $p<0.05$. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

41 The analysis was extended to review gains on the letter sound identification subtask for the 372 students for the matched sample. Mean gain scores for letter sound identification were as follows: on average, intervention A students increased by 22.4 CLSPM from baseline to endline, followed by intervention B students with an increase of 21.0 CLSPM, and then by comparison group students with an increase of 8.6 CLSPM. Both intervention gain scores were found to be significantly higher than those of the comparison group, though there was no statistically significant difference found among intervention groups.

42 There is an auto stop rule in all the timed EGRA subtasks. In this case, the test was discontinued if a student was unable to correctly name any the first ten letters on the stimulus.

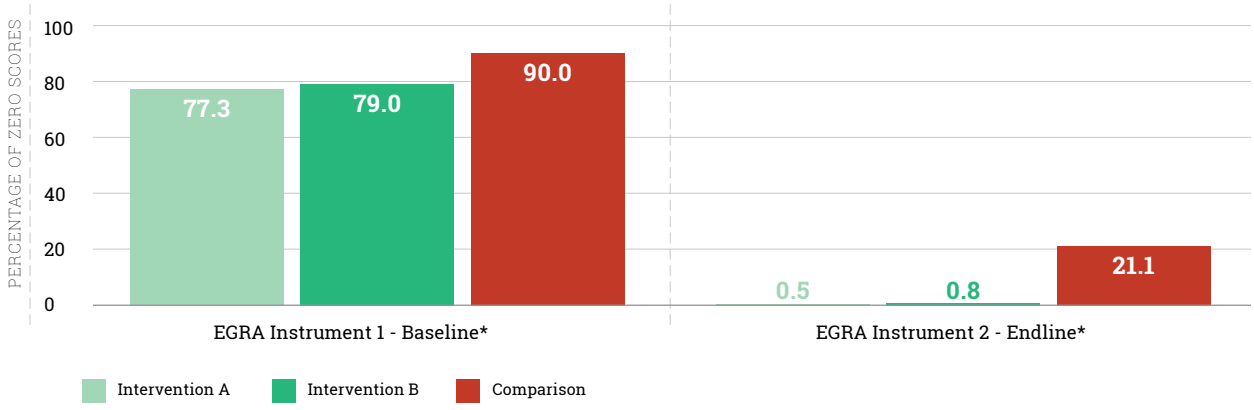
Figure 11: **Mean Results by Group at Baseline and Endline—Letter Sound Identification (CLSPM)**[43]



As with the previous subtasks, results on the letter sound identification subtask varied across time points and groups. On the EGRA instrument 1 at baseline, the three groups performed comparably: they all averaged less than 1.0 CLSPM. On the EGRA instrument 2 at endline, both intervention groups had statistically significantly higher average letter sound identification fluency rates than did the comparison group. **Students in interventions A and B outperformed students in the comparison group by 15.0 and 14.3 CLSPM, respectively.** No significant difference was detected between average fluency rates for interventions A and B.

Figure 12 shows the proportions of students receiving zero scores at baseline and endline on the letter sound identification subtask. At baseline, the proportion of zero scores among the three groups was statistically significantly different, although the majority of students in all three groups were unable to answer a single item correctly; 77.3 percent of intervention A students, 79.0 percent of intervention B students, and 90.0 percent of comparison group students had zero scores. **At endline, the proportion of students receiving zero scores at endline was statistically significantly different; 0.5 percent of intervention A students and 0.8 percent of intervention B students received zero scores at endline in contrast with 21.1 percent of students in the comparison group.**

Figure 12: **Percentage of Students Receiving Zero Scores by Group at Baseline and Endline—Letter Sound Identification (%)**[44]



43 An asterisk (*) indicates the average score was significantly higher than the comparison group at p<0.05. Two asterisks (**) indicate that intervention A's average score was significantly higher than that of both the comparison group and intervention B at p<0.05. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}$=540; $n_{\text{Intervention A}}$=198; $n_{\text{Intervention B}}$=252; $n_{\text{Comparison}}$=90.

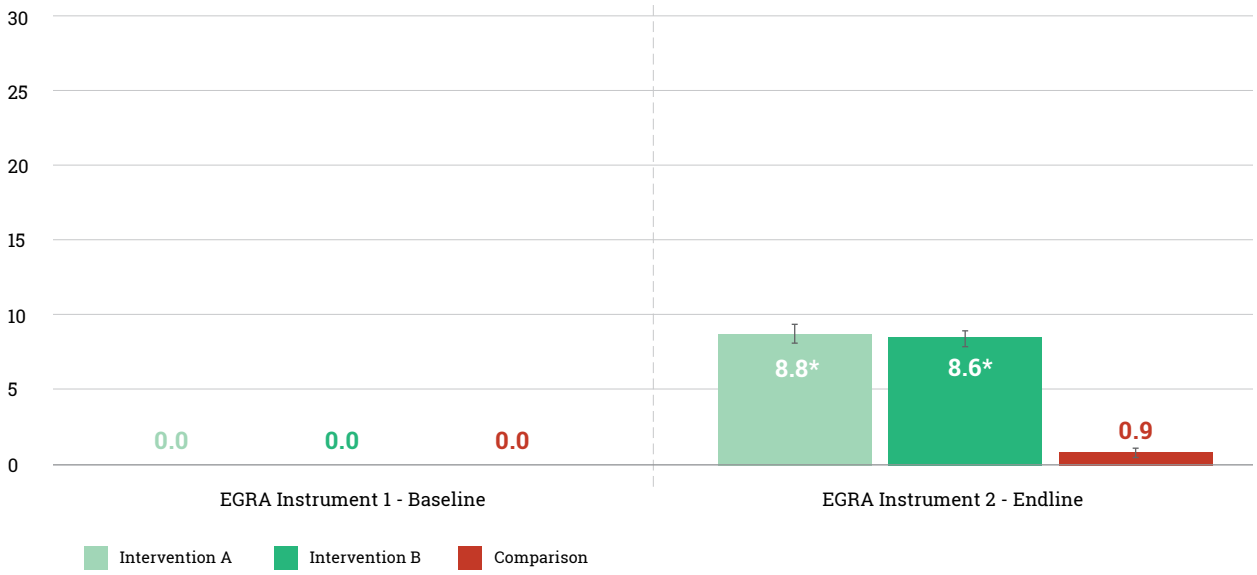
44 An asterisk (*) indicates a significant difference in the proportion of zero scores across groups at p<0.05. *N* sizes: $N_{\text{All students}}$=540; $n_{\text{Intervention A}}$=198; $n_{\text{Intervention B}}$=252; $n_{\text{Comparison}}$=90.

Nonword Reading[45]

The nonword reading subtask measures students' decoding ability by presenting them with words that they would not be able to recognize due to familiarity. Many students in the early grades learn to memorize or recognize a range of familiar words. Thus, to assess their decoding skills, students are presented with invented nonsense words, which requires them to sound out each letter and syllable to decode a word. During this timed subtask, the assessor presented each student with 50 nonwords and asked him or her to read as many as possible in one minute.[46] Results for this subtask are reported as a fluency rate of CNWPM.

Mean results for the nonword reading subtask are presented in Figure 13. Student performance on this subtask varied by group and time point. At baseline, the average score for each of the three groups was comparable, and low: less than 1.0 CNWPM. **At endline, both intervention groups had statistically significant higher average scores than did the comparison group, though no significant difference was detected among intervention groups.** On average, intervention A students read 8.8 CNWPM, while intervention B students read 8.6 CNWPM. Comparison group students read 0.9 CNWPM, on average.

Figure 13: **Mean Results by Group at Baseline and Endline—Nonword Reading (CNWPM)[47]**



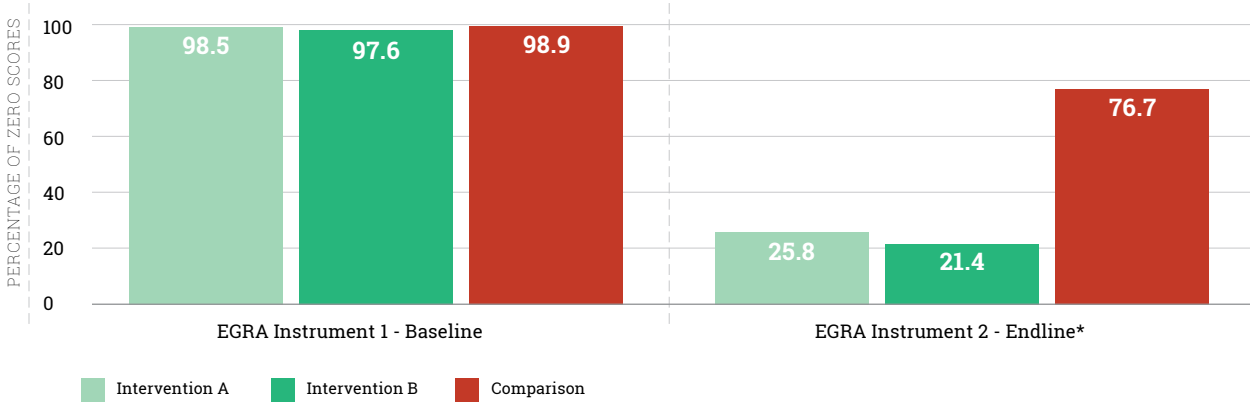
The percentage of students receiving zero scores at baseline and endline is presented in Figure 14. At baseline, over 97.0 percent of students in each group received zero scores, with no statistically significant differences among groups. At endline, students in intervention B had the lowest proportion of students with zero scores, followed by intervention A, and then the comparison group (21.4 percent, 25.8 percent, and 76.7 percent, respectively). The proportion of students receiving zero scores was statistically significantly varied by group at endline.

45 The analysis was extended to review gains on the nonword reading subtask for the 372 students for the matched sample. Mean gain scores for nonword reading were as follows: on average, intervention A students increased by 8.5 CNWPM from baseline to endline, followed by intervention B students with an increase of 8.3 CNWPM, and then by comparison students with an increase of 1.1 CNWPM. Both intervention gain scores were found to be significantly higher than those of the comparison group, though there was no statistically significant difference found among intervention groups.

46 After one minute, the student was asked to stop. The subtask was discontinued if a student was unable to correctly read any the first five nonwords.

47 An asterisk (*) indicates the average score was significantly higher than the comparison group at $p<0.05$. Two asterisks (**) indicate that intervention A's average score was significantly higher than that of both the comparison group and intervention B at $p<0.05$. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

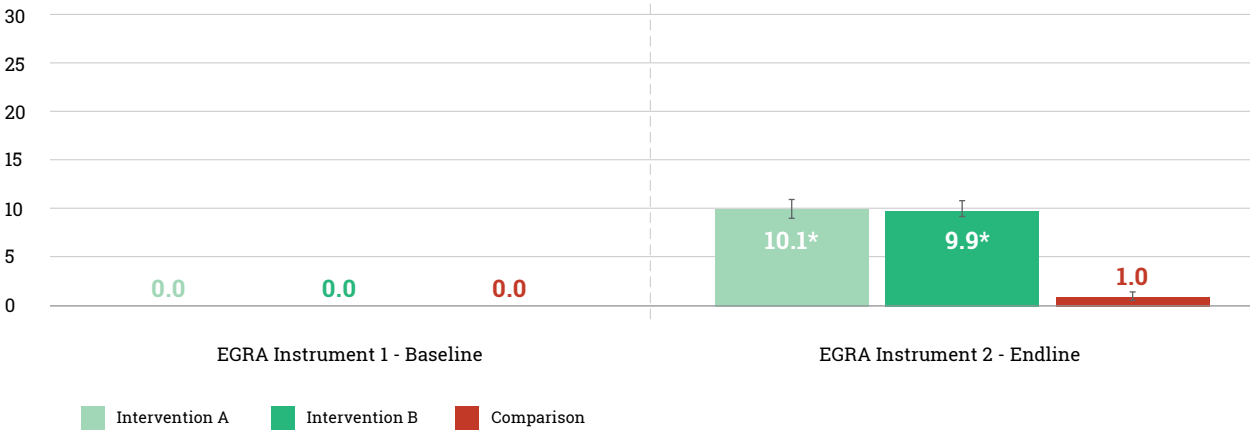
Figure 14: **Percentage of Students Receiving Zero Scores by Group at Baseline and Endline—Nonword Reading (%)**[48]



Oral Reading Fluency

The ORF subtask measures students' overall reading competence. It is the culmination of translating letters into sounds, merging sounds to become words, linking words to become sentences, relating the text to meaning, and making inferences to fill in missing information. A student's ORF score is dependent on the skills assessed in previous subtasks since students need to have some mastery of letter sounds and decoding to read fluently. Students had the opportunity to read up to 50 words in the ORF passage on the EGRA instrument 1 at baseline and up to 52 words on the EGRA instrument 2 at endline. Results for this subtask are reported as a fluency rate of CWPM. Mean results are presented in Figure 15.

Figure 15: **Mean Results by Group at Baseline and Endline—Oral Reading Fluency (CWPM)**[49]



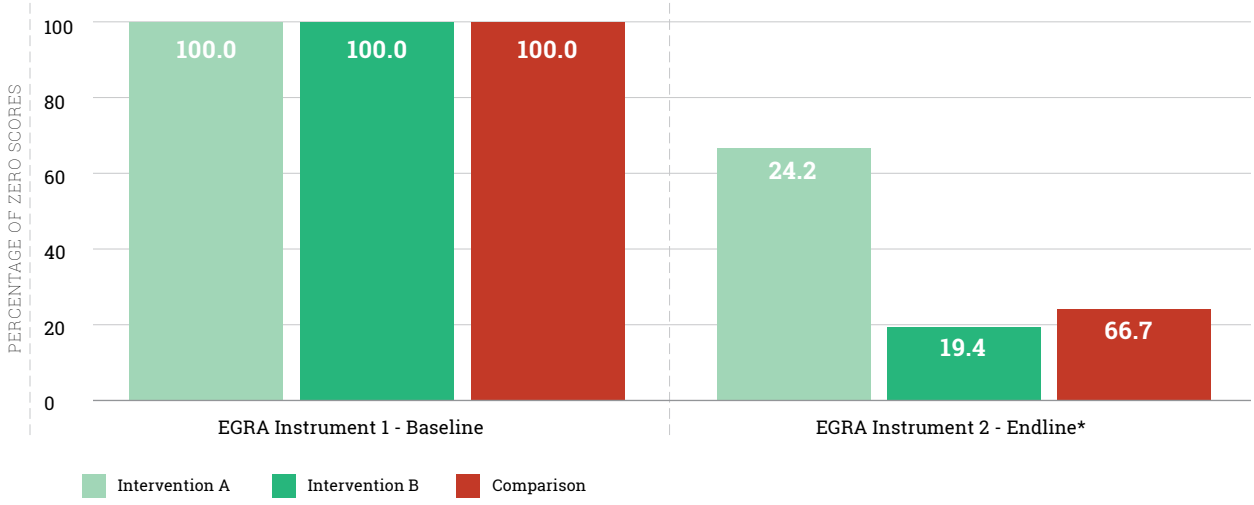
48 An asterisk (*) indicates a significant difference in the proportion of zero scores across groups at $p<0.05$. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

49 An asterisk (*) indicates the average score was significantly higher than the comparison group at $p<0.05$. Two asterisks (**) indicate that intervention A's average score was significantly higher than that of both the comparison group and intervention B at $p<0.05$. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

At baseline, no student in any of the three groups could read a single word on the ORF subtask. **At endline, the difference in the average results of students across the three groups was statistically significant: students in both intervention groups had significantly higher average ORF scores than students in the comparison group.** Intervention A students read, on average, 10.1 CWPM, while intervention B students read 9.9 CWPM. By contract, students in the endline comparison group read, on average, 1.0 CWPM. No statistically significant difference was detected among average fluencies across intervention groups.

The percentages of students receiving zero scores on the ORF subtask are presented in Figure 16. At baseline, all students received zero scores. **At endline, there was a statistically significant difference in the proportion of zero scores among groups. The comparison group had the highest proportion of students with zero scores—66.7 percent.** This was followed by 24.2 percent of students in intervention A and 19.4 percent of students in intervention B.

