Togo McGovern-Dole International Food for Education and Child Nutrition Program - STARS

Final Evaluation

5/13/2025

STARS Final Evaluation Report

Program: McGovern-Dole International Food for Education and Child Nutrition
Agreement Number: USDA Award No. FFE-693-2019/016-00
Funding Year: Fiscal Year 2022–2023
Project Duration: 2019–2025
Implemented by: Catholic Relief Services – United States Conference of Catholic Bishops (CRS)

Evaluation Authored by: School-to-School International
Parnika Bhatia, Drew Schmenner, Jarret Guajardo

Cover photo by: Catholic Relief Services/Staff

Table of Contents

List of Figures

List of Tables

List of Acronyms

AAM	Assessor Accuracy Measure
BMI	Body Mass Index
CE1	Cours élémentaire première année, or Grade 3
CNWPM	Correct Nonwords per Minute
CP2	Cours préparatoire deuxième année, or Grade 2
CRS	Catholic Relief Services – United States Conference of Catholic Bishops
CWPM	Correct Words per Minute
EGRA	Early Grade Reading Assessment
FY	Fiscal Year
IHfRA	Innovative Hub for Research in Africa
IYCF	Infant and Young Child Feeding assessment
MAD	Minimum Acceptable Diet
MDD	Minimum Dietary Diversity
McGovern-Dole	McGovern-Dole International Food for Education and Child Nutrition Program
PMP	Performance Monitoring Plan
SEDL	Southwest Educational Development Laboratory
SO	Strategic Objective
SOW	Statement of Work
STARS	Santé, Transformation et Apprentissage pour une Réussite Scolaire
STS	School-to-School International
ToC	Theory of Change
UNICEF	United Nations Children's Fund
USDA	United States Department of Agriculture
WHO	World Health Organization

Executive Summary

Project Background and Purpose

Catholic Relief Services (CRS) implemented a McGovern-Dole International Food for Education and Child Nutrition (McGovern-Dole) Program, *Santé, Transformation et Apprentissage pour une Réussite Scolaire* (STARS), in Togo. Funded by the United States Department of Agriculture (USDA), the project aims to improve literacy and primary education in Togo's Savanes and Kara regions by reducing student hunger.[1] It is designed to achieve these goals by providing school meals, training teachers and school administrators, improving water and sanitation facilities, providing school infrastructure, and building skills and knowledge.

CRS began implementing the STARS project activities in fiscal year (FY) 2020.[2] STARS aimed to reach 36,341 primary school students at 138 schools in its first year and expand to 46,925 students by FY24, totaling 71,248 students for the project's life due to anticipated enrollment increases. The objectives of STARS align with the standard strategic objectives (SO) of the McGovern-Dole Program:

- SO 1: Improved literacy of school-aged children
- SO 2: Increased use of health and dietary practices of school-aged children

This report presents the findings of the STARS endline evaluation, which is a follow-up to the baseline and midterm evaluations. The evaluation establishes endline values for all performance indicators, generates data for comparative analysis, and validates project strategies and assumptions.

Evaluation Design

The external evaluation of STARS was conducted intermittently over five years. Baseline data collection for the evaluation occurred in November 2020, midterm data collection in November 2022, and endline data collection in 2024. School-to-School International (STS) was contracted as the external evaluator to undertake the baseline, midterm, and endline evaluation of the STARS project. A regional data collection firm, Innovative Hub for Research in Africa (IHfRA), managed the fieldwork. CRS reviewed and helped refine all tools before use.

The evaluation at all three timepoints used a mixed-methods approach, including quantitative and qualitative data collection tools. Quantitative tools included an Early Grade Reading Assessment (EGRA) composed of five subtasks, student survey, head teacher survey, parent survey, and classroom observation. Quantitative data were collected from 80 schools where the project is intervened. IHfRA enumerators administered the EGRA and student survey to 20 randomly selected students—10 boys and 10 girls— enrolled in grade 3 at each school.[3] Enumerators then collected additional data at each site, including a survey with the school's head teacher, a parent survey with three parents of students who also had a child younger than two; and a school and classroom observations tool. Qualitative tools included a focus group discussion (FGD) and key informant

[1] In English: "Health, Transformation and Learning for School Success"
[2] CRS received approval from USDA to begin some activities prior to the submission of the baseline report due to lengthy delays in data collection resulting from the global Covid-19 pandemic.
[3] There were cases where there were less than 16 students available at the school. In this case, all available students were sampled. The following schools had less than 20 students: EPP DJABONLI, EPP KOUTEOU, EPP MONDOFOALI, EPP NANDJONKARGOU, EPP DAKALFAM, EPP DJABIGNON, EPP SANLOAGA, EPP KOUTEGOU, EPP DJANTCHOGOU, and EPP DJANKPENTENE. In schools with more than 20 students, a random number generator was used to select students.

interviews (KIIs). Two teacher FGDs and two COGEP FGDs were conducted in each prefecture in the selected schools; five FGDs with *Mamans lumières* were conducted. KIIs were conducted with school directors, mayors, MoE pedagogical advisors or inspectors, and local partners.

The evaluation sought to measure outcomes of the two standard SOs of the McGovern-Dole Program—1.) improved literacy of school-aged children, and 2) increased use of health and dietary practices of school-aged children—as well as two broader McGovern-Dole research questions:

- What systems of community health care governance are the most effective at sustaining the delivery of health interventions through school meal programs?
- What are the differences in educational outcomes from school meal programs between malnourished or undernourished children and those who are not?

Strategic Objectives: Findings and Conclusions

The project's largest impact can be seen in reading skills such as letter sound and initial sound identification, the best practices used by school personnel, and the improvement of school sanitation facilities.

Findings and Conclusions Related to SO 1: Improved Literacy of School-aged Children

Findings Related to Improvements in Literacy Skills for All Students

- **Zero Scores for all Students:** The percentage of girls and boys scoring zero—not answering a single item correctly on a subtask—significantly decreased from baseline to endline on three subtasks: Letter Sound Identification (down from 38.0 percent to 19.0 percent), Initial Sound Identification (72.0 percent to 57.0 percent), and Oral Reading Fluency subtasks (71.0 percent to 56.0 percent). Despite these gains, the proportion of students scoring zero on subtasks remained high.
- **Mean Scores for all Students:** Mean scores significantly improved from baseline to endline for both girls and boys on two subtasks: Initial Sound Identification and Letter Sound Identification. On average, students correctly responded to about two out of the 10 items on the Initial Sound Identification subtask, a statistically significant increase from baseline (1.4 out of ten items) but still ultimately low. On the Letter Sound Identification subtask, students correctly identified 9.2 letters out of 100 on average, a significant increase from baseline.
- **Thresholds for Reading Comprehension**: The predetermined threshold for Reading Comprehension was three correct out of five questions. Although most students (99.9 percent) did not meet this threshold at endline, approximately 4.9 percent were able to answer at least one question correctly.
- **Effective School Supports:** Respondents point to four complementary factors driving the observed improvements in children's reading skills. Supplying each pupil with their own

Highlight

The literacy findings suggest that **the STARS project had the greatest impact on lower-level literacy skills and are beginning to show effects on the higher-order skill of word reading.** **B**oth boys and girls were significantly less likely to receive zero scores—to not answer a single item correctly on a subtask—on the Letter Sound Identification, Initial Sound Identification, and Oral Reading Fluency subtasks.

leveled reading books and workbooks transformed learning – with printed texts in hand, children could practice daily, and teachers could track progress one-on-one. The Ministry of Education and inspection teams also began tying teachers' application of new methods to their school's performance rating, an accountability mechanism that encouraged teachers to embed new literacy pedagogies in every class. Respondents also credit regular school performance competitions, with prizes for the top two readers, fostering friendly competition and motivating pupils. Finally, participants cited the structured and sequential approach to instructional modules, with practice lessons and peer feedback, as effective**.**

Conclusion: Students showed statistically significant improvements in foundational literacy and fluency outcomes on many measures since baseline, which FGD and KII respondents attribute to the provision of dedicated reading materials, enhanced teaching training with supervised follow-up, reading competitions, and structured, hands-on instructional modules. Despite statistically significant improvement in some areas, overall literacy is still low.

Highlight

The vast majority of students (99.9 percent) did not meet the STARS project's benchmark for Reading Comprehension, which was to answer three of five reading comprehension questions correctly.

Findings Related to Gender Disparities in Improvements in Literacy Skills

- **Mean Scores for Boys**: The scores for boys but not girls improved on the Oral Reading Fluency subtask. At the endline, students read 2.4 words out of 57 on average. While this increased from baseline, the difference was not statistically significant for students overall. Only boys—who read 2.9 words on average at endline—showed a statistically significant increase on the Oral Reading Fluency subtask from baseline.
- **Accuracy Scores for Boys**: Boys scored significantly higher accuracy scores than girls did on three subtasks: Initial Sound Identification (0.6 percentage points more accurate), Letter Sound Identification (0.5 percentage points more accurate), and Oral Reading Fluency (0.1 percentage points more accurate). Girls scored higher than boys on the two remaining subtasks—Reading Comprehension and Listening Comprehension—but neither was statistically significant.
- **Zero Scores for Boys**: Boys received fewer zero scores on all five subtasks than girls. Differences ranged from a 3.0-percentage point difference on the Initial Sound Identification and Listening Comprehension subtasks to a 10.0-percentage point difference on the Oral Reading Fluency subtask.