Figure 16: **Percentage of Students Receiving Zero Scores by Group at Baseline and Endline—Oral Reading Fluency (%)**[50]



Reading Comprehension

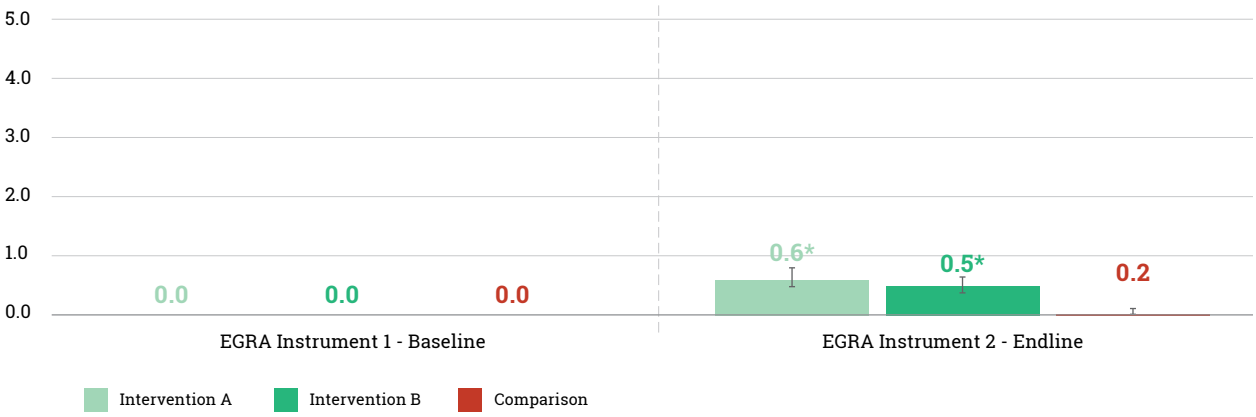
Comprehension is the purpose of reading. Once a child learns the sound-letter relationship (alphabetic principle) and becomes able to decode and read with automaticity, he or she becomes increasingly able to understand the meaning of a text. This subtask assesses that ability.

For the reading comprehension subtask, the assessor removed the passage used in the ORF subtask and then asked each student up to five comprehension questions based on what he or she had read. The number of questions asked depended on how many words each student read on the ORF subtask. For instance, if a student read just the first ten words, he or she would be asked only the first comprehension question. Similarly, if a student read all words on the ORF subtask, he or she would be asked all five questions. Students who received zero scores on the ORF subtask also received zero scores on the reading comprehension subtask because no questions were presented to them. Additionally, any student who could not correctly answer a single reading comprehension question received a zero score on this subtask.

50 An asterisk (*) indicates a significant difference in the proportion of zero scores across groups at $p<0.05$. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

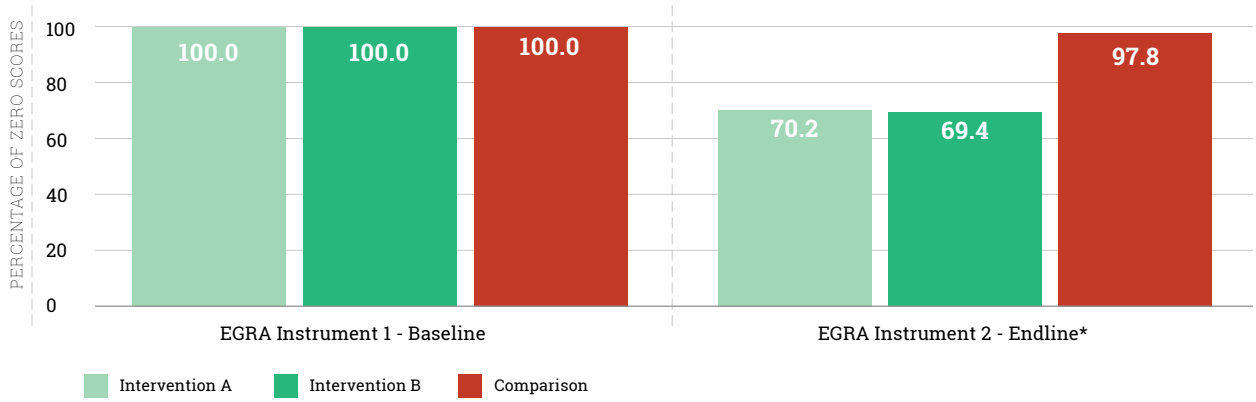
Mean scores are presented in Figure 17. On the EGRA instrument 1 at baseline, no student was presented with a reading comprehension question since all students received zero scores on the ORF subtask. **On the EGRA instrument 2 at endline, the difference in the average results of students across the three groups was statistically significant: both intervention groups had significantly higher average reading comprehension scores than did the comparison group, though no significant difference was detected among intervention groups.** On average, students in interventions A and B correctly answered 0.6 and 0.5 questions, respectively. Comparison group students correctly answered fewer than 0.1 questions.

Figure 17: **Mean Results by Group at Baseline and Endline—Reading Comprehension (Correct out of Five)**[51]



The proportions of students receiving zero scores at baseline and endline are presented in Figure 18. At baseline, all students across all groups received zero scores on the reading comprehension subtask. **At endline, there was a statistically significant difference in the proportion of zero scores of the three groups.** While the majority of students in all three groups still received zero scores on this subtask, smaller proportions of zero scores were observed in the intervention groups: 70.2 percent of intervention A students and 69.4 percent of intervention B students, respectively, received zero scores. These proportions were lower than the proportion of students receiving zero scores in the comparison group—97.8 percent.

Figure 18: **Percentage of Students Receiving Zero Scores by Group at Baseline and Endline—Reading Comprehension (%)**[52]



51 An asterisk (*) indicates the average score was significantly higher than the comparison group at $p<0.05$. Two asterisks (**) indicate that intervention A's average score was significantly higher than that of both the comparison group and intervention B at $p<0.05$. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

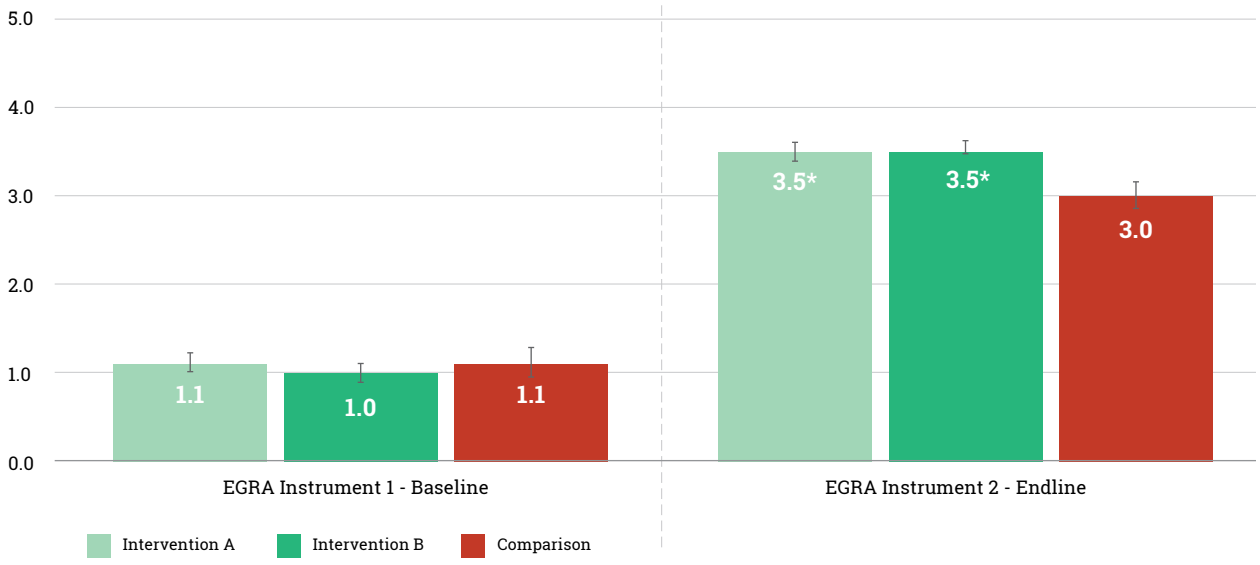
52 An asterisk (*) indicates a significant difference in the proportion of zero scores across groups at $p<0.05$. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

Listening Comprehension

The untimed listening comprehension subtask measures students' ability to comprehend the meaning of a story read to them orally. Students do not need to know how to read in order to answer listening comprehension questions. As a result, this subtask is an important measure of students' pre-reading abilities because it helps detect obstacles to learning to read, such as limited language proficiency, auditory problems, attention deficit, and other difficulties. In this subtask, the assessor reads a short passage to the student and asks them to answer comprehension questions based on what they heard. Results for this subtask are presented as the number of questions correctly answered out of five.

Mean scores for this subtask are presented in Figure 19. On the EGRA instrument 1 at baseline, the average score for each of the three groups was about one correct question out of five; there was no statistically significant difference detected among the three groups. **At endline, students in both intervention groups averaged significantly higher reading comprehension scores than did students in the comparison group.** Average endline listening comprehension scores were the same for interventions A and B students—3.5 correct questions—while students in the comparison group correctly answered 3.0 questions. No significant difference was detected among intervention groups' scores at endline.

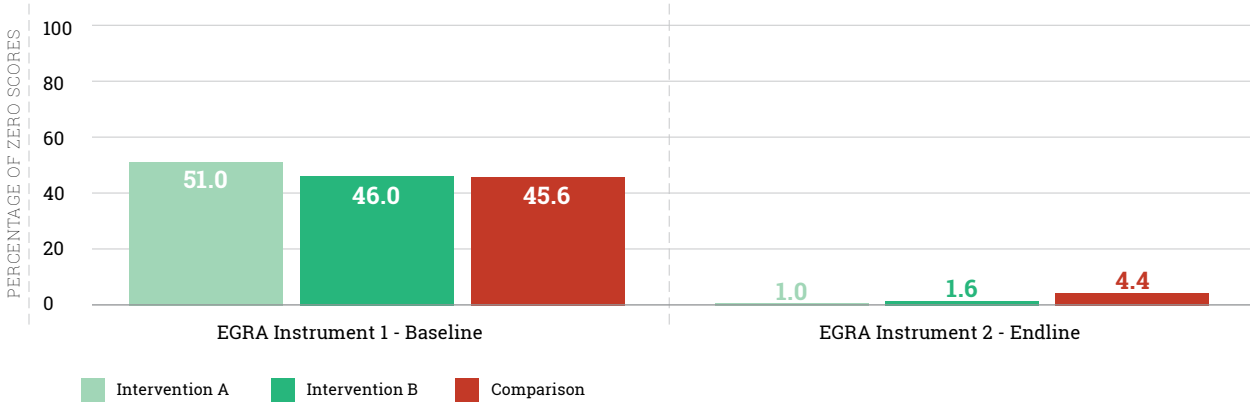
Figure 19: **Mean Results by Group at Baseline and Endline—Listening Comprehension (Correct out of Five)**[53]



The proportion of students receiving zero scores at baseline and endline is presented in Figure 20. At baseline, approximately half of all students in the three groups received zero scores on this subtask; no statistically significant differences were observed among groups. At endline, the proportions of students receiving zero scores were close to zero, again with no significant difference detected among the three groups.

53 An asterisk (*) indicates the average score was significantly higher than the comparison group at $p<0.05$. Two asterisks (**) indicate that intervention A's average score was significantly higher than that of both the comparison group and intervention B at $p<0.05$. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: $N_{\text{All students}}=540$; $n_{\text{Intervention A}}=198$; $n_{\text{Intervention B}}=252$; $n_{\text{Comparison}}=90$.

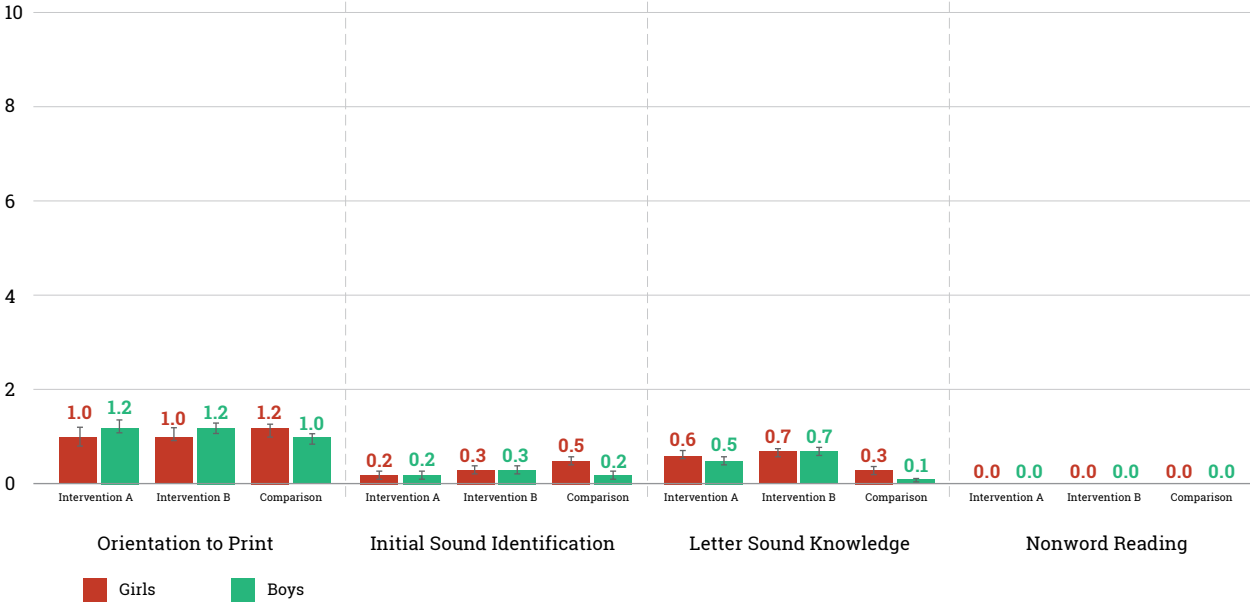
Figure 20: **Percentage of Students Receiving Zero Scores by Group at Baseline and Endline—Listening Comprehension (%)**[54]



EGRA Results by Subgroup

EGRA results were also analyzed by students' gender across groups to understand if the OCLR project impacted girls and boys differently. Specifically, this section responds to the supplemental question *How did the project influence certain subsets of the student population more than others based on identifiable contextual factors?* In the baseline group, there were 245 girls and 295 boys in the sample; in the endline group, there were 254 girls and 286 boys (see Annex Table H.1). Mean scores on the EGRA instrument 1 at baseline and the EGRA instrument 2 at endline are presented in Figures 21 and 22 for the orientation to print, initial sound identification, letter sound identification, and nonword reading subtasks by gender and intervention group.

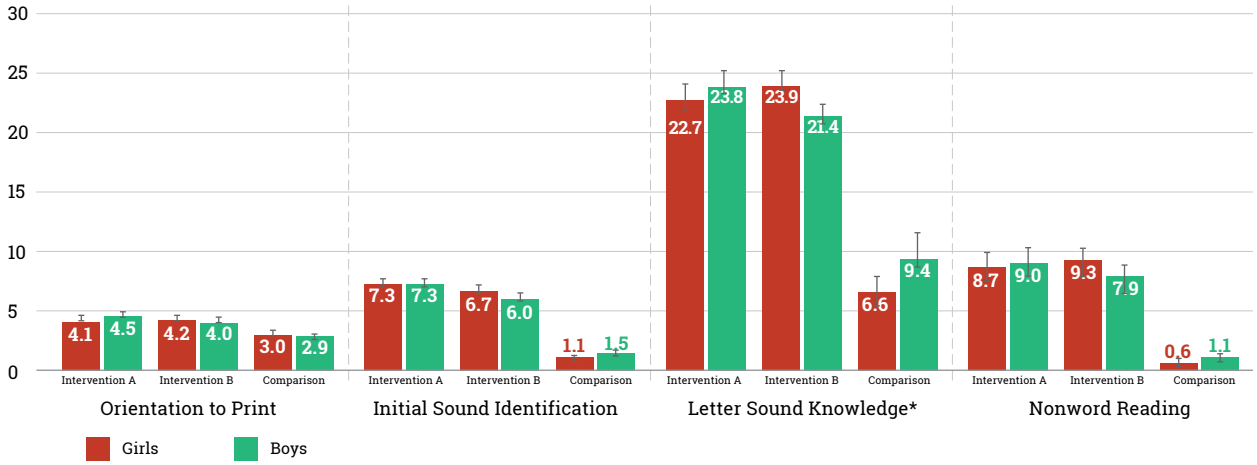
Figure 21: **Mean Results by Gender and Group on the EGRA Instrument 1 at Baseline—Orientation to Print, Initial Sound Identification, Letter Sound Identification, and Nonword Reading**[55]



54 An asterisk (*) indicates a significant difference in the proportion of zero scores across groups at p<0.05. *N* sizes: $N_{\text{All students}}$=540; $n_{\text{Intervention A}}$=198; $n_{\text{Intervention B}}$=252; $n_{\text{Comparison}}$=90.