Conclusion: Gendered differences in performance were seen in literacy outcomes at the endline evaluation, as boys demonstrated greater foundation skills than girls.

Findings and Conclusions Related to SO 2: Increased Use of Health and Dietary Practices of School-aged Children

- **Improved Sanitation Facilities**: Sanitation facilities at the 80 schools sampled in the endline evaluation significantly improved from the baseline. The number of schools with no toilets available decreased from 29 at baseline to 15 at endline. Meanwhile, the number of schools with composting toilets increased from 10 at baseline to 26 at endline.

Highlight

School sanitation facilities improved over the life of the project. Specifically, the project increased the number of schools with toilets (14 added), increased the number of schools with composting toilets (16 added), and ensured all schools (100.0 percent) had functional water sources.

Conclusion: The project saw success in the improvement of school sanitation facilities.

- **Functional Waters Sources:** Of the 54 schools with water sources observed at the endline, 100 percent were functioning. This is a continued improvement from baseline and midterm, when 75.0 percent and 95.9 percent of sources were functioning, respectively.
- **Handwashing Promoted**: Qualitative data collection respondents reported that the project successfully promoted the habit of washing hands, especially once latrines were introduced in schools that previously lacked them. However, respondents felt that handwashing remained insufficient because some schools still lacked access to clean, running water. One woman in a focus group added that there are still far too few handwashing stations at her school.

Conclusion: The functionality of water sources significantly improved from baseline to endline; however, more work is possible for their ongoing development.

Significant Findings and Conclusions Related to Expected Outcomes

Findings related to Expected Outcome 1: Improve literacy outcomes by strengthening school systems and community support

Highlight

When the project began, teachers were not consistently using quality teaching practices for literacy lessons – however, by the end, **more than 70 percent of teachers used a majority of the quality literacy practices** identified by the STARS project in each lesson.

- **Homework**: One-third (33.0 percent) of parents reported helping their children with homework the week before the survey, a statistically insignificant change from baseline.
- **Educational Activities at Home**: More than half (54.3 percent) of parents reported participating in three or more educational activities with their children at home, a statistically insignificant change from baseline.

Conclusion: Parent's self-reported engagement in their children's education at home did not improve.

- **Less Representation**: More than 43.0 percent of students reported that their teachers never or rarely told positive stories about female characters, compared to 41.7 percent for male characters—indicating a slight disparity in how girls are represented in classroom narratives.
- **More Chores:** A plurality of students (40.0 percent) reported that girls clean school toilets or latrines "most of the time" or "always," compared to just 28 percent for boys.

Conclusion: Gender norms and practices in schools may contribute to unequal learning environments.

Expected outcome 2: Improve the quality of literacy instruction by building the capacity of teachers and administrators and providing sufficient literacy materials

- **Quality Teaching Practices**: The proportion of teachers observed during the classroom observation demonstrating at least five of nine "quality teaching practices" during a lesson improved significantly from baseline (0.0 percent) to endline (70.6 percent).
- **Supervision Tools in Use**: The number of quality supervision tools used at schools increased from baseline to endline. At baseline, 41.8 percent of school officials used five or fewer tools, while at endline 62.2 percent used eight or more.

Conclusion: Teachers' and schools' use of quality instructional practices and materials improved from baseline to endline.

- **Textbook Access**: Only 18.8 percent of classrooms had a one-to-one ratio of French textbooks to students, while more than 67.0 percent had a ratio worse than one-to-two, severely limiting students' opportunities for independent reading and engagement with written French.
- **Student Fluency**: Just 0.9 percent of students reported speaking French very well, while 95 percent said they could at most say a few things or have a basic conversation—suggesting low oral language proficiency at the endline.
- **Home Language Use**: Only 1.0 percent of students reported speaking French at home; the vast majority use local languages like Konkomba, Ngam-gam, and Gourma.

Conclusion: Evidence from classroom observations and survey data reveals systemic limitations in both the availability of materials and the linguistic environment:

Expected outcome 3: Improve student attentiveness and attendance by providing daily school lunches and ensuring a safe school environment

- **Attentive Students**: Classroom observations at endline revealed that 69.1 percent of students were attentive compared to 59.7 percent at baseline. However, this represents only a marginally insignificant improvement ($p < 0.1$) and fails to reach the life of the project target of 75 percent.
- **Higher Attendance**: According to survey data at endline, an average of 89.9 percent of students were present on the day of the evaluation, a statistically significant increase from 84.7 percent at baseline. This surpasses the STARS project target of 89.2 percent.
- **Less Illness**: At the endline, only 10.2 percent of parents stated that at least one of their children missed school in the past month due to illness, a statistically significant decrease from baseline (14.9 percent).

Conclusion: School-age children missed fewer days of school and were more attentive when they were in school, possibly as a result of providing daily school lunches and ensuring a safe school environment.

Expected outcome 4: Improve health and dietary practices of targeted beneficiaries by increasing awareness of nutrition, health, and hygiene behaviors combined with water and sanitation infrastructure improvements

- **Minimum Acceptable Diet**: At the endline, 12.4 percent of parents had children 6–23 months old who met the minimum acceptable diet (MAD) threshold, a significant decrease compared to baseline (20.1 percent).

Conclusion: The nutrition of young children did not improve, a goal outlined in Sub IR 1.2.1.

Research Question 1: Findings and Conclusions:

What are the differences in educational outcomes from school meal programs between malnourished or undernourished children and those who are not?

- **Students are Not Underweight**: Body mass index (BMI) scores for students measured as part of a special study at the endline show that 88.8 percent of students were not considered underweight, with only 12.2 percent of students classified at or below the fifth percentile for their age.
- **No Correlation to Literacy Outcomes**: A special study introduced at midterm found that BMI did not correlate with literacy outcomes for those students sampled then or at the endline.

Conclusion: Healthy weights are not yet contributing to greater learning outcomes.

Research Question 2: Findings and Conclusions:

What systems of community health care governance are the most effective at sustaining the delivery of health interventions through school meal programs?

- **Stakeholders' and volunteers' willingness to participate without financial expectations was a key strength**: Community members and project volunteers continued to engage in project activities "without receiving anything in return." As a result, "in the community, without incentives, the work will be done. That's what was done with those lumières." Similarly, respondents reported alignment with an existing government decree that envisioned these committees as unpaid, "essential for the functioning of schools," and that sustaining this approach would ensure permanence beyond the project's lifespan:"...if we align ourselves behind this strategy [...], it could be good, and it must be sustainable." Finally, in a direct question on readiness to work without reward, a COJEP member unambiguously stated: "Are you ready to work without expecting anything in return? Yes, for the good of the whole community."
- **Teachers and Head Teachers have shown themselves as effective actors in their roles as educators**: At the endline, they were observed in higher numbers to be using quality teaching practices and supervision tools.

Sustainability:

- **Beneficiary groups have appropriated activities.** As one community member phrased it, they have so fully taken up these activities that they "continue, and will continue until 1000 years." Likewise, the "mères lumière" trained in nutrition, hygiene and child-feeding practices continue to meet and sensitize their peers autonomously: "[...] despite the fact that the project has come to an end, there are a good number of women who continue to raise awareness among their peers [...] so that these good practices in the community can be promoted and kept up every day."
- **Saving-and-credit groups bolstered financial resilience.** By mobilizing internal loans and supporting local education and health expenditures, these groups help people save money

to meet their children's needs at school and for health care. These peer-to-peer mechanisms have become self-reproducing and require minimal ongoing outside support.

- **Durable, high-quality infrastructure has fostered enduring practices in sanitation and school feeding**. Canteen kitchens built to international hygiene standards and storage facilities are the clearest example. Importantly, community committees have remained actively engaged in upkeep and maintenance.
- **Engagement with decentralized state structures has helped ensure sustainability.** Continuity of hot-meal provision through 2028 exemplifies the institutional bridge helping ensure communities remain in place.

Recommendations

Based on endline results, STS proposes the following recommendations for future Food for Education projects.

Increase students' exposure to French in all settings to increase literacy levels.
Increasing literacy needs to center around increasing time devoted to reading in French during the day. This could take the form of engaging parents and guardians to encourage reading in French at home, encouraging teachers to collaborate across subjects to incorporate French reading into other lessons, or providing a variety of vocabulary-related materials in French and local languages.

Bilingual Education Programs and TaRL
For students who are not fully proficient in the language of instruction, implementing bilingual education programs could help bridge the gap between their native language (L1) and the language of instruction (L2). This approach would support improved comprehension and contribute to better learning outcomes.

Further improvements in school water and sanitation sources are warranted.
Although notable improvements in school facilities were observed at endline, upgrades of water facilities remain necessary. Future project interventions could make an impact by improving the number of handwashing facilities with clean, running water.

To ensure long-term sustainability, future interventions should focus on the minimum acceptable diet among children 6–23 months old.
Food for Education projects should investigate why minimum dietary diversity (MDD) has fallen between baseline and endline and why minimum meal frequency (MMF) declined during the same period. Future interventions should consider educational and food provision components that target these dynamics.

Collaborate more deeply with decentralized state structures.

Qualitative data collection revealed multiple respondents who called for deeper collaboration with decentralized state structures to avoid "repetition of past mistakes" and secure national recognition of successful approaches. One manager observed: "Given the lessons learned [...] intensify collaboration with the decentralized structures of the State and even see the State involve them more [...] for the success of the project."