55 An asterisk (*) indicates the average score was significantly different across groups at p<0.05. Two asterisks (**) indicate that the average score of intervention A was significantly higher than that of both the comparison group and intervention B at p<0.05. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: Intervention A: Girls *n*=91, Boys *n*=107; Intervention B: Girls *n*=113, Boys *n*=139.

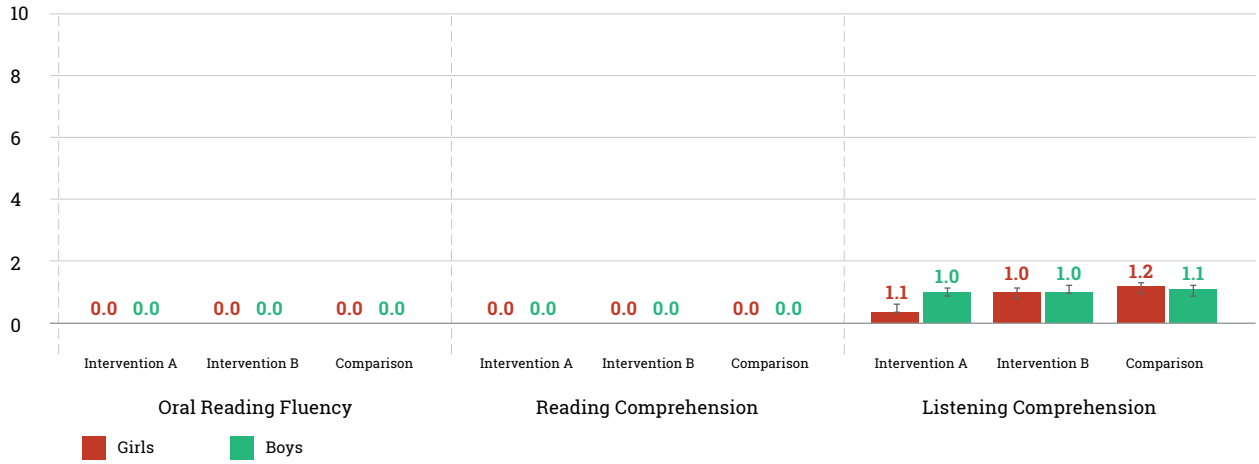
Figure 22: **Mean Results by Gender and Group on EGRA Instrument 2 at Endline—Orientation to Print, Initial Sound Identification, Letter Sound Identification, and Nonword Reading**[56]



No statistically significant differences were found between genders on the orientation to print, initial sound identification, and nonword reading subtasks across time points and group. However, on the letter sound identification subtask, results varied across time points, group, and gender. On the EGRA instrument 1 at baseline, the average fluency rate across gender and group was less than 1.0 CLSPM; there was no statistically significant difference by gender and group. On the EGRA instrument 2 at endline, average letter sound fluency was significantly different between boys and girls in intervention B where, on average, girls outperformed boys (23.9 vs. 21.4 CLSPM). There was no significant difference in letter sound fluency found between girls and boys in intervention A or the comparison group.

Mean results on the EGRA instrument 1 at baseline and the EGRA instrument 2 at endline are presented in Figures 23 and 24 for the ORF, reading comprehension, and listening comprehension subtasks by gender and group. On the EGRA instrument 1 at baseline and the EGRA instrument 2 at endline, there were no statistically significant differences in performance by gender and group on the ORF, reading comprehension, or listening comprehension subtask. In other words, boys and girls in each group and time point performed comparably.

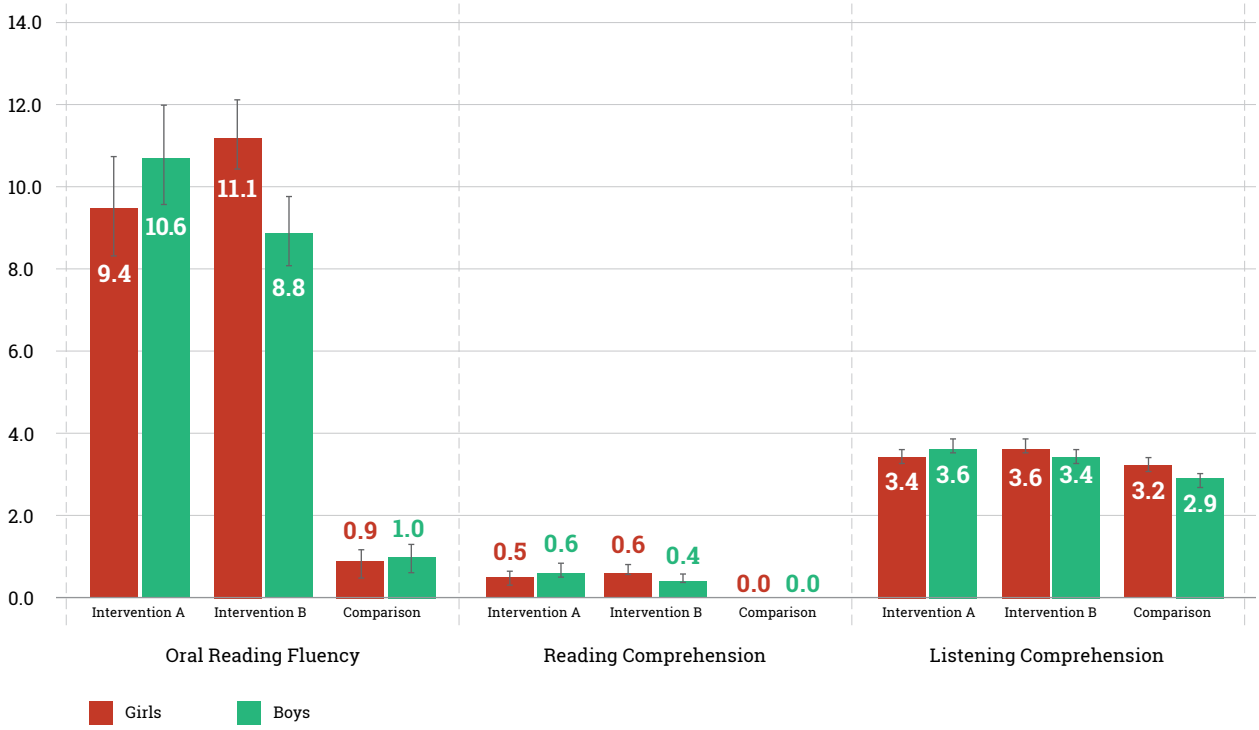
Figure 23: **Mean Results by Gender and Group on the EGRA Instrument 1 at Baseline—ORF, Reading Comprehension, and Listening Comprehension**[57]



56 An asterisk (*) indicates the average score was significantly different across groups and gender at p<0.05. Note: Mean score estimates are noted above each bar in the graph. The CI is indicated by the lines at the top of each bar. CI indicates a range of values that is likely to encompass the true value. *N* sizes: Intervention A: Girls *n*=95, Boys *n*=103; Intervention B: Girls *n*=123, Boys *n*=129.

57 *N* sizes: Intervention A: Girls *n*=91, Boys *n*=107; Intervention B: Girls *n*=113, Boys *n*=139.

Figure 24: **Mean Results by Grade and Group on the EGRA Instrument 2 at Endline— ORF, Reading Comprehension, and Listening Comprehension**[58]



VIII. Additional Results

To better understand possible variations in student learning experiences and context within the OCLR project, questions from the student questionnaire administered with the endline EGRA were compiled into five composites, or groups of questions related to each other. Each composite consists of a series of items related to a specific theme that may have affected students' early grade reading skill acquisition. Composites were then assigned a maximum score equal to the total number of items in the composite.[59]

The composites for the OCLR project include

1. Language exposure
2. Socioeconomic status
3. Family reading support
4. Teacher reading support
5. Disposition to reading

Descriptive statistics for the student questionnaire composites are presented in Tables 6 and 7 (see Annex D for full composite questions, response options, and frequencies).

58 *N* sizes: Intervention A: Girls *n*=95, Boys *n*=103; Intervention B: Girls *n*=123, Boys *n*=129.

59 Nonresponses were given a '0.'

Table 7: **Average Student Questionnaire Composite Scores by Intervention Group**

	Intervention A	Intervention B	Comparison	Total
Language exposure	5.9	5.9	5.7	**5.8**
Socioeconomic status	8.2	8.1	8	**8.1**
Family reading support	3.1	3.1	2.7	**3**
Teacher reading support	2.6	2.7	2.4	**2.6**
Disposition to reading	2.6	2.6	2.5	**2.6**

Results indicate that, across intervention groups for all composites, students had similar learning experiences and context. A correlational analysis was conducted to determine what relationships, if any, existed between mean composite scores and EGRA gain scores for the orientation to print, initial sound identification, letter sound identification, and nonword reading subtasks. The analysis, which used the matched sample, revealed a significant relationship between language exposure and students' letter sound identification gain scores, with the relationship strongest amongst comparison group students (.351) and relatively weaker with the intervention group students (.158 and .171 in interventions A and B, respectively) (see Annex F). **This suggests that students who had greater exposure to Bamanankan at home and at school tended to have higher gains on the letter sound identification, regardless of the intervention.**

IX. Scalability

Stakeholders are increasingly interested in assessing the scalability of interventions, in addition to their results or impacts. To scale up a project means to expand, replicate, adapt, and sustain a successful project in a new geographic area and to reach more beneficiaries over time.[60] ACR GCD grantees have implemented small-scale pilot projects, and an important consideration after each project is the feasibility of replicating or expanding the technology-based innovation and project models to a different or larger population or area.

To inform this decision, STS conducted a scalability assessment guided by the following research question: *Are this project and technology suitable for scaling?* STS used an indirect approach that relies on qualitative descriptions of project performance around seven parameters of sustainability:

- Credibility
- Observability
- Relevance
- Relative advantage
- Ease of transfer and adoption
- Testability
- Sustainability of funding

60 Cooley, L., & Linn, J. F. (2014). *Taking Innovations to Scale: Methods, Applications and Lessons*. Results for Development Institute. Washington, D.C. Retrieved from https://www.usaid.gov/sites/default/files/documents/1865/v5web_R4D_MSI-BrookingsSynthPaper0914-3.pdf

The seven parameters were adapted from the USAID-funded Scalability Assessment Tool developed by Management Systems International.[61] The tool includes seven sections and 28 questions. STS used data from EOP interviews, EGRA results, literature reviews, and project M&E to assess scalability parameters. These results are meant to inform local program staff, stakeholders, and donors about key considerations before scaling the OCLR project's model and technologies to a larger or different beneficiary population.

Credibility

An intervention or innovation must be credible to be supported and taken to scale, through either replication or expansion. This aspect of scalability assesses whether various stakeholders—including potential adopters, funders, implementers, and beneficiaries—believe that the model has a strong evidence base that may include existing empirical research or anecdotal information.

Key Considerations

1. What evidence was used to develop the intervention?
2. What evaluations have been conducted on the intervention?
3. In what social contexts does the intervention work?
4. What individuals and institutions support the intervention?

The OCLR project combined teacher training in the balanced literacy approach with access to training materials on tablets and mobile phones using the Stepping Stone app. Both components had previously been implemented in Mali. The balanced literacy approach was used in Mali during the USAID PHARE project, during which EDC and the MEN developed guidance on the seven teaching strategies to help Malian teachers improve their literacy instruction. The five-year program worked nationally, reaching more than 40,000 classrooms and 600,000 students. Furthermore, EDC had already tested and implemented the Stepping Stone app in Mali through the USAID PAJE-Nièta project. The project distributed more than 4,000 mobile phones—pre-loaded with the Stepping Stone app and containing numerous literacy and numeracy activities—to out-of-school youth.[62] The Stepping Stone app has since been tested on a variety of projects and is used by 12,000 children in Mali, according to EDC staff members who have been involved in the development of the Stepping Stone app from its beginning. Evaluations of the PAJE-Nièta project suggested a potentially high impact of the Stepping Stone app: 77.0 percent of youth who were part of the project showed improved reading skills on EDC's Out-of-School Literacy Assessment.[63]

Although no previous evaluations were conducted on the specifics of the OCLR project model—specifically, its pairing of teacher training in the balanced literacy approach, coupled with access to supplementary training materials on tablets—USAID PHARE collected and analyzed M&E data on the impact of Malian students' reading competencies after receiving instruction using the balanced literacy approach.[64] In those studies, there was no statistically significant difference between baseline and endline results on several of the early grade reading skills, such as reading comprehension or writing of familiar words. However, students did show improvement on essential pre-reading skills, such as recognizing graphemes and phonics. The OCLR project attempts to address

61 Cooley, L., & Linn, J. F. (2014). *Taking Innovations to Scale: Methods, Applications and Lessons*. Results for Development Institute. Washington, D.C. Retrieved from https://www.usaid.gov/sites/default/files/documents/1865/v5web_R4D_MSI-BrookingsSynthPaper0914-3.pdf

62 Education Development Center, Inc. (n.d.). *History of Stepping Stone*. Retrieved from http://sstone.edc.org/en/what-is-stepping-stone-4/history-of-stepping-stone/

63 Education Development Center, Inc. (n.d.). *Positive Youth Development in Mali: Better Educated, Economically Productive and Civically Engaged Rural Youth*. Retrieved from http://idd.edc.org/sites/idd.edc.org/files/Mali%20Out%20of%20School%20Project%20Summary.pdf

64 Education Development Center, Inc. (2015, March). *Balanced Literacy Transforms Classrooms: Evidence from Mali*. Retrieved from http://idd.edc.org/sites/idd.edc.org/files/RRN%20Mali%20Balanced%20Literacy%20Brief.pdf

some of the challenges identified with USAID PHARE, including lack of adequate follow-up after teacher training, insufficient support to teachers, and lack of supplementary reading materials.

The OCLR project's intervention could be implemented in different social contexts given specific conditions. RARE depended on strong buy-in from the MEN and a tested, pre-developed literacy teaching strategy to implement the OCLR project. The teacher training component of the OCLR project could be easily adapted to different parts of Mali, though is unclear if, in different social contexts, the local MEN representatives would allow for the balanced literacy approach to replace preexisting literacy teaching approaches. The challenges addressed by access to supplementary training materials on tablets using the Stepping Stone app, however, are highly relevant to a variety of social contexts. Throughout West and Sub-Saharan Africa, national ministries of education lack the resources to provide adequate follow-up and support to teachers after training. Having access to information and communications technologies (ICT), like the Stepping Stone app, could alleviate this challenge and would be a great help in reinforcing teacher training content. Furthermore, the Stepping Stone app can be utilized in internet-deprived areas due to its offline feature, meaning this component of the project could be replicated in urban or rural areas.

There is strong support for the OCLR project and balanced literacy approach within Mali. Various stakeholders in Mali, including the MEN and EDC, were engaged in developing the balanced literacy approach through USAID PHARE and supported its incorporation into the OCLR project. Local representatives were also highly involved in the OCLR project; they closely engaged with start-up activities, helped implementation through the CAPs, reviewed training materials, assisted in developing the EGRAs used to evaluate the project, and participated in baseline and endline evaluations. Furthermore, USAID/Mali expressed support for the model and its scale-up through the USAID SIRA project.

Credibility Conclusion

Within Mali, there is high credibility for the OCLR project model given the government's investment in the balanced literacy approach as a key teaching strategy in its primary schools. The OCLR project used previously tested components and combined them into an intervention model that addressed challenges in earlier iterations, such as the USAID PHARE project. There is also high engagement from government, NGOs, and USAID/Mali stakeholders. Although the use of ICT to reinforce learnings from teacher training is a highly adaptable and relevant strategy across a variety of social contexts, it is unclear if the implementation of the balanced literacy approach would be replicable outside of Mali.

Observability

For an intervention or innovation to be scaled, it should have observable results that show efficacy or impact. Observability of results is key to providing nontechnical audiences with proof that an intervention or innovation achieved its intended outcomes and therefore will have positive impacts on beneficiaries.

Key Considerations

1. Are the results visual and observable?
2. What is the relationship (if any) between results and the intervention?
3. Is there any emotional appeal associated with the evidence?