Examine gender constraints within target communities.
Girls' underperformance compared with boys deserves further exploration and may warrant a specific focus in future interventions to address the underlying causes of these gender disparities. Across evaluation time points—baseline to midterm to endline—gender gaps in learning outcomes remain stagnant or even grew. Future project interventions should focus on resources to help close this gap, such as gender-responsive approaches. For instance, adopting an inclusive and differentiated approach that addresses the needs of all learners, such as TaRL (Teaching at the Right Level). This method can be effective for fostering equity in the classroom while avoiding stereotypes, ensuring that it applies equally to all students. Future interventions should be mindful not to foster girls' development at the expense or neglect of the boys who fail to meet the benchmark.

1. Introduction and Purpose

1.1. Project Context

Geography and Demographics

The Republic of Togo, located in West Africa, had an estimated population of 9.5 million in 2024, with 40 percent of the population under the age of 14 (World Bank, 2024).[4] The country remains one of Africa's smallest and most densely populated, particularly in urban centers.

Political and Economic Landscape

Togo experienced political upheaval in the 1990s, which led to diplomatic and economic isolation. Although diplomatic ties were restored in the mid-2000s, the country continues to grapple with the long-term effects of this period. While poverty rates have declined in recent years, economic growth remains uneven, with stark disparities between urban and rural populations.

Figure 1: Map of CRS Togo Intervention Prefectures

Togo's economy is primarily agriculture-based, and rural communities face economic hardship. In 2023, an estimated 58.8 percent of rural households lived below the poverty line, compared to a significantly lower percentage in urban areas (INSEED, 2023).[5] This divide is also reflected in social services, such as education, healthcare, and access to clean water and sanitation.

Education Challenges

Togo has tried to improve educational access, but regional and gender disparities persist, particularly in the northern regions. According to 2023 government data, out-of-school children of primary school age are still predominantly from rural areas (88.1 percent), compared to 11.9 percent in urban areas (UNESCO, 2023).[6] The regions of Savanes and Kara have the highest concentrations of out-of-school children, 27.9 percent and 27.0 percent, respectively. Most of these children come from low-income families, and 53 percent are girls (UNICEF, 2023).[7]

Girls from low-income households face additional barriers to education. While they have an 89 percent probability of entering primary school, their chances of completing it drop to just 60

[4] CONFEMEN. (2019). *PASEC 2019: Performance of Education Systems in Francophone Countries*. Conférence des Ministres de l'Éducation des États et Gouvernements de la Francophonie.
[5] INSEED. (2023). *National Institute of Statistics, Economic and Demographic Studies: Togo Poverty Report 2023*. Lomé, Togo.
[6] UNESCO. (2023). *Global Education Monitoring Report: Togo's Education System and Challenges*. Paris, France.
[7] UNICEF. (2023). *State of Education in West and Central Africa: Togo Country Report*.

percent. Factors such as early marriage, household responsibilities, and long distances to school contribute to these low retention rates (UNICEF, 2023).[8]

Additionally, early grade reading outcomes remain a significant concern. According to 2019 and 2024 studies by the Conférence des Ministres de l'Éducation des États et Gouvernements de la Francophonie (CONFEMEN), more than 75 percent of grade 2 students in Togo do not read at an acceptable level (CONFEMEN, 2019; CONFEMEN, 2024).[9] This has implications for students' long-term educational success, as early literacy is a critical predictor of future learning outcomes.

Disparities in Sanitation, Health, Nutrition, and School Feeding Programs

The rural-urban divide is also evident in health indicators and access to sanitation. According to UNICEF (2024):

- 89.1 percent of urban households in Togo have access to improved water sources, compared to only 48.4 percent of rural households (UNICEF, 2024).[10]
- The gap is even wider for improved sanitation, with 28.6 percent of urban households having access to improved sanitation facilities, compared to 7.4 percent of rural households (WHO/UNICEF Joint Monitoring Programme, 2024).[11]

Limited access to clean water and sanitation contributes to high rates of waterborne diseases, disproportionately affecting school-aged children and hindering their educational attendance and performance (WHO, 2024).[12]

Child nutrition and school feeding programs significantly influence educational attendance and performance. According to WFP (2024):

- Only 11 percent of public elementary schools in Togo currently benefit from school feeding programs, far below the 17 percent national target set for 2025 (WFP, 2024).
- Malnutrition remains a primary cause of high school dropouts and low attendance rates, especially affecting girls in rural regions (WFP, 2024).

Limited access to adequate nutrition and regular school meals adversely impacts children's cognitive development, attendance, and academic achievement (WFP, 2024).