Results from the OCLR project are positive. The project targeted pre-reading and foundational skills, and on the subtasks that assessed those, intervention group students had statistically significantly higher scores at endline than did their peers in the comparison group. Specifically, on the initial sound identification subtask, intervention A students correctly identified 7.3 initial sounds, intervention B students correctly identified 6.3 initial sounds, and comparison group students correctly identified only 1.3 initial sounds. On the letter sound identification subtask, intervention A students identified 23.3 CLSPM, intervention B students identified 22.6 CLSPM, and comparison group students identified 8.3 CLSPM. Intervention group students also had higher decoding skills at endline than their comparison group peers: on the nonword reading subtask, intervention A and B students identified 8.8 CNWPM and 8.6 CNWPM, respectively, in contrast to only 0.9 by the comparison group students. Students in both intervention groups also had statistically significantly higher scores on the ORF, reading comprehension, and listening comprehension subtask than did comparison group students. However, only on the initial sound identification subtask did students in intervention A—those whose teachers had access to tablets with the Stepping Stone app—have statistically significantly higher scores than students in intervention B. As a result, it is not clear if the technology component of the project—the tablet and its content—added value.

Given the inclusion of a randomly selected comparison group, it is reasonable to assume a relationship between the EGRA results and the OCLR project intervention. However, because the OCLR project targeted teacher classroom instruction, the observability of the results would be strengthened by an analysis that incorporated standardized classroom observation data and EGRA results. Although there were increases in early grade reading skills among students whose teachers were trained by the OCLR project, without detailed information on how much each teacher used the approach in their classroom, the relationship between the intervention and EGRA results is inferable but not empirically proven.

Project managers and pedagogical counselors believed that intervention A teachers, who had access to the tablets, were more comfortable with balanced literacy techniques than were their peers who received training but did not have access to the tablets. Project managers also stated, anecdotally, that teachers with tablets were more likely to employ specific components of the balanced literacy approach than others. However, these findings are not corroborated by the EGRA results, and without more rigorous classroom observation data, it is not clear if intervention A teachers adopted the balanced literacy approach in different ways than did intervention B teachers. Project management did express that the number of videos available to teachers on the Stepping Stone app was too few, and they believed that more videos of improved quality would have been beneficial.

Observability Conclusion

Students whose teachers were trained by the OCLR project had stronger early grade reading skills at endline than did their peers whose teachers did not receive training, particularly in the pre-reading and foundational skills that the project intervention targeted. However, EGRA results suggest that there was not a difference in performance between students whose teachers had access to the project's technological components and those whose teachers did not. Before the technology component of the OCLR project is scaled, further research should investigate teacher classroom behaviors and the relationship between EGRA gains, as well as explore offering more training content on the Stepping Stone app.

Relevance

An intervention must be relevant to the context in which it is being implemented to be scalable. It should effectively address a problem that is recognizable and considered important by stakeholders.

Key Considerations

1. What is the level of significance of the problem that the intervention is trying to address?
2. Does the intervention address a priority on the policy agenda for potential adopters?
3. Does the intervention address a need felt by the potential beneficiaries?

In Mali, reading assessments have indicated that at least 70 percent of primary school students are unable to read at grade level.[65] Research suggests that low learning levels are impeding economic growth, as a 10.0 percent increase in the share of students reaching basic literacy would translate into a 0.3 percentage-point greater annual growth rate for the country.[66] Further, results from a 2009 EGRA in Mali showed that a majority of students were unable to correctly read a single word at the end of Grade 2. According to statistics, only 38.8 percent of Malian females and 56.0 percent of Malian males aged 15 to 24 are literate; a 2009 study showed that only 17.9 percent of Grade 2 students in Mali have books at home.[67]

The MEN faces challenges in ensuring effective in-service training to teachers at a reasonable cost that contribute to the low levels of literacy in Mali. According to RARE, most teachers do not receive sufficient in-service training to ensure they have mastered pedagogical innovations, and training activities are often implemented gradually, making periodic refresher training a necessity.[68] Given financial limitations and the challenges observed during the USAID PHARE rollout of the balanced literacy approach, the problems addressed by the OCLR project is significant.

The OCLR project addresses a priority of the policy agenda in Mali, that of promoting basic quality education for every Malian through the GoM's Decennial Program for the Development of Education. To reach such an objective, the Malian system prioritizes the development of reading and writing skills.[69] EOP interviews conducted with five stakeholders from the MEN confirmed that improving early grade literacy, especially through teacher training programs, was indeed a policy priority for the GoM. On a national scale, Mali has opted for a bilingual curriculum approach to education in the early grades—teaching students to read in mother tongue languages in Grade 1 with an eventual transition to French-language instruction in Grades 2 and 3. The OCLR project worked with Grade 1 and 2 teachers to help promote literacy in Bamanankan in early grades, thus supporting Mali's policy of mother tongue, bilingual instruction.

The OCLR project's components also address policy priorities for USAID/Mali. USAID's global education strategy, released in 2011, emphasizes the need for improved literacy teaching techniques and instructional approaches, and promotes initial teacher preparation and professional development.[70] In EOP interviews, USAID/Mali stakeholders reiterated the relevance of the OCLR project's components to their priorities within Mali; they also repeated that they have encouraged this intervention model by incorporating components into the USAID SIRA project. Mother tongue language instruction is also emphasized by the United Nations Educational, Scientific, and Cultural Organization (UNESCO) through its strategy for languages and multilingualism.[71]

65 Gillies, J., & Quijada, J.J. (2008). *Opportunity to learn: A high impact for improving educational outcomes in developing countries*. Academy for Educational Development. Washington, D.C. Retrieved from https://www.fhi360.org/sites/default/files/media/documents/Opportunity%20to%20Learn%20-%20English.pdf

66 Hanushek, E., & Woessmann, L. (2009). *Do better schools lead to more growth?: Cognitive skills, economic outcomes, and causation*, NBER Working Paper 14633. Cambridge, MA: National Bureau of Economic Research.

67 Œuvre Malienne d'Aide à l'Enfance du Sahel. (2014). *Your Child, Reading, and You: Technical Proposal*. (Unpublished proposal for funding)

68 Réseau d'Acteurs pour le Renouveau de l'Education. (2014). *Using Mobile Technology to Improve the Teaching-Learning of Reading-Writing in Bilingual Curriculum Schools of the Education Region of Bougouni, Mali*. (Unpublished proposal for funding)

69 United Nations Children's Fund. (2015, February). *Mali. Country programme document 2015-2019*. Retrieved from https://www.unicef.org/about/execboard/files/2015-PL2-Mali_CPD-Final_approved-EN.pdf

70 United States Agency for International Development. (2012, April). *2011 USAID Education Strategy Reference Materials*. Retrieved from http://pdf.usaid.gov/pdf_docs/pdact680.pdf

71 United Nations Educational, Scientific, and Cultural Organization. (n.d.). *Languages in Education Strategy*. Retrieved from http://www.unesco.org/new/en/education/themes/strengthening-education-systems/languages-in-education/strategy/

The project worked to address multiple challenges in the Malian primary education context, as articulated by beneficiaries. Teachers, pedagogical counselors, and school directors interviewed at EOP confirmed that the project worked to close vital gaps in Malian primary education and emphasized that these gaps were highly relevant to their communities.

Some of the key challenges expressed by these beneficiaries were not, however, specifically targeted by the OCLR project. For example, in interviews, community members noted that teacher absenteeism and teacher strikes are a major impediment in the primary education system. Nevertheless, in EOP interviews, the OCLR project management noted that combining interventions at the school level—under the OCLR project—with community-level interventions—implemented by OMAES—would be more effective for improving students' learning outcomes and addressing a range of relevant literacy challenges for primary school students.

Relevance Conclusion

The relevance of the OCLR project is strong. Primary level literacy instruction and teacher training are a key focus of the MEN. The OCLR project model provides solutions that address the challenges of low early grade literacy skills and resource scarcities for teacher training follow-up and reinforcement. Although the project's objectives are highly relevant to the GoM, USAID/Mali, and local community education policy priorities, it may not fully address the commonly cited issues of teacher absenteeism and teacher strikes in Malian primary schools. To fully address this may require collaboration with the OMAES family- and community-engagement literacy project.

Relative Advantage

Relative advantage relates to whether the intervention offers an improvement over current or alternative solutions to the problem.

Key Considerations

1. How adequate are the current solutions to the problem?
2. Is this intervention more effective than the current solution?
3. Is this intervention more effective than other innovative models established?

The GoM has historically focused much of its early grade literacy efforts on improving reading skills through in-service teacher training and improved pedagogical instruction. Other projects that were implemented or are currently being implemented in Mali, such as USAID SIRA, USAID PHARE, and Ecole et Langues Nationales en Afrique focus on teaching instruction. These approaches have been relatively successful in their ability to improve teaching instruction and early grade literacy skills, but stakeholders and beneficiaries had outstanding challenges that these projects were not able to solve. Instead of designing a completely different model, the OCLR project leveraged the preexisting balanced literacy approach and teacher training model and combined it with tablets equipped with the Stepping Stone app to solve issues of lack of adequate follow-up after teacher training and insufficient support to teachers. By giving teachers tablets equipped with training materials and resources to help them correctly implement the pedagogical approach, the OCLR project provided a potentially low-cost solution to the challenges of the previously implemented interventions.

In EOP interviews, the OCLR project managers noted that a key strength of the project was the ability of teachers who received tablets with the Stepping Stone app to access follow-up training materials at any time. They also noted, however, that the training materials that the project provided during this pilot were insufficient—teachers had access to three videos, whereas project management would have preferred to develop ten or more videos. The quality and range of content in the videos could have been improved, as well. They mentioned that offering videos that showed a range of teaching levels—modeling low-, average-, and high-performing implementation of strategies—would allow viewers to observe the differences between strong and poor implementation of the balanced literacy approach.

Relevance Advantage Conclusion

The OCLR project has a relative advantage over existing solutions that attempt to improve teaching instruction and early grade literacy skills. RARE incorporated the preexisting balanced literacy approach and teacher training model that has previously been implemented in Mali through other donor-funded projects and the MEN; it added a technology component to solve issues of lack of adequate follow-up after teacher training and insufficient support to teachers. However, it is not clear if the amount of content created and distributed to teachers via the Stepping Stone app was sufficient for teachers to master fully the pedagogical approach. The OCLR project should, in future iterations, consider developing more training materials for teachers to better support improved teaching instruction.

Ease of Transfer and Adoption

Ease of transfer and adoption relates to whether the characteristics and components of the intervention lend themselves to being adopted by organizations other than the original implementer. This parameter of scalability looks at how complex or resource-heavy an intervention is, as well as whether specific elements of the intervention may be deemed inappropriate or unattractive to other implementers.

Key Considerations[72]

1. What is the level of technical sophistication of the intervention's components and activities?
2. What is the level of complexity of the intervention?
3. What level of supervision and monitoring is needed?

The components of the OCLR project required mixed levels of technical sophistication. The Stepping Stone app, a low-cost platform that can be used in a variety of contexts, does not require high levels of technical literacy for users to be able to access and interact with the content. In EOP interviews, EDC noted that the Stepping Stone app could easily be updated to accommodate different languages and devices. However, because the Stepping Stone app is a proprietary platform, it is unlikely that the programming and maintenance of the software could be done without technical expertise from EDC. EDC staff created, processed, and uploaded all training materials onto the Stepping Stone app, which would constrain the adoption of this platform by new projects without their support.

Additionally, the balanced literacy approach and training materials were developed for use in Mali over many years and required high levels of engagement from the local MEN stakeholders and technical experts, as well as support from international NGOs. As RARE and EDC were involved in the rollout of the balanced literacy approach through the USAID PHARE project, they held high levels of technical knowledge that allowed them to

72 In the original tool, this section includes 11 questions. This analysis includes the questions deemed most relevant for the intervention model and context.

incorporate the pedagogical approach into the OCLR project more seamlessly. The relationships that RARE and EDC had cultivated with the MEN through the previous project also contributed to their ability to engage the MEN in the OCLR project approach and leverage local government support. It is unclear if other organizations without a technical understanding of the balanced literacy approach or preexisting relationships with those who oversee local and national primary pedagogical approaches would be able to effectively adopt the OCLR project model, either in Mali or in other country contexts.

Overall, the intervention is relatively complex due to the multiplicity of its components: filming teachers in classrooms, holding teacher training, distributing balanced literacy resources and IRI kits, uploading materials to the Stepping Stone app, and ensuring that relevant follow-up support is conducted.

The OCLR project required moderate levels of supervision and monitoring. During its pilot implementation period, RARE conducted in-person training with all participating teachers, coordinated regular follow-up sessions with teachers by pedagogical counselors, and collected M&E and FOI data in teacher classrooms. The Stepping Stone app contains an internal tracking system that captures usage data, which minimizes the level of in-person M&E needed to monitor user progress. For example, the app tracks how often and how many times a teacher has read or listened to content. RARE team members manually collected this data from tablets; user internet access was inconsistent, and it was not possible to remotely upload log data to a central location from which RARE could analyze user activity in real time. Furthermore, the usage logs from the Stepping Stone app had technical challenges that made it difficult for the OCLR project staff to use them as a sole source of monitoring participant progress. Future implementers would need to take into account the cost-benefit of more routine M&E visits to teachers to observe their classroom practices, as well as the necessary investments in improving the Stepping Stone app's ability to effectively capture usage data, including potentially recording and transmitting teaching practices for project management to monitor remotely.

Ease of Transfer and Adoption Conclusion

It would be relatively challenging for other organizations to transfer and adopt the OCLR project model. The project components each required significant preexisting knowledge from the RARE and EDC teams—including in-depth understanding of the balanced literacy approach and the Stepping Stone app. It is unclear if a different organization, either in Mali or another country, could easily adopt and implement the model without support from technical experts in the pedagogical approach, knowledge of the Stepping Stone app, or strong relationships with the local ministry of education. Nevertheless, the technology component of the OCLR project has the potential to significantly reduce the amount of time and resources that a future implementer would need to spend on M&E, as the Stepping Stone app, with technical updates, could capture and transmit monitoring and classroom observation data remotely.

Testability

The testability parameter examines how easy it is for organizations to pilot the intervention on a small scale before full adoption. Testability assesses whether potential adopters would need to commit significant resources or time to test the model if they chose to pilot it in a new context.

Key Consideration

1. Is the model able to be tested on a limited scale?

The OCLR project model is not easily tested in a new context. The Stepping Stone app is a content-delivery option available on basic tablets and mobile phones; because the technology and training materials had already been tested and used in Mali, the OCLR project could be rolled out in other Bamanankan-speaking areas of the project on a small scale. Adaptation of the technology component to other contexts has already been tested; in addition to Mali, the Stepping Stone app has been rolled out in other countries, including Zambia, for the USAID Time to Learn project, which engaged 2,250 teachers in 2014.[73] Nevertheless, EDC staff would need to program the software in a new language and create, process, and insert all relevant components into the Stepping Stone app. The project cannot be implemented elsewhere without support from EDC or other technically competent partners. It is unclear what level of time or financial investment would be required to update the platform for a new language.

Furthermore, the OCLR project model necessitates adequate follow-up training materials for teachers to be able to confidently and effectively implement the literacy approach in their classroom without in-person follow-up. If a potential adopter were to implement the OCLR project model in a new context, they would need to make a significant financial investment in the development of training materials—print, digital, and video—to replicate the model.

Testability Conclusion

The OCLR project model is easily replicable in Bamanankan-speaking areas of Mali. However, potential adopters would need to invest significant time and financial resources to replicate and test the project in a new language or country context.

Sustainability of Funding

Sustainability of funding refers to how cost effective the intervention is and whether there are funds available to scale the intervention, either through government or other organizations.

Key Considerations

1. Is the model more cost effective than other solutions?
2. What kind of funding commitment is required to scale the model?
3. Is there any potential for internal revenue from the model (e.g., service fees)?

No comprehensive cost-effectiveness analysis was conducted on the OCLR project; instead, a cost analysis was performed. A cost analysis is often a component of scalability assessments, as it helps decision makers and stakeholders understand the feasibility of replication within given budgetary constraints. Because ACR GCD grantees implement new approaches, they often allot significant financial resources to developing new materials that could be used on a recurring basis. To better understand the funding requirements of the OCLR project, a cost analysis was conducted to present the total cost of the intervention, including the investments that would be needed for project replication or scale-up.

73 Education Development Center, Inc. (n.d.). *History of Stepping Stone*. Retrieved from http://sstone.edc.org/en/what-is-stepping-stone-4/history-of-stepping-stone/

USAID guidance on conducting cost analyses on early grade reading projects suggests that the "ingredients method"[74] be used to calculate costs in the following categories:

- Management and associated technical costs
- Development costs
- Implementation costs

Project staff completed a costing template with guidance from World Vision and STS. Invoiced costs from the beginning of the project through June 2017 were presented based on the three categories listed above.[75] Table 8 provides a detailed breakdown of costs by category under the OCLR project's key project activities.