Implications for the STARS Project

The *Santé, Transformation et Apprentissage pour une Réussite Scolaire (STARS)* project addressed educational disparities in rural Togo, particularly in the Savanes and Kara regions. The project targets early-grade literacy by improving reading instruction, providing learning materials, and enhancing teacher training. Given the socio-economic and infrastructural challenges, the project also works to engage communities and parents in supporting children's education, particularly girls and students from low-income households. The STARS project also built latrines, improved WASH stations, and provided canteens to motivate regular attendance.

[8] UNICEF. (2024). *Togo: Water, Sanitation, and Hygiene (WASH) Assessment Report 2024*.
[9] CONFEMEN. (2024). *PASEC 2024: Early Grade Reading Assessment Results in Togo*. Conférence des Ministres de l'Éducation des États et Gouvernements de la Francophonie.
[10] UNICEF. (2024). *Togo: Water, Sanitation, and Hygiene (WASH) Assessment Report 2024*.
[11] WHO/UNICEF Joint Monitoring Programme. (2024). *Progress on Household Drinking Water, Sanitation, and Hygiene 2024 Update*.
[12] WHO. (2024). *Togo Health and Sanitation Report 2024*.

1.2. Project Description

Catholic Relief Services (CRS) is implementing the STARS project in Togo. Funded by the United States Department of Agriculture's (USDA) McGovern-Dole International Food for Education and Child Nutrition (McGovern-Dole) Program, STARS aims to combat hunger while enhancing literacy and primary education. The McGovern-Dole program operates globally, offering school meals, teacher training, and additional educational support to promote enrollment and improve learning outcomes.

Running from fiscal year (FY) 2020 to FY2024, STARS has been extended at no additional cost through September 2025. With a $20 million budget, CRS initially targeted 36,341 primary school students in 138 schools, expanding its reach to 46,925 students by the final year due to projected enrollment growth. The project delivers educational and nutritional assistance in Togo's northern regions, specifically in Kpendjal and Oti-Sud (Savanes region) and Dankpen (Kara region). The program focuses on the following key objectives:

- Improve literacy outcomes by strengthening school systems and community support.
- Improve the quality of literacy instruction by building the capacity of teachers and administrators and providing sufficient literacy materials.
- Improve student attentiveness and attendance by providing daily school lunches and ensuring a safe school environment.
- Improve targeted beneficiaries' health and dietary practices by increasing awareness of nutrition, health, and hygiene behaviors combined with improvements in water and sanitation infrastructure.
- Increase the capacity of the government and other key actors to improve school feeding, health, and nutrition and prioritize literacy in education.

CRS collaborates with various partners to ensure the effective implementation of STARS. Alongside community members and national and local government entities, CRS works with key organizations such as the World Food Programme (WFP) for school feeding initiatives and policy advocacy, UNICEF for supporting school governance, preschool teacher training, coordinating water point and latrine construction, and promoting hygiene (WASH) and child protection programs. In addition, STARS received deworming medicine from the National Agency for Grassroots Development (ANADEB) administered by community health workers of the Ministry of Health. This partnership-driven approach enhanced the program's impact on school communities. Findings from STARS are shared with stakeholders through dissemination workshops, webinars, and reports to support continuous improvements in education and child nutrition.

Table 1: STARS Project Stakeholders

Students	Community leaders/Mayors
Parents	Ministry of Primary and Secondary Education
Teachers	Ministry of Grassroots Development
School administrators	Ministry of Health and Social Protection
Food preparers	National Federation of Parents
School Management Committee members	Inter-ministerial committee members
Parent-Teacher Association (APE) members	Implementing partners
Savings and Internal Lending Community (SILC) members	World Food Program and UNICEF
Lead mothers	World Bank
Child Promotion Agents	USDA
Community Health Workers	

1.3. Results Framework

The following frameworks describe the expected activities, outputs, and outcomes to be measured through the three evaluations.