Table 8: **Cost Analysis**[76]

Activity	Management	Development	Implementation
EGRA baseline data collection	$ -	$ 172	$ 19,281
EGRA endline data collection	$ -	$ 117	$ 11,767
Fidelity of implementation tool development and piloting	$ -	$ 1,604	$ -
Project field meetings	$ -	$ -	$ 2,178
Project start-up and in-country launch activities	$ -	$ -	$ 8,959
School director and teacher training including materials development	$ -	$ 13,917	$ 30,396
Training of pedagogical counselors	$ -	$ 1,853	$ 7,061
Monitoring visits to schools	$ -	$ -	$ 17,373
Total	**$ 130,542**	**$ 17,664**	**$ 97,014**
Percentage of total (%)	53.2	7.2	39.6

The management category includes costs that are not directly related to implementation and are likely to vary widely based on who is overseeing the intervention. **Management costs for the OCLR project represented 53.2 percent of the costs expended** and included the cost of maintaining the project office in Bamako, personnel salaries, and other indirect rates and fees.[77]

74 RTI International. (2015). Measurement and research support to education strategy goal 1: Early grade reading costing template and guidance. Washington, D.C.: USAID. Retrieved from http://pdf.usaid.gov/pdf_docs/PBAAF458.pdf

75 The total grant amount for the OCLR project was $329,265. At the end of the grant on June 30, 2017, the project had invoiced $245, 220.

76 Due to rounding, totals may not equal the sum of activity costs.

77 Management costs are inclusive of a 17.0% flat fee charged for Negotiated Indirect Cost Recovery Agreement (NICRA), which captures indirect costs, including regional management and technical support, the local RARE country representative, overseas operations management (RARE headquarters), program quality and support (RARE headquarters), and shipping and procurement costs. This also captures miscellaneous headquarters-based services that were provided to the project, including finance, internal auditing, human resources, executive management, board, global knowledge, and information management. This analysis assumes that no NICRA expenses were also billed as independent line items, although it should be recognized that some double-counting may have occurred.

Development includes the costs related to the development of materials, survey instruments, programs, and other content that would not need to be redeveloped in the scale-up of a project. **The development costs for the OCLR project represented 7.2 percent of the costs expended.** The major expenses within this category were the development of training materials and the EGRA instruments. These costs are one-off expenditures that would not be incurred again if a similar project were implemented in Bamanankan-speaking areas in Mali.

The implementation cost category is arguably the most relevant for stakeholders who are considering scaling up a project or intervention. This category includes all of the recurrent activities and costs that would need to be expensed should the project be replicated, including: materials printing and distribution, training, M&E, events and presentations, workshops, and human resources activities. **For the OCLR project, implementation costs represented about 39.6 percent of the total project cost.** Within this cost category, the largest expenses were trainings and monitoring visits to schools.

Projects sometimes benefit from in-kind services, institutional support, or preexisting relationships with stakeholders or governments that may provide the project with tangible benefits, although it may be difficult or impossible to monetize the costs. Examples of this include local volunteers, strong capacity or support from a large NGO, or relationships with local governments that could ease logistics and procedures. The OCLR project benefitted from its longstanding relationship with EDC which, although a subcontractor of the project, may have provided additional nontechnical or nonexpensed support. Notably, development costs were low for the OCLR project because the Stepping Stone app was already developed and tested, as were many of the balanced literacy approach training materials. RARE also had strong connections with the MEN officials and USAID/Mali, which provided support and guidance at points throughout the project.

Scalability of Funding Conclusion

Management costs for the OCLR project represented a larger proportion of the overall cost of the project than implementation costs, and development costs represented a very minimal proportion of the cost of the project. A scale-up of the project would not require significant investments in development unless it is determined that reading materials and the Stepping Stone app require improvements, indicating that it may be cost effective to replicate the project. More rigorous cost and impact data should be collected to understand better the cost effectiveness of the model better.

X. Conclusions

RARE leveraged existing literacy approaches, materials, and software to deliver a pilot project that aimed to improve reading and writing instruction in Grades 1 and 2 by training teachers—with the overall goal of improving pre-reading and foundational skills of students. The project provided in-person training and follow-up visits to teachers; half of the trained teachers received tablets with supplemental training materials that they could access on-demand through the Stepping Stone app. EGRA results indicate students whose teachers were trained by the OCLR project had higher pre-reading and foundational skills at endline, though there was no evidence that students whose teachers had access to tablets outperformed students whose teachers did not have access to tablets.

The following lessons should be considered for any future interventions incorporating components of the OCLR project.

Lessons Learned

Traditional teacher-training models can impact the pre-reading and foundational skills of students, though monitoring and follow-up support are expensive and time-consuming.

Across endline subtasks, intervention group students outperformed their peers in the comparison group. This corroborates the commonly accepted knowledge that teacher training, particularly in the balanced literacy approach, can have positive impacts on early grade reading skills of students. However, the pedagogical counselors hired to provide follow-up support to teachers after the in-person training experienced challenges, such as lack of time to visit all teachers regularly and unreliable forms of transportation to reach schools. Nearly 40.0 percent of the total cost of the project was allocated to implementation, much of which was spent on monitoring visits. Projects with similar scopes should explore leveraging technology to conduct remote monitoring sessions rather than investing heavily in in-person monitoring sessions.

Although technology may minimize the cost of in-person support after teacher training, projects should ensure that content is sufficient before substituting in-person support with remote, technology-based support. In other words, technology does not replace good content.

The OCLR project provided half of the trained teachers with tablets with the Stepping Stone app, which gave them access to three 90-minute videos demonstrating effective implementations of the balanced literacy approach in a classroom setting. EGRA results did not provide evidence that students whose teachers had access to the tablets outperformed those whose teachers did not have access to the tablets. This indicates that the technology and literacy content as delivered did not provide an added value to the OCLR project. Project management noted that they did not believe that the amount of content provided—three 90-minute videos—was sufficient; in future projects, they hope to develop ten or more videos. The project spent only 7.2 percent of its total budget on development. More investment in digital content in the future might yield different impacts on student learning.

Projects that target teacher classroom behavior should invest in rigorous, standardized classroom observations over multiple time points.

Although the OCLR project tracked FOI data and conducted visits to observe teacher practices, data were not made available to correlate standardized classroom observation with EGRA results. As a result, this analysis can only infer that the higher performance of students whose teachers were trained through the OCLR project was a result of improved teaching practices and adoption of the balanced literacy approach. Without supporting data, it is unclear to what extent teachers truly adopted the practices and, thus, whether different levels of uptake of the approach impacted students' EGRA scores in different ways. In the future, the OCLR project could explore standardizing classroom observation data collection or leveraging technology to conduct observations remotely.

Annexes

Annex A: **EGRA Instrument 1 at Baseline**

Enumerator Name	

La date et l'heure

Date	
Time	

Le site de l'ecolé

CAP	
Commune	
Ecole	
Code	

L'identification de l'élève

L'identification de l'élève	

Le consentement

I ni sɔgɔma! Ne tɔgɔ ye _____ I ɲɔgɔn demisɛnninw bɛ ne bolo. Kalanjɛ, farikoloŋɛnajɛ ani ntolatan ka di u ye. E dun, e tɔgɔ ? Mun de ka di e ye ?

[Attendez la réponse de l'enfant. Si l'enfant semble à l'aise, passez directement au consentement verbal.

N'i ma taa kalanyɔrɔ la don min, i bɛ mun kɛ? (Le jour où tu ne vas pas à l'école, que fais-tu ?)

Veuillez lire, à haute voix, la déclaration suivante à l'élève pour obtenir son consentement verbal:

N bɛ n nakun fɔ i ye. Kalan minisiriso y'a ɲini ka denmisɛnninw ka kalanjɛ kɛcogo kiimɛ. E sugandira k'i sendon o kiimɛni na. Nafaba de bɛ i sendonni in na ; nka n'a man di i ye, diyagoya tɛ.

An bɛna lamɛnni ni kalanjɛ tulon dɔw kɛ.

I bɛ waati min kɛ fɛn dɔw kalanni na, o bɛ jateminɛ. Nka ɲɔgɔndan tɛ. Ne ni e bɛ min kɛ, o tɛ foyi falen i ka kuruw la kalanso kɔnɔ. N bɛna ɲininkali dɔw k'i la fana aw ka du kan. Maa si tɛn'a dɔn ko e ka jaabiw don. N'i t'a fɛ ka ɲininkali min jaabi, i b'o to yen. N b'a fɔ i ye hali bi, i diyagoyalen tɛ k'i sendon kiimɛni in na, n'a ma bɛn i ma.

I sɔn na wa? *[Attendez la réponse de l'élève avant de poser la prochaine question. Si l'élève dit 'oui' à la question, posez la question suivante. Si l'élève dit 'non', remerciez l'élève et passez au prochain élève.]*

An bɛ se k'a daminɛ wa?

Consentement verbal obtenu: ☐

Information de l'élève

Nom de l'élève

Age de l'élève

L'élève est dans quelle classe ?

☐ 1ere ☐ 2eme ☐ 3eme

Le sexe de l'élève

☐ Fille ☐ Garçon

Orientation a la Lecture

[Montrez à l'élève la Feuille A]

Lisez les instructions suivantes et enregistrez les réponses de l'élève :

I tɛna masalabolo in kalan fɔlɔ, sisan n'i bɛna masalabolo in kalan i bɛ a kalan ka taa fan jumɛ fɛ.

([L'élève déplace son doigt de la gauche à la droite])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Sisan, ne b'a fɛ, i ka masalabolo in dameninɛ jira.

([L'élève pose le doigt sur la 1ére ligne, le mot le plus à gauche « Bi / Bi sɔgɔmada »])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Sisan, ne b'a fɛ, i ka masalabolo in laban jira.

([L'élève pose le doigt sur « la / lakɔli la »])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Sisan, sira fɔlɔ laban jira.

(L'élève déplace son doigt sur le premier « Mun »])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

N'i sera sira fɔlɔ laban na, i bɛna sira min kalan o kɔ, o jira.

([L'élève déplace son doigt vers le mot placé le plus à gauche de la seconde ligne – « b'i ».])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Sisan, kumasen fɔlɔ laban jira.

([L'élève pose son doigt sur « ye »])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Identification du son initial

Instructions à l'élève :

Nin ye degeli de ye min bɛ kɛ baro senfɛ. N bɛna daŋɛ kelen fɔ i ye. Nb'a fɔ ka segin a kan. O kɔfɛ, i mana siginiden min mankan mɛn daŋɛ in daminɛ na, i b'o fɔ n ye. I sɔnna wa?

Misali la "fa" daŋɛ bɛ daminɛ ni « fff » mankan ye. O tɛ wa ? « fa » daŋɛ bɛ fɔlɔ ni mankan jumɛn ye ? « fa» ?
[Attendre que l'élève répète le son "fff". S'il ne répond pas, dites-lui, "« fa » daŋɛ bɛ daminɛ ni « fff » mankan ye.*]*

An ka misali wɛrɛw lajɛ :

"sisi" daŋɛ bɛ daminɛ ni mankan jumɛn ye? « sisi » ?

(Si l'élève répond correctement, dites-lui a ka ɲi kosɛbɛ. « sisi » daŋɛ bɛ daminɛ ni « sss » mankan ye.)

(Si l l'élève ne répond pas, dites-lui « « sisi » daŋɛ bɛ daminɛ ni « sss » mankan ye.)

« taga » daŋɛ bɛ daminɛ ni mankan jumɛn ye? « taga » ?

(Si l'élève répond correctement, dites-lui «a ka ɲi kosɛbɛ! « taga » daŋɛ bɛ daminɛ ni « t' » mankan ye.)

(Si l'élève ne répond pas, dites-lui « « taga » daŋɛ bɛ daminɛ ni « t' » mankan ye.)

« Ami» daŋɛ bɛ daminɛ ni mankan jumɛn ye ? « Ami » ?

(Si l'élève répond correctement, dites-lui «a ka ɲi kosɛbɛ! « ami» daŋɛ bɛ daminɛ ni « a » mankan ye.)

(Si l'élève ne répond pas, dites-lui « Ami » daŋɛ bɛ daminɛ ni « a » mankan ye.)

Ne bɛ min nɔfɛ i y'o faamu wa ? Sisan, ne bɛna daŋɛ wɛrɛw kalan i ye. N bɛ daŋɛ bɛɛ kelen kelen kalan siŋɛ fila. I tulomajɔ kosɛbɛ. I bɛ fɔlɔ ka mankan min mɛn daŋɛ daminɛ na, i b'o fɔ n ye. I sɔnna wa ?

Ne pas corriger l'élève pendant le test.

En cas de non-réponse ou d'hésitation de sa part, après 3 secondes, relancer la question. Si l'élève ne réponds pas, marquer la case « Pas de réponse » et passez au prochain item.

1. « ba » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « ba » ? (/b'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

2. « di » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « di » ? (/d'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

3. « gafe » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « gafe » ? (/g'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

4. « Umu » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « Umu » ? (/uuu/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

5. « so » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « so » ? (/ssss/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

6. « pili » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « pili » ? (/p'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

7. « kɔ » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « kɔ » ? (/k'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

8. « malo » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « malo » ? (/mmm/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

9. « ɲɛ » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « ɲɛ » ? (/ɲ'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

10. « walan » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « walan » ? (/w'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

Le son de la lettre

Siginidenw ni siginidenkuluw filɛ ka ɲɛ. Siginiden ninnu kalan i k'u mankan fɔ n ye. Misali la, nin siginiden in : *[Indiquer le "a": dans la ligne des exemples]*Ale bɛ kalan /a/ i n'a fɔ "naji" daɲɛ kɔnɔ.

An k'a waleya sisan. Nin siginiden in kalan *[Indiquer le "l" dans le rang des exemples]:*

Si l'élève répond correctement, dites: A ka ɲi kosɛbɛ, siginiden in bɛ kalan /l/ i n'a fɔ "lɛfɛ" daɲɛ kɔnɔ.

Si l'élève ne répond pas correctement, dites: Ayi, siginiden in bɛ kalan /l/ i n'a fɔ "lɛfɛ" daɲɛ kɔnɔ.

An ka misali wɛrɛw lajɛ. Nin siginiden in kalan *[Indiquer le "aa" dans le rang des exemples]:*

Si l'élève répond correctement, dites: A ka ɲi kosɛbɛ, siginiden in bɛ kalan /aa/ i n'a fɔ "naani" daɲɛ kɔnɔ.

Si l'élève ne répond pas correctement, dites: Ayi, siginiden in bɛ kalan /aa/ i n'a fɔ "naani" daɲɛ kɔnɔ.

An ka misali wɛrɛw lajɛ tun. Nin siginiden in kalan *[Indiquer le "en" dans le rang des exemples]:*

Si l'élève répond correctement, dites: A ka ɲi kosɛbɛ, siginidenkulu in bɛ kalan /en / i n'a fɔ « den » daɲɛ kɔn

Si l'élève ne répond pas correctement, dites: Ayi, siginidenkulu in bɛ kalan /en/ i n'a fɔ « den » daɲɛ kɔnɔ

I y'a faamu wa? An bɛ se ka taa a fɛ ? Ni ne ko "a daminɛ", i kɛtɔ ka siginiden fɛn o fɛn kalan, i b'i bolo da o kan. I b'u kalanni daminɛ numanfɛ ka taa kininfɛ sira ni sira. I y'a faamu kosɛbɛ wa? I bolo da sigiden fɔlɔ kan. I labɛnnen don wa? I b'a lajɛ k'u kalan ka ɲɛ teliya la. A damiɛ!

b	a	u	l	s	o	ɔ	L	u	c
k	d	nj	h	t	e	l	ii	m	ɔ
r	u	c	ns	p	ee	ɛ	n	e	b
n	an	ɛ	L	ɔn	t	M	oo	l	g
nt	o	uu	h	u	d	W	r	g	l
k	nc	s	f	a	n	An	a	k	nf
ɛ	w	on	L	ng	s	np	y	a	s
j	a	ɛn	l	ɛɛ	p	nb	y	a	ɔ
m	ŋ	z	nk	b	e	U	L	d	aa
g	a	r	ŋ	ɔɔ	o	K	un	l	en

Time Remaining

Autostop? ☐

Mots inventés

Daŋɛ dɔw filɛ, lala i ma deli ka minnu ye. Nka ne tun b'a fɛ i k'a lajɛ k'u kalan. Misali la, daŋɛ fɔlɔ in bɛ kalan «gɛ» *[Indiquer le mot* « gɛ» *avec le doigt].* I bɛ se ka segin daŋɛ fɔlɔ in kalanni kan wa ?