Figure 2: Results Framework

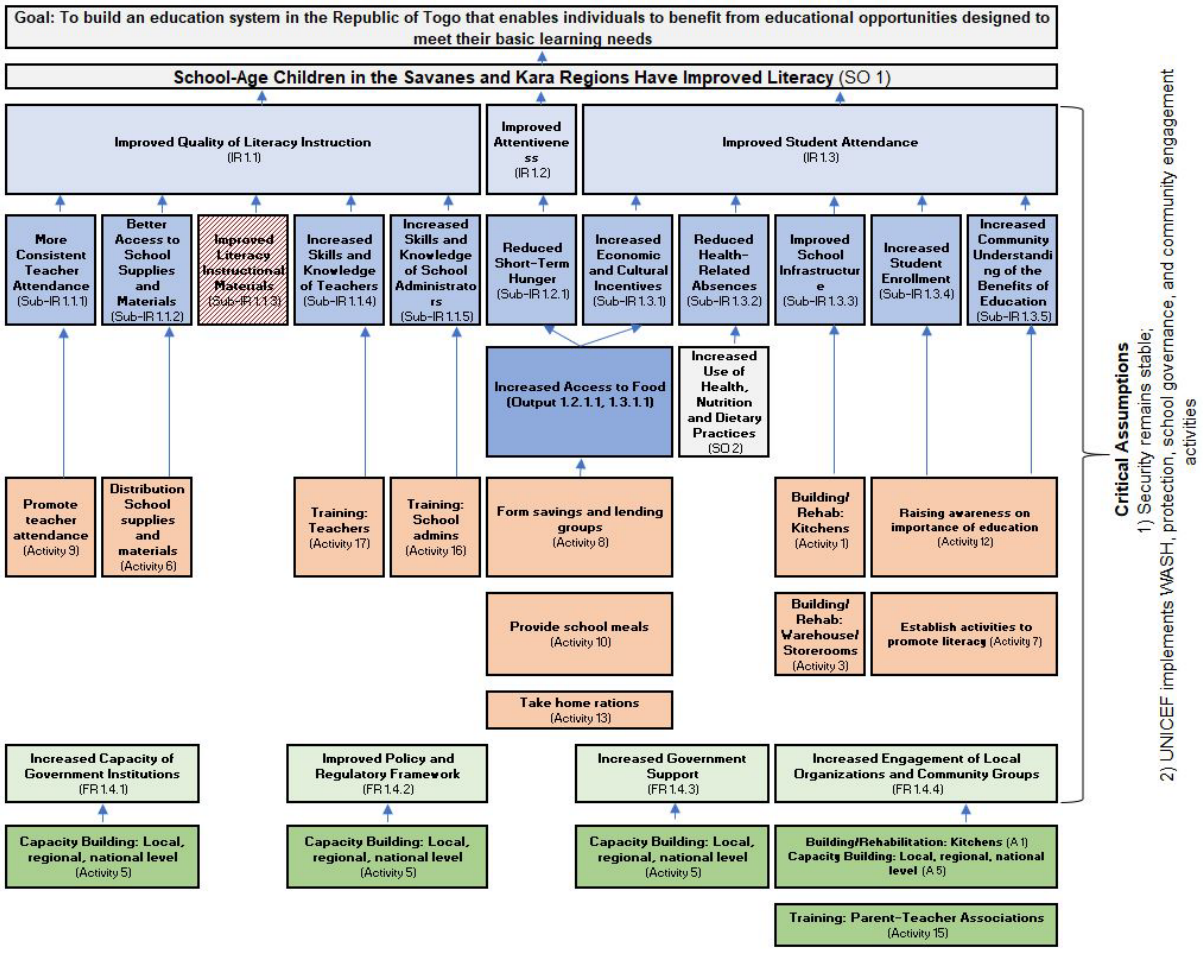
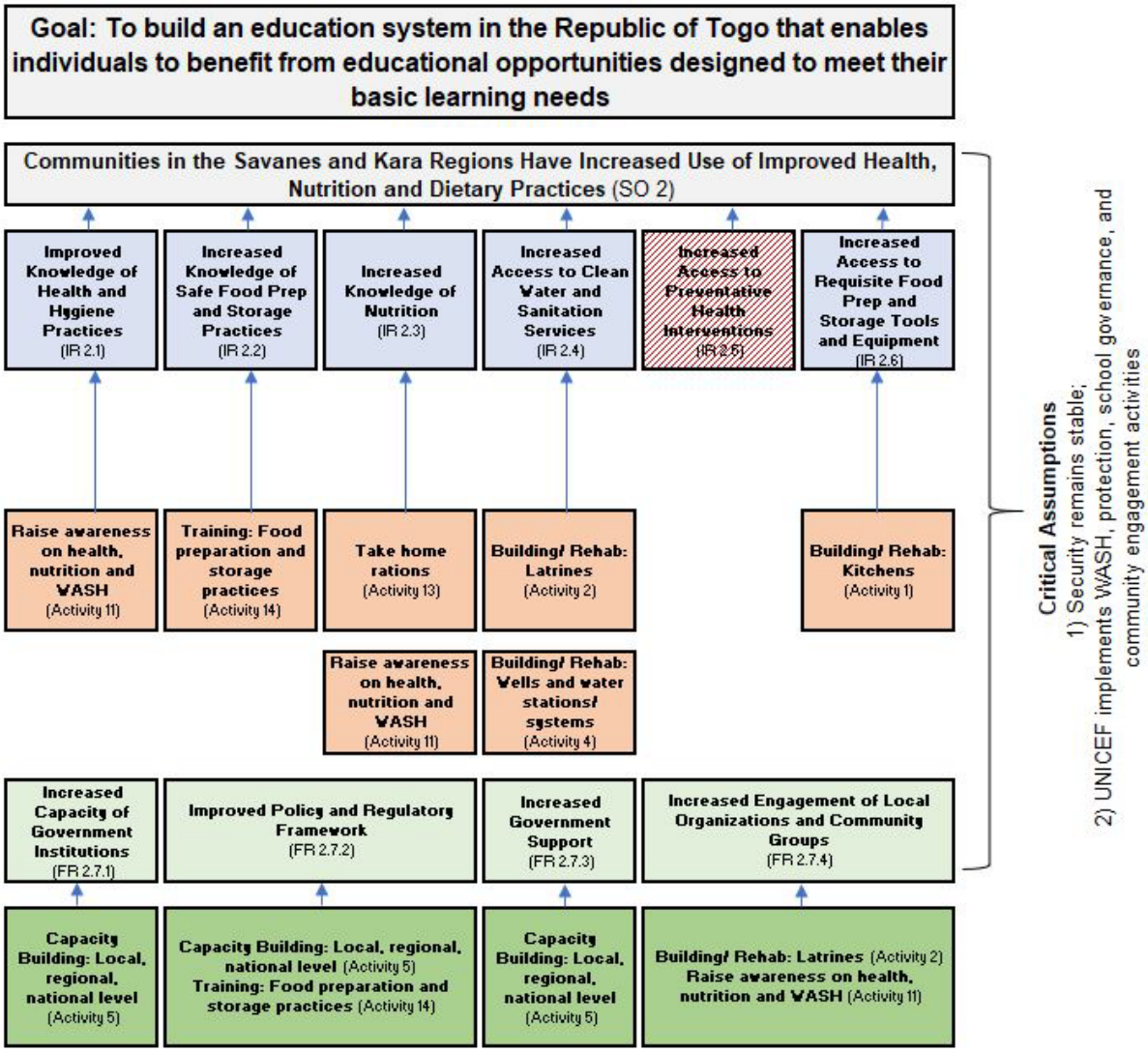