[Après sa réponse, ou après 3 secondes dans le cas de non-réponse, montrez-lui comment faire.]

Daŋɛ in dun ? *[indiquer le mot « zii » avec le doigt].* I bɛ se k'o kalan wa ?

[Après sa réponse, ou après 3 secondes dans le cas de non-réponse, montrez-lui comment faire.]

Nin dun ? *[indiquer le mot « hu » avec le doigt].* I bɛ se k'o kalan wa ?

[Après sa réponse, ou après 3 secondes dans le cas de non-réponse, montrez-lui comment faire.]

I y'a faamu wa ? N bɛ min nɔfɛ i y'o faamu wa? Ni ne ko "a daminɛ", i bɛ siraw ta kelen kelen k'u kalan k'a daminɛ numanfɛ ka taa kininfɛ. N'i sera sira dɔ laban na, i b'o nɔkanta daminɛ. I labɛnnen don wa? I b'a lajɛ k'u kalan ka ŋɛ teliya la. A daminɛ!

zi	fe	do	lu	tee
laa	bii	kee	mo	sawa
ki	gibɔ	lezo	fuki	cuto
gamo	Luba	yow	basɔ	pifo
pa	kiwɔ	zaa	yenu	jowe
guu	Mire	maja	dɛca	nsɔ
yɛbu	lina	ŋipɛ	tansa	yonpe
wɛn	Mudo	sipu	poora	ŋasi
zuso	wɛɛ	Loo	lunan	njew
ŋope	Nbeli	luro	pini	leko

Time Remaining

Autostop? ☐

Lecture du texte 1

Sisan, n b'a fɛ i ka maana in kalan. I b'i kan bɔ kosɛbɛ A lajɛ i k'a kalan ka ɲɛ teliya la; o kɔ ne bɛ ɲininkali dɔw kɛ i la. Ni ne ko i k'a daminɛ, i b'a daminɛ yan *(Mettez la feuille de la Section 5 devant l'élève (F/5). Montrez du doigt le premier mot du passage).* I labɛnna wa ? An k'a daminɛ.

[Faites démarrer le chrono en appuyant sur le bouton START / STOP]

Samiyɛ	waati	don.	Ji	sigira
Sibi	bɔlɔnw	kɔnɔ.	Dɔgɔ	don,
Fati	ye	a	ka	ɔrɔbu
kura	don.	A	n'a	terimuso
Umu	taara	sugu	la.	U
bɛ	taama	na.	Sɔɔnin,	Fati
binna.	A	kasira.	A	y'a
ka	ɔrɔbu	lajɛ.	A	seginna
so.	A	ba	ye	ɔrɔbu
kura	wɛrɛ	di	a	ma.

Time Remaining

Autostop? ☐

Questions de Compréhension

[Reprendre le texte]

Sisan, i bɛna ɲininkali damadɔ jaabi maana in kan.

1. Ko in kɛra san waati jumɛn ? ([Samiyɛ])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

2. Fati ye mun don? ([ɔrɔbu])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

3. Mun ye Fati sɔrɔ ? ([A binna])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

4. Jɔn kasira ? ([Fati])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

5. Fati binna. A ka ɔrɔbu bɛ cogo di ? ([ɔrɔbu nɔgɔlen])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

Compréhension à l'audition

Sisan, ne bɛna maana kelen kalan i ye siɲɛ kelen. O kɔ, n bɛ ɲininkali damadɔ k'i la maana in kan. I bɛ maana in lamɛn kosɛbɛ. I bɛ tila ka ɲininkaliw jaabi i fɛrɛ ma' I sɔnna wa? N b'a fɛ i ka min kɛ i y'o faamu wa? An k'a daminɛ. A lamɛn kosɛbɛ:

Bi ye seli ye.

Ma ye Buba n'a dɔgɔmuso Fanta ka fini kuraw labɛn.

U y'u pari ka taa warabafilɛso la.

U taara mɔbili ɲini sirada la.

U mɛɛnna u ma mɔbili sɔrɔ bawo mɔbili bɛɛ falen don.

Laban na, mɔbili dɔ sɔrɔla.

U selen warabafilɛso la Buba ni Fanta ye ji suma san.

O kɔfɛ, u ye waraba, sama, bama ani bagan caman wɛrɛw ye.

U ye fotow ta ani k'u teriw ye.

Seli diyara dɛ !

1. Buba ni Fanta taara min ? ([Warabafilɛso la.])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

2. Munna u ma mɔbili sɔrɔ joona ? ([Bawo mɔbili bɛɛ falen don. Bawo selidon don.])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

3. Bagan jumɛnw bɛ sɔrɔ warabafilɛso la ? ([Waraba, sama, banba (hali n'a ye bagan fila fɔ)/ Kungokɔnɔ baganw.])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

4. Jɔnw ye fotow ta ? ([Buba ni Fanta])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

5. Munna u ye ji suma san ? ([Bawo minnɔgɔ b'u la.])

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

Entretien sur l'environnement de l'élève.

An tilala ka ban. An tɔ ye ɲininkali damadɔw ye e kan, aw ka du kan, i ka kalan kan ani aw ka so kan.

1. E si ye san joli ye? *((Enregistrez le nombre d'années. Si l'élève ne sais pas ou ne répond pas, enregistrez 99.))*

2. I bangena kalo jumɛn ? *((Enregistrez le mois. Si l'élève ne sais pas ou ne répond pas, enregistrez 99.))*

3. I bangena san jumɛn ? *((Enregistrez l'année. Si l'élève ne sais pas ou ne répond pas, enregistrez 99.))*

4A. I bɛ kan jumɛn (w) fɔ so ? ([Jaabi caman bɛ se ka di])

☐ Bamanankan ☐ Fulfuldé ☐ Songhoi ☐ Bomu ☐ Français ☐ Arabe ☐ Autre

☐ Pas de réponse / ne sais pas

4B. *Si la réponse est autre, précisez:*

5. Kalanjɛ gafe dɔ b'i bolo wa?

☐ Oui ☐ Non ☐ Pas de réponse / ne sais pas

6A. Gafe wɛrɛw, kunnafonisɛbɛnw walima fɛn kalanta wɛrɛ b'i bolo k'a bɔ kalanso taw wa ?

☐ Oui ☐ Non ☐ Pas de réponse / ne sais pas

6B. Misali damadɔw di ((Pas besoin d'enreigstrer la réponse))

7A. I bɛ kan jumɛn (w) fɔ so ? ([Jaabi caman bɛ se ka di])

☐ Français ☐ Bamanankan ☐ Fulfuldé ☐ Songhoi ☐ Bomu ☐ Arabe ☐ Autre

☐ Pas de réponse / ne sais pas

7B. *Si la réponse est autre, précisez:*

8. K'a bɔ e la, mɔgɔ wɛrɛ bɛ a' ka du kɔnɔ min bɛ se kalanjɛ la wa ?

☐ Oui ☐ Non ☐ Pas de réponse / ne sais pas

9A. Jɔn ni jɔn bɛ se kalanjɛ la aw ka so ? *((Plusieurs réponses sont autorisée])*

☐ Oui ☐ Non ☐ Pas de réponse / ne sais pas

9B. *Si la réponse est autre, précisez:*

10. Arajo b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

11. Telefɔni b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

12. Yeelen (kuran) b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

13. Tele b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

14. Firigo (jisumanyalan) b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

15. Sokɔnɔŋɛgɛn b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

16. Nɛgɛso b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

17. Moto b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

18. Wotoro walima kurun walima pinasi b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

19. Mɔbili, kamiyɔn, 4x4, sɛnɛkɛmansin b'aw ka so wa ?

☐ Oui ☐ Non ☐ Pas de réponse

20. E ye zaridɛn kɛ yanni e ka don lakɔli la wa ?

☐ Oui ☐ Non ☐ Pas de réponse / ne sais pas

21. I bɛ kalanso jumɛn na ɲinan ?

☐ 1ère année ☐ 2ème année ☐ 3ème année ☐ 4ème année

22. E tun bɛ kilasi jumɛn na salon?

☐ Jardin d'enfants ☐ 1ère année ☐ 2ème année ☐ 3ème année ☐ 4ème année

☐ Pas à l'ecole ☐ Pas de réponse / ne sais past ☐ Autres

23. Yala karamɔgɔ bɛ to ka baara d'e ma ka kɛ so wa ?

☐ Oui ☐ Non ☐ Pas de réponse / ne sais pas

24. [Ni 8 jaabi ye ɔwɔ ye] Yala mɔgɔ b'i dɛmɛ ka baara in kɛ tuma dɔw wa ?

☐ Oui ☐ Non ☐ Pas de réponse / ne sais pas

25. Salon, e ye kalan bila ka tɛmɛ dɔgɔkun kelen kan wa ?

☐ Oui ☐ Non ☐ Pas de réponse / ne sais pas

Annex B: **EGRA Instrument 2 at Endline**

Enumerator Name	

La date et l'heure

Date	
Time	

Le site de l'ecolé

CAP	
Commune	
Ecole	
Nomd'eleve	
ID	
Groupe	
Code	

Autre élève

Nom d'élève: ((nom et prénom))	

Verification de l'élève

Est-ce [Nom d'élève]? ☐ Oui ☐ Non

Raison d'absence

Pourquoi est-ce que l'élève ne peut pas passer l'EGRA? ☐ L'enfant est absent ☐ L'enfant ne va plus à l'école

☐ Son handicap empêche l'enfant de faire passer l'évaluation ☐ L'élève est malade ☐ Autre

Si autre, précisez:	

Nom d'élève remplacement

Nom d'élève qui est le empaçant(e): ((nom et prénom))	

L'identification de l'élève

L'identification de l'élève	

Le consentement

I ni sɔgɔma! Ne tɔgɔ ye _____ I ɲɔgɔn demisɛnninw bɛ ne bolo. Kalanjɛ, farikoloɲɛnajɛ ani ntolatan ka di u ye. E dun, e tɔgɔ ? Mun de ka di e ye ?

[Attendez la réponse de l'enfant. Si l'enfant semble à l'aise, passez directement au consentement verbal. S'il hésite ou a l'air peu à l'aise, posez la deuxième question avant de passer au consentement verbal].

N'i ma taa kalanyɔrɔ la don min, i bɛ mun kɛ? (Le jour où tu ne vas pas à l'école, que fais-tu ?)

Veuillez lire, à haute voix, la déclaration suivante à l'élève pour obtenir son consentement verbal:

N bɛ n nakun fɔ i ye. Kalan minisiriso y'a ɲini ka denmisɛnninw ka kalanjɛ kɛcogo kiimɛ. E sugandira k'i sendon o kiimɛni na. Nafaba de bɛ i sendonni in na ; nka n'a man di i ye, diyagoya tɛ.

An bɛna lamɛnni ni kalanjɛ tulon dɔw kɛ.

I bɛ waati min kɛ fɛn dɔw kalanni na, o bɛ jateminɛ. Nka ɲɔgɔndan tɛ. Ne ni e bɛ min kɛ, o tɛ foyi falen i ka kuruw la kalanso kɔnɔ. N bɛna ɲininkali dɔw k'i la fana aw ka du kan. Maa si tɛn'a dɔn ko e ka jaabiw don. N'i t'a fɛ ka ɲininkali min jaabi, i b'o to yen. N b'a fɔ i ye hali bi, i diyagoyalen tɛ k'i sendon kiimɛni in na, n'a ma bɛn i ma.

I sɔn na wa? *[Attendez la réponse de l'élève avant de poser la prochaine question. Si l'élève dit 'oui' à la question, posez la question suivante. Si l'élève dit 'non', remerciez l'élève et passez au prochain élève.]*

An bɛ se k'a daminɛ wa?

Consentement verbal obtenu: ☐

Information de l'élève

Nom de l'élève

Age de l'élève

L'élève est dans quelle classe ?

☐ 1ere ☐ 2eme ☐ 3eme ☐ 4eme ☐ 5eme

Le sexe de l'élève

☐ Fille ☐ Garçon

Orientation a la Lecture

[Montrez à l'élève la Feuille A]

Lisez les instructions suivantes et enregistrez les réponses de l'élève :

I tɛna masalabolo in kalan fɔlɔ, sisan n'i bɛna masalabolo in kalan i bɛ a kalan ka taa fan jumɛ fɛ.

([L'élève déplace son doigt de la gauche à la droite])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Sisan, ne b'a fɛ, i ka masalabolo in damininɛ jira.

([L'élève pose le doigt sur la 1ére ligne, le mot le plus à gauche « Bi / Bi sɔgɔmada »])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Sisan, ne b'a fɛ, i ka masalabolo in laban jira.

([L'élève pose le doigt sur « la / lakɔli la »])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Sisan, sira fɔlɔ laban jira.

(L'élève déplace son doigt sur le premier « Mun »])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

N'i sera sira fɔlɔ laban na, i bɛna sira min kalan o kɔ, o jira.

([L'élève déplace son doigt vers le mot placé le plus à gauche de la seconde ligne – « b'i ».])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Sisan, kumasen fɔlɔ laban jira.

([L'élève pose son doigt sur « ye »])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Identification du son initial

Instructions à l'élève :

Nin ye degeli de ye min bɛ kɛ baro senfɛ. N bɛna daŋɛ kelen fɔ i ye. Nb'a fɔ ka segin a kan. O kɔfɛ, i mana siginiden min mankan mɛn daŋɛ in daminɛ na, i b'o fɔ n ye. I sɔnna wa?

Misali la "fa" daŋɛ bɛ daminɛ ni « fff » mankan ye. O tɛ wa ? « fa » daŋɛ bɛ fɔlɔ ni mankan jumɛn ye ? « fa» ?
[Attendre que l'élève répète le son "fff". S'il ne répond pas, dites-lui, "« fa » daŋɛ bɛ daminɛ ni « fff » mankan ye.*]*

An ka misali wɛrɛw lajɛ :

"sisi" daŋɛ bɛ daminɛ ni mankan jumɛn ye? « sisi » ?

(Si l'élève répond correctement, dites-lui a ka ɲi kosɛbɛ. « sisi » daŋɛ bɛ daminɛ ni « sss » mankan ye.)

(Si l l'élève ne répond pas, dites-lui « « sisi » daŋɛ bɛ daminɛ ni « sss » mankan ye.)

« taga » daŋɛ bɛ daminɛ ni mankan jumɛn ye? « taga » ?

(Si l'élève répond correctement, dites-lui « a ka ɲi kosɛbɛ! « taga » daŋɛ bɛ daminɛ ni « t' » mankan ye.)

(Si l'élève ne répond pas, dites-lui « « taga » daŋɛ bɛ daminɛ ni « t' » mankan ye.)

« Ami» daŋɛ bɛ daminɛ ni mankan jumɛn ye ? « Ami » ?

(Si l'élève répond correctement, dites-lui « a ka ɲi kosɛbɛ! « ami » daŋɛ bɛ daminɛ ni « a » mankan ye.)

(Si l'élève ne répond pas, dites-lui « Ami » daŋɛ bɛ daminɛ ni « a » mankan ye.)

Ne bɛ min nɔfɛ i y'o faamu wa ? Sisan, ne bɛna daŋɛ wɛrɛw kalan i ye. N bɛ daŋɛ bɛɛ kelen kelen kalan siŋɛ fila. I tulomajɔ kosɛbɛ. I bɛ fɔlɔ ka mankan min mɛn daŋɛ daminɛ na, i b'o fɔ n ye. I sɔnna wa ?

Ne pas corriger l'élève pendant le test.

En cas de non-réponse ou d'hésitation de sa part, après 3 secondes, relancer la question. Si l'élève ne réponds pas, marquer la case « Pas de réponse » et passez au prochain item.

1. « ba » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « ba » ? (/b'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

2. « di » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « di » ? (/d'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

3. « gafe » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « gafe » ? (/g'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

4. « Umu » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « Umu » ? (/uuu/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

5. « so » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « so » ? (/ssss/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

6. « pili » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « pili » ? (/p'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

7. « kɔ » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « kɔ » ? (/k'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

8. « malo » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « malo » ? (/mmm/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

9. « ɲɛ » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « ɲɛ » ? (/ɲ'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

10. « walan » daɲɛ bɛ daminɛ ni mankan jumɛn ye ? « walan » ? (/w'/)

☐ A ka ɲi ☐ A man ɲi ☐ jaabi ma di

Le son de la lettre

Siginidenw ni siginidenkuluw filɛ ka ɲɛ. Siginiden ninnu kalan i k'u mankan fɔ n ye. Misali la, nin siginiden in : *[Indiquer le "a": dans la ligne des exemples]*Ale bɛ kalan /a/ i n'a fɔ "naji" daɲɛ kɔnɔ.

An k'a waleya sisan. Nin siginiden in kalan *[Indiquer le "l" dans le rang des exemples]:*

Si l'élève répond correctement, dites: A ka ɲi kosɛbɛ, siginiden in bɛ kalan /l/ i n'a fɔ "lɛfɛ" daɲɛ kɔnɔ.

Si l'élève ne répond pas correctement, dites: Ayi, siginiden in bɛ kalan /l/ i n'a fɔ "lɛfɛ" daɲɛ kɔnɔ.

An ka misali wɛrɛw lajɛ. Nin siginiden in kalan *[Indiquer le "aa" dans le rang des exemples]:*

Si l'élève répond correctement, dites: A ka ɲi kosɛbɛ, siginiden in bɛ kalan /aa/ i n'a fɔ "naani" daɲɛ kɔnɔ.

Si l'élève ne répond pas correctement, dites: Ayi, siginiden in bɛ kalan /aa/ i n'a fɔ "naani" daɲɛ kɔnɔ.

An ka misali wɛrɛw lajɛ tun. Nin siginiden in kalan *[Indiquer le "en" dans le rang des exemples]:*

Si l'élève répond correctement, dites: A ka ɲi kosɛbɛ, siginidenkulu in bɛ kalan /en / i n'a fɔ « den » daɲɛ kɔn

Si l'élève ne répond pas correctement, dites: Ayi, siginidenkulu in bɛ kalan /en/ i n'a fɔ « den » daɲɛ kɔnɔ

I y'a faamu wa? An bɛ se ka taa a fɛ ? Ni ne ko "a daminɛ", i kɛtɔ ka siginiden fɛn o fɛn kalan, i b'i bolo da o kan. I b'u kalanni daminɛ numanfɛ ka taa kininfɛ sira ni sira. I y'a faamu kosɛbɛ wa? I bolo da sigiden fɔlɔ kan. I labɛnnen don wa? I b'a lajɛ k'u kalan ka ɲɛ teliya la. A damiɛ!

b	a	u	l	s	o	ɔ	L	u	c
k	d	nj	h	t	e	l	ii	m	ɔ
r	u	c	ns	p	ee	ɛ	n	e	b
n	an	ɛ	L	ɔn	t	M	oo	l	g
nt	o	uu	h	u	d	W	r	g	l
k	nc	s	f	a	n	An	a	k	nf
ɛ	w	on	L	ng	s	np	y	a	s
j	a	ɛn	l	ɛɛ	p	nb	y	a	ɔ
m	ŋ	z	nk	b	e	U	L	d	aa
g	a	r	ŋ	ɔɔ	o	K	un	l	en

Time Remaining

Autostop?

Mots inventés

Daŋɛ dɔw filɛ, lala i ma deli ka minnu ye. Nka ne tun b'a fɛ i k'a lajɛ k'u kalan. Misali la, daŋɛ fɔlɔ in bɛ kalan «gɛ» *[Indiquer le mot* « gɛ» *avec le doigt].* I bɛ se ka segin daŋɛ fɔlɔ in kalanni kan wa ?

[Après sa réponse, ou après 3 secondes dans le cas de non-réponse, montrez-lui comment faire.]

Daŋɛ in dun ? *[indiquer le mot « zii » avec le doigt].* I bɛ se k'o kalan wa ?

[Après sa réponse, ou après 3 secondes dans le cas de non-réponse, montrez-lui comment faire.]

Nin dun ? *[indiquer le mot « hu » avec le doigt].* I bɛ se k'o kalan wa ?

[Après sa réponse, ou après 3 secondes dans le cas de non-réponse, montrez-lui comment faire.]

I y'a faamu wa ? N bɛ min nɔfɛ i y'o faamu wa? Ni ne ko "a daminɛ", i bɛ siraw ta kelen kelen k'u kalan k'a daminɛ numanfɛ ka taa kininfɛ. N'i sera sira dɔ laban na, i b'o nɔkanta daminɛ. I labɛnnen don wa? I b'a lajɛ k'u kalan ka ŋɛ teliya la. A daminɛ!

zi	fe	do	lu	tee
laa	bii	kee	mo	sawa
ki	gibɔ	lezo	fuki	cuto
gamo	Luba	yow	basɔ	pifo
pa	kiwɔ	zaa	yenu	jowe
guu	Mire	maja	dɛca	nsɔ
yɛbu	lina	ŋipɛ	tansa	yonpe
wɛn	Mudo	sipu	poora	ŋasi
zuso	wɛɛ	Loo	lunan	njew
ŋope	Nbeli	luro	pini	leko

Time Remaining

Autostop? ☐

Lecture du texte

Sisan, n b'a fɛ i ka maana in kalan. I b'i kan bɔ kosɛbɛ A lajɛ i k'a kalan ka ɲɛ teliya la; o kɔ ne bɛ ɲininkali dɔw kɛ i la. Ni ne ko i k'a daminɛ, i b'a daminɛ yan *(Mettez la feuille de la Section 5 devant l'élève (F/5). Montrez du doigt le premier mot du passage).* I labɛnna wa ? An k'a daminɛ.

[Faites démarrer le chrono en appuyant sur le bouton START / STOP]

Ami	ye	kalanden	ye.	A
kɔrɔkɛ	tɔgɔ	ye	Musa.	Jiri
caman	b'a	ka	foro	la.
Mangoro,	buyagi	ani	papaye	bɛ
sɔrɔ	yen.	Sugu	don,	a
b'a	ka	wotoro	fa	jiriden
na.	A	kelen	tɛ	se
k'a	ka	jiriden	bɛɛ	feere.
A	kunna	ka	di	Ami
bɛ	yan.	U	bɛ	jiriden
bɛɛ	feere			

Time Remaining

Autostop? ☐

Questions de Compréhension

[Reprendre le texte]

Sisan, i bɛna ɲininkali damadɔ jaabi maana in kan.

1. Ami ye mun ye ? ([kalanden, Musa dɔgɔnin]

☐ Correct ☐ Incorrect ☐ Pas de Réponse

2. Mun bɛ Musa ka foro la ? ([jiri, mangoro, buyagi, lenburu ni papaye])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

3. Musa bɛ jiridenw ta mun na ? ([wotoro])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

4. Jɔn kelen tɛ se ka jiriden bɛɛ feere ? ([Musa])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

5. Ami bɛ mun kɛ ? ([Ami bɛ Musa dɛmɛ, Ami b'a kɔrɔkɛ dɛmɛ.])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Compréhension à l'audition

Sisan, ne bɛna maana kelen kalan i ye siɲɛ kelen. O kɔ, n bɛ ɲininkali damadɔ k'i la maana in kan. I bɛ maana in lamɛn kosɛbɛ. I bɛ tila ka ɲininkaliw jaabi i fɛrɛ ma' I sɔnna wa? N b'a fɛ i ka min kɛ i y'o faamu wa? An k'a daminɛ. A lamɛn kosɛbɛ:

Sitan ka kalanyɔrɔ bɛ sugu kɛrɛfɛ.

A ni a ka so ka jan dɔɔnin.

Dɔgɔkun tile fɔlɔ o tile fɔlɔ, a taatɔ u ka kalanyɔrɔ la, a b'i jɔ ka bonbon san Yakuba ka bitiki kɔnɔ.

Sɔgɔmada in na, Yakuba bɛ nakɔ la.

Sitan selen bitigi la, a y'a sɔrɔ a tun da tugulen don.

A nisɔngoyara kosɛbɛ.

A ka kan ka taa kalanyɔrɔ la.

Sira la a ni Yakuba bɛ bɛn a bɔtɔ nakɔ la .

A bɛ lenburuba di Sitan ma.

A nisɔndiyara.

1. Sitan ka kalanyɔrɔ bɛ min ? ([Sugu kɛrɛfɛ])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

2. Sitan bɛ mun san sira la ? ([Bɔnbɔn])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

3. Yakuba bɛ min Sɔgɔmada in na ? ([Nakɔ la.])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

4. Jɔn ye lenburuba di Sitan ma ? ([Yakuba])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

5. Dɔgɔkun don jumɛn Sitan bɛ bonbon san ? ([Ntɛnɛn.])

☐ Correct ☐ Incorrect ☐ Pas de Réponse

Annex C: Student Questionnaire

Questionnaire d'élève - intervention comparison Part 1

Sisan An bɛ na ɲininkali dɔ w k' i la I yɛrɛ kunkan aw ka so kan an'I kalanjɛ kan.
Maintenant, nous allons vous poser quelques questions à propos de toi, de ta famille, et de la lecture.

1. Kalanyɔrɔ la, karamɔgɔ bɛ bamanankan fɔ i ye wa?
   A l'école, est-ce que ton maitre te parle en Bamanankan?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

2. Kalanyɔrɔ la, i teriew bɛ bamanankan fɔ i ye wa?
   A l'école, est-ce que ton maitre te parle en Bamanankan?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

3. Kalanyɔrɔ la, i ni teriew bɛ bamanankan fɔ wa?
   A l'école est-ce que tu parles avec tes amis en Bamanankan?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

4. So, i ni balimakɛw ni balimamusow bɛ bamanankan fɔ wa?
   A la maison, est-ce que tu parles avec tes frères et soeurs en Bamanakan?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

5. So, i ni balimakɛw ni balimamusow bɛ bamanankan fɔ wa?
   A la maison, est-ce que tu parles avec les adultes en Bamanakan?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

6. Kalanyɔrɔ la, gafew bɛ yen minnu sɛbɛnnen don bamanakan na wa?
   A l'école, y a-t-il des livres, magazines/journaux en Bamanakan?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

Questionnaire d'élève - intervention comparison Part 2

7. Arajo bɛ aw ka so wa? *As-tu une radio à la maison?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

8. Telewizon bɛ aw ka so wa? *As-tu une télévision à la maison?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

9. Telefɔni bɛ aw ka so wa? *Y a-t-il un téléphone/portable à ta maison?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

10. Kuran bɛ aw ka so wa? *As-tu de l'éléctricité/courant à la maison?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

11. Sokɔnɔŋɛgɛn bɛ aw ka so wa? *As-tu une toilette dans ta maison?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

12. Nɛgɛso bɛ aw ka so wa? *Y a-t-il un vélo ou un moto à ta maison?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

13. Bolimafɛn bɛ aw ka so wa?
Y a-t-il un véhicule motorisé à ta maison (un voiture, un camion, un 4x4 ou un tracteur)?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

14. Kunun wulafɛ, e ye wati joli kɛ sokɔnɔbaaraw la ?
Hier soir, combien de temps as-tu passé en faisant des tâches menagères (à la maison)?

☐ Fu ☐ Dɔɔnin ☐ Kosɛbɛ ☐ Jaabi ma di

15. I ba bɛ kalan kɛ bamanankan na wa? *Ta mère, lis-t-elle en Bamanankan?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

16. I fa bɛ kalan kɛ bamanankan na wa? *Ton père, lis-t-il en Bamanankan?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

17. Mɔgɔ bɛ aw ka so, min bɛ nsiirinw kalan i ye wa?
Est-ce qu'il y a quelqu'un chez toi (parent, frère/soeur, grandparent) qui te lis des contes ou des histoires?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

18. Mɔgɔ bɛ aw ka so, min b'i dɛmɛ k'i ka kalansodegeliw kɛ wa?
Est-ce qu'il y a quelqu'un chez toi (parent, frère/soeur, grandparent) qui t'aide avec tes devoirs scolaries?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

19. Siɲɛ joli, i ni ka sokɔnɔmɔgɔ dɔ bɛ kalanjɛ kɛ? *A quelle fréquence lis-tu avec un membre de ta famille chez toi?*

☐ Don o don ☐ Tuma ni tuma ☐ Abada ☐ Jaabi ma di

20. Aw ka so kɔnɔ, jon de b'i dɛmɛ tuma ni tuma kalanjɛ la walima k'i ka degeliw kɛ?
Chez toi, qui t'aide le plus souvent avec la lecture ou tes devoirs?

☐ Mansa kelen walima fila ☐ Balimakɛ walima balimamuso ☐ Dukɔnɔmɔgɔwɛrɛ (mɔkɛ, mɔmuso, tɛnɛnmuso walima benkɛ)

☐ N tɛ dɛmɛ sɔrɔ kalanjɛ ni delgeliw la so ☐ Jaabi ma di

21. Kalanjɛ kɛli ka di e ye so wa? *Aimes-tu lire chez toi?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

22. Kalanjɛ kɛli ka di e ye Kalanyɔrɔ la wa? *Aimes-tu lire à l'école?*

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

23. Kalanyɔrɔ la, e kelen bɛ gafew kalan siɲɛ joli?
A l'école, à quelle fréquence lis-tu des livres silencieusement tout(e) seul(e)?

☐ Don o don ☐ Tuma ni tuma ☐ Abada ☐ Jaabi ma di

24. Kalanyɔrɔ la, karamɔgɔ bɛ ɲinikaliw k'i la i ka gafe kalannenw kan wa?
A l'école, est-ce que ton maitre te pose des questions concernant ce que tu lis?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

25. N'i ma se ka yɔrɔ minnu kalan, karamɔgɔ b'i dɛmɛ k'olu kalan wa?
Ton maitre, est-ce qu'il t'aide si tu ne peux pas lire quelque chose?

☐ Ɔwɔ ☐ Ayi ☐ N t'a dɔni ☐ Jaabi ma di

26. Karamɔgɔ b'i bila sɛbɛnni na siɲɛ joli kalanyɔrɔ la?
A quelle fréquence ton maitre te demande d'écrire à l'ecole?

☐ Don o don ☐ Tuma ni tuma ☐ Abada ☐ Jaabi ma di

Annex D: Student Questionnaire Results and Composites

Table D.1: **Language Exposure Composite**

Composite Questions	Response Options	Intervention Group					
		Intervention A		Intervention B		Comparison	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
At school, does your teacher speak to you in Bamanankan?	No	2	1.0	4	1.6	4	4.5
	Yes	195	99.0	247	98.4	84	95.5
	Don't Know	0	0.0	0	0.0	0	0.0
At school, do your friends speak to you in Bamanankan?	No	2	1.0	2	0.8	4	4.4
	Yes	194	99.0	249	99.2	86	95.6
	Don't Know	0	0.0	0	0.0	0	0.0
At school, do you speak to your friends in Bamanankan?	No	2	1.0	1	0.4	3	3.4
	Yes	194	99.0	249	99.6	86	96.6
	Don't Know	0	0.0	0	0.0	0	0.0
At home, do you speak to your siblings in Bamanankan?	No	3	1.5	3	1.2	2	2.2
	Yes	192	98.5	247	98.4	87	97.8
	Don't Know	0	0.0	1	0.4	0	0.0
At home, do you speak to the adults in your home in Bamanankan?	No	8	4.1	6	2.4	1	1.1
	Yes	188	95.9	242	97.2	87	97.8
	Don't Know	0	0.0	1	0.4	1	1.1
At school, are there books, magazines, or newspapers in Bamanankan?	No	6	3.2	13	5.2	19	21.8
	Yes	180	95.2	235	94.8	68	78.2
	Don't Know	3	1.6	0	0.0	0	0.0

Table D.2: **Socioeconomic Status Composite**

Composite Questions	Response Options	Intervention Group					
		Intervention A		Intervention B		Comparison	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
At your house, do you have a radio?	No	15	7.6	10	4.0	1	1.1
	Yes	182	92.4	238	96.0	89	98.9
	Don't Know	0	0.0	0	0.0	0	0.0
At your house, do you have a tv?	No	42	21.3	54	21.8	27	30.0
	Yes	155	78.7	194	78.2	63	70.0
	Don't Know	0	0.0	0	0.0	0	0.0
At your house, do you have a telephone or mobile phone?	No	2	1.0	0	0.0	1	1.1
	Yes	195	99.0	248	100.0	89	98.9
	Don't Know	0	0.0	0	0.0	0	0.0
At your house, do you have electricity?	No	8	4.1	9	3.6	1	1.1
	Yes	189	95.9	239	96.4	89	98.9
	Don't Know	0	0.0	0	0.0	0	0.0
Do you have a toilet inside your house?	No	12	6.1	8	3.2	2	2.2
	Yes	185	93.9	239	96.8	88	97.8
	Don't Know	0	0.0	0	0.0	0	0.0
At your house, do you have a bicycle or motorcycle?	No	13	6.6	10	4.0	2	2.2
	Yes	184	93.4	238	96.0	88	97.8
	Don't Know	0	0.0	0	0.0	0	0.0
At your house, do you have four-wheeler (a car, truck, 4x4, or tractor)?	No	46	23.5	69	27.8	30	33.3
	Yes	150	76.5	179	72.2	60	66.7
	Don't Know	0	0.0	0	0.0	0	0.0
Last night, how much time did you spend on household chores?	None	56	29.0	66	27.3	32	36.0
	Some	81	42.0	91	37.6	28	31.5
	A lot	56	29.0	85	35.1	29	32.6
Can your mother read in Bamanankan?	No	48	24.5	87	35.1	31	34.4
	Yes	147	75.0	161	64.9	59	65.6
	Don't Know	1	0.5	0	0.0	0	0.0
Can your father read in Bamanankan?	No	42	21.4	63	25.4	20	22.2
	Yes	152	77.6	182	73.4	69	76.7
	Don't Know	2	1.0	3	1.2	1	1.1

Table D.3: **Family Reading Support Composite**

Composite Questions	Response Options	Intervention Group					
		Intervention A		Intervention B		Comparison	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Does someone from home (parent, sibling, or grandparent) read stories to you?	No	74	37.6	87	35.2	33	37.5
	Yes	122	61.9	160	64.8	55	62.5
	Don't Know	1	0.5	0	0.0	0	0.0
Does someone from home (parent, sibling, or grandparent) help you with your school work?	No	22	11.2	30	12.0	22	24.4
	Yes	175	88.8	221	88.0	68	75.6
	Don't Know	0	0.0	0	0.0	0	0.0
At home, how often do you read with a family member?	Never	16	8.2	17	7.0	20	23.0
	Sometimes	128	65.3	152	62.6	47	54.0
	Every day	52	26.5	74	30.5	20	23.0
At home, who most often helps you with your reading or homework?	I usually do not get help with reading or homework at home	18	9.2	27	10.8	21	23.6
	One or both parents	23	11.8	37	14.9	18	20.2
	Brother or sister	150	76.9	174	69.9	50	56.2
	Other family member (grandparents, aunts, or uncles)	4	2.1	11	4.4	0	0.0

Table D.4: **Teacher Reading Support Composite**

Composite Questions	Response Options	Intervention Group					
		Intervention A		Intervention B		Comparison	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
At school, does your teacher ask you questions about what you are reading?	No	14	7.2	15	6.0	8	9.9
	Yes	181	92.8	235	93.3	67	82.7
	Don't Know	0	0.0	2	0.8	6	7.4
Does teacher help you when you are unable to read something?	No	9	4.5	7	2.8	5	5.9
	Yes	189	95.5	245	97.2	72	84.7
	Don't Know	0	0.0	0	0.0	8	9.4
How often does your teacher asks you to write in school?	Never	1	0.5	1	0.4	5	6.1
	Sometimes	95	48.2	105	42.7	43	52.4
	Every day	101	51.3	140	56.9	34	41.5

Table D.5: **Disposition to Reading Composite**

Composite Questions	Response Options	Intervention Group					
		Intervention A		Intervention B		Comparison	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Do you like reading at home?	No	1	0.5	1	0.4	2	2.2
	Yes	197	99.5	251	99.6	87	97.8
	Don't Know	0	0.0	0	0.0	0	0.0
Do you like reading at school?	No	0	0.0	1	0.4	0	0.0
	Yes	198	100.0	251	99.6	90	100.0
	Don't Know	0	0.0	0	0.0	0	0.0
At school, how often do you read books quietly by yourself?	Never	10	5.3	22	9.0	21	25.0
	Sometimes	131	68.9	151	61.9	44	52.4
	Every day	49	25.8	71	29.1	19	22.6

Annex E: EGRA Descriptive Statistics and Additional Tables

Table E.1: **Orientation to Print (Correct out of Six)**

Intervention Group	N/n	Baseline			Endline		
		Mean	SD	Zero Score (%)	Mean	SD	Zero Score (%)
Intervention A	198	1.1	1.1	36.4	4.3	1.3	2.0
Intervention B	252	1.1	1.1	36.5	4.1	1.4	2.4
Comparison	90	1.1	1	33.3	2.9	1.5	7.8
Total: All students	**540**	**1.1**	**1.1**	**35.9**	**4**	**1.4**	**3.1**

Table E.2: **Initial Sound Identification (Correct out of Ten)**

Intervention Group	N/n	Baseline			Endline		
		Mean	SD	Zero Score (%)	Mean	SD	Zero Score (%)
Intervention A	198	0.2	0.6	84.3	7.3	2.9	2.5
Intervention B	252	0.3	0.7	81.3	6.3	3.4	2.0
Comparison	90	0.3	1.1	82.2	1.3	1.7	37.8
Total: All students	**540**	**0.3**	**0.8**	**82.6**	**5.9**	**3.6**	**8.1**

Table E.3: **Letter Sound Identification (CLSPM)**

Intervention Group	N/n	Baseline			Endline		
		Mean	SD	Zero Score (%)	Mean	SD	Zero Score (%)
Intervention A	198	0.5	1.5	77.3	23.3	10.6	0.5
Intervention B	252	0.7	2.0	79.0	22.6	8.7	0.8
Comparison	90	0.2	0.9	90.0	8.3	8.3	21.1
Total: All students	**540**	**0.6**	**1.7**	**80.2**	**20.5**	**10.8**	**4.1**

Table E.4: **Nonword Reading (CNWPM)**

Intervention Group	N/n	Baseline			Endline		
		Mean	SD	Zero Score (%)	Mean	SD	Zero Score (%)
Intervention A	198	0.0	0.2	98.5	8.8	8.4	25.8
Intervention B	252	0.0	0.2	97.6	8.6	7.6	21.4
Comparison	90	0.0	0.2	98.9	0.9	2.2	76.7%
Total: All students	**540**	**0.0**	**0.2**	**98.1**	**7.4**	**7.9**	**32.2**

Table E.5: **ORF (CWPM)**

Intervention Group	N/n	Baseline			Endline		
		Mean	SD	Zero Score (%)	Mean	SD	Zero Score (%)
Intervention A	198	0.0	0.0	100.0	10.1	11.3	24.2
Intervention B	252	0.0	0.0	100.0	9.9	9.6	19.4
Comparison	90	0.0	0.0	100.0	1.0	2.3	66.7
Total: All students	**540**	**0.0**	**0.0**	**100.0**	**8.5**	**10.1**	**29.1**

Table E.6: **Reading Comprehension (Correct out of Five)**

Intervention Group	N/n	Baseline			Endline		
		Mean	SD	Zero Score (%)	Mean	SD	Zero Score (%)
Intervention A	198	0.0	0.0	100.0	0.6	1.0	70.2
Intervention B	252	0.0	0.0	100.0	0.5	0.9	69.4
Comparison	90	0.0	0.0	100.0	0.0	0.1	97.8
Total: All students	**540**	**0.0**	**0.0**	**100.0**	**0.4**	**0.9**	**74.4**

Table E.7: **Listening Comprehension (Correct out of Five)**

Intervention Group	N/n	Baseline			Endline		
		Mean	SD	Zero Score (%)	Mean	SD	Zero Score (%)
Intervention A	198	1.1	1.3	51.0	3.5	1.1	1.0
Intervention B	252	1.0	1.2	46.0	3.5	1.1	1.6
Comparison	90	1.1	1.4	45.6	3.0	1.3	4.4
Total: All students	**540**	**1.0**	**1.3**	**47.8**	**3.4**	**1.1**	**1.9**

Table E.8: **Mean Scores by Intervention Group and Gender on EGRA Instrument 1 at Baseline**

Composite Questions	Intervention A		Intervention B		Comparison	
	Girls	Boys	Girls	Boys	Girls	Boys
	n=91	n=107	n=113	n=139	n=41	n=49
Orientation to print (correct out of six)	1.0	1.2	1.0	1.2	1.2	1.0
Initial sound identification (correct out of ten)	0.2	0.2	0.3	0.3	0.5	0.2
Letter sound identification (CLSPM)	0.6	0.5	0.7	0.7	0.3	0.1
Nonword reading (CNWPM)	0.0	0.0	0.0	0.0	0.0	0.0
Oral reading fluency (CWPM)	0.0	0.0	0.0	0.0	0.0	0.0
Reading comprehension (correct out of five)	0.0	0.0	0.0	0.0	0.0	0.0
Listening comprehension (correct out of five)	1.1	1.0	1.0	1.0	1.2	1.1

Table E.9: **Mean Scores by Intervention Group and Gender at EGRA Instrument 2 at Endline**

Composite Questions	Intervention A		Intervention B		Comparison	
	Girls	Boys	Girls	Boys	Girls	Boys
	n=95	n=103	n=123	n=129	n=36	n=54
Orientation to print (correct out of six)	4.1	4.5	4.2	4.0	3.0	2.9
Initial sound identification (correct out of ten)	7.3	7.3	6.7	6.0	1.1	1.5
Letter sound identification (CLSPM)	22.7	23.8	23.9	21.4	6.6	9.4
Nonword reading (CNWPM)	8.7	9.0	9.3	7.9	0.6	1.1
Oral reading fluency (CWPM)	9.4	10.6	11.1	8.8	0.9	1.0
Reading comprehension (correct out of five)	0.5	0.6	0.6	0.4	0.0	0.0
Listening comprehension (correct out of five)	3.4	3.6	3.6	3.4	3.2	2.9

Table E.10: **Zero Scores by Intervention Group and Gender at Endline**

Composite Questions	Intervention A		Intervention B		Comparison	
	Girls	Boys	Girls	Boys	Girls	Boys
	n=95	n=103	n=123	n=129	n=36	n=54
Orientation to print (correct out of six)	4.2	0.0	3.3	1.6	13.9	3.7
Initial sound identification (correct out of ten)	3.2	1.9	0.8	3.1	52.8	27.8
Letter sound identification (CLSPM)	1.1	0.0	0.0	1.6	22.2	20.4
Nonword reading (CNWPM)	26.3	25.2	16.3	26.4	75.0	77.8
Oral reading fluency (CWPM)	23.2	25.2	14.6	24.0	72.2	63.0
Reading comprehension (correct out of five)	69.5	70.9	65.0	73.6	97.2	98.1
Listening comprehension (correct out of five)	2.1	0.0	1.6	1.6	2.8	5.6

Annex F: Correlation Analysis Results

Table F.1: **Correlation Table—Intervention A**

Composite Category	Orientation to Print (correct out of six)	Initial Sound Identification (correct out of ten)	Letter Sound Identification (CLSPM)	Nonword Reading (CNWPM)	Oral Reading Fluency (CWPM)	Reading Comprehension (correct out of five)	Listening Comprehension (correct out of five)
Language exposure	-0.023	0.131	.158*	0.136	0.077	0.069	-0.024
Socioeconomic status	-0.009	0.118	0.001	-0.029	-0.112	-0.105	-0.104
Family reading support	-0.089	0.020	-0.037	-0.128	-0.100	-0.086	-.145*
Teacher reading support	0.093	.158*	0.106	0.106	0.101	0.095	-0.023
Disposition to reading	0.027	.181*	.149*	0.136	.148*	0.132	-0.024

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Table F.2: **Correlation Table—Intervention B**

Composite Category	Orientation to Print (correct out of six)	Initial Sound Identification (correct out of ten)	Letter Sound Identification (CLSPM)	Nonword Reading (CNWPM)	Oral Reading Fluency (CWPM)	Reading Comprehension (correct out of five)	Listening Comprehension (correct out of five)
Language exposure	.193**	0.105	.171**	.146*	0.085	0.023	.220**
Socioeconomic status	.148*	.141*	0.008	0.086	0.095	0.076	0.062
Family reading support	0.069	0.069	0.048	0.089	0.060	0.110	-0.040
Teacher reading support	0.018	0.084	0.077	0.089	0.101	0.021	0.031
Disposition to reading	-0.065	0.028	-0.030	0.045	0.091	0.093	0.028

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Table F.3: **Table—Comparison**

Composite Category	Orientation to Print (correct out of six)	Initial Sound Identification (correct out of ten)	Letter Sound Identification (CLSPM)	Nonword Reading (CNWPM)	Oral Reading Fluency (CWPM)	Reading Comprehension (correct out of five)	Listening Comprehension (correct out of five)
Language exposure	0.061	0.178	.351**	.233*	0.188	-0.051	0.129
Socioeconomic status	-0.011	0.053	.261*	0.171	0.152	-0.116	-0.058
Family reading support	0.007	0.211	.231*	0.178	0.185	-0.127	0.063
Teacher reading support	-.259*	0.150	0.125	0.041	0.032	-0.084	-0.067
Disposition to reading	-0.201	0.179	0.083	0.092	0.098	-0.096	-0.064

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

Annex G: EGRA Reliability Results

Table G.1: **Reliability Results for EGRA Instrument 1 at Baseline**

Subtask	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Orientation to print (correct out of six)	0.223	0.176
Initial sound identification (correct out of ten)	0.254	0.221
Letter sound identification (CLSPM)	0.338	0.279
Nonword reading (CNWPM)	0.081	0.304
Oral reading fluency (CWPM)	0.000	0.306
Reading comprehension (correct out of five)	0.000	0.306
Listening comprehension (correct out of five)	0.279	0.148
	EGRA Coefficient Alpha	**0.297**

Table G.2: **Reliability Results for EGRA Instrument 2 at Endline**

Subtask	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Orientation to print (correct out of six)	0.494	0.796
Initial sound identification (correct out of ten)	0.617	0.800
Letter sound identification (CLSPM)	0.740	0.784
Nonword reading (CNWPM)	0.783	0.760
Oral reading fluency (CWPM)	0.740	0.756
Reading comprehension (correct out of five)	0.628	0.776
Listening comprehension (correct out of five)	0.308	0.826
	EGRA Coefficient Alpha	**0.810**

Annex H: Sample Comparison (Total Sample v. Matched Sample)

Table H.1: **Sample by Gender and Group**

Description		Total Sample	Matched Sample
Gender at baseline	Girls	245	-
	Boys	295	-
Gender at endline*	Girls	254	173
	Boys	286	199
Group at endline	Intervention A	198	148
	Intervention B	252	173
	Comparison	90	51
Total: All students		**540**	**372**

* Not all replacement students at endline replaced a student of the same gender from baseline.

Table H.2: **Total Sample v. Matched Sample – EGRA Instrument 1 at Baseline Results Comparison**

Intervention Group	Matched Sample (n=372)			Total Sample (n=540)		
	Intervention A	Intervention B	Comparison	Intervention A	Intervention B	Comparison
Orientation to print (correct out of six)	1.3	1.2	1.0	1.1	1.1	1.1
Initial sound identification (correct out of ten)	0.2	0.3	0.4	0.2	0.3	0.3
Letter sound knowledge (CLSPM)	0.7	0.8	0.2	0.5	0.7	0.2
Nonword reading (CNWPM)	0.0	0.0	0.0	0.0	0.0	0.0
Oral reading fluency (CWPM)	0.0	0.0	0.0	0.0	0.0	0.0
Reading comprehension (correct out of five)	0.0	0.0	0.0	0.0	0.0	0.0
Listening comprehension (correct out of five)	1.1	1.1	0.8	1.1	1.0	1.1

Table H.3: **Total Sample v. Matched Sample – EGRA Instrument 2 at Endline Results Comparison**

Intervention Group	Matched Sample (n=372)			Total Sample (n=540)		
	Intervention A	Intervention B	Comparison	Intervention A	Intervention B	Comparison
Orientation to print (correct out of six)	4.3	4.0	2.8	4.3	4.1	2.9
Initial sound identification (correct out of ten)	7.4	6.4	1.4	7.3	6.3	1.3
Letter sound knowledge (CLSPM)	23.1	21.8	8.8	23.3	22.6	8.3
Nonword reading (CNWPM)	8.5	8.3	1.1	8.8	8.6	0.9
Oral reading fluency (CWPM)	9.9	9.8	1.0	10.1	9.9	1.0
Reading comprehension (correct out of five)	0.6	0.5	0.0	0.6	0.5	0.0
Listening comprehension (correct out of five)	3.5	3.5	2.9	3.5	3.5	3.0