Figure 3: Results Framework



Theory of Change

CRS' overarching Theory of Change for the STARS project is **IF** the school system is strengthened and delivers quality literacy instruction, **IF** communities and parents support their child's education and invest in health, **IF** children benefit from safe and nutritious meals and **IF** schools provide a safe and stimulating learning environment, **THEN** children in Savanes and Kara regions will attend school regularly, thrive and learn, as evidenced by assessment results.

School Feeding

IF national and local government authorities coordinate their actions toward their vision to expand school's canteens nationwide, **IF** policy and regulatory frameworks are strengthened with clear roles set out for the management of school feeding, **IF** school governance structures at the community level, such as the SMC and APEs, and community members hold government officials and school administrators accountable to improve schools governance, **THEN** schools systems will be strengthened and contribute to sustain the access to a quality meal for each pupil.

Evidence: A first pathway for improved nutrition centers on the contribution of school feeding to overall daily food intake, which depends on the quality of the school feeding transfer as well as intra-

household allocation of food.[13] Improved intake can contribute to improved physical and psychosocial health and nutrition outcomes. The second channel involves other school health and nutrition services that are provided in schools alongside school feeding that may address health and nutrition outcomes directly (e.g., de-worming) or indirectly by improving nutrition knowledge, attitudes, and practices (e.g., nutrition education and SBC on healthy food choices).

Quality of literacy instruction

IF the educational system prioritizes literacy improvement, **IF** teachers and school administrators improve competency and commitment to teach literacy and to be accountable to parents, **IF** schools have access to appropriate teaching and reading materials, and **IF** communities sustain literacy activities both in and out of the classroom, **THEN** the school system will deliver quality literacy instruction and student literacy will improve.

Evidence: Strong evidence exists of the linkage between teacher training and the supply of materials with improved literacy outcomes. The most research-based strategy at both the pre- and in-service levels to improve children's learning achievement is to improve teachers' delivery of literacy.[14] This design will, therefore, heavily emphasize supply-side interventions that directly support teacher development. There is also evidence of a strong relationship between teacher presence and student achievement, particularly in first and second grade.

Child health and attendance

IF students have access to improved water and sanitation infrastructures, **IF** household members have increased awareness, and improved practice, of key nutrition, health and hygiene behaviors, **IF** parents are supportive of student's education and empowered to pay related costs, **IF** children consume improved diets at home and at school, **IF** schools provide a safe and enabling environment, **THEN** children will remain in good health and will attend school regularly.

Evidence: WASH interventions reduce the number of sick students, particularly when water supply is assured in combination with other approaches.[15] Stunting, iron and iodine deficiencies negatively impact growth and brain development and lead to more frequent and severe illness. Promotion of key nutrition and health behaviors are among the most effective interventions for reducing the burden of childhood malnutrition. Likewise, the use of key preventative and curative health services, including vitamin A supplementation, is a proven strategy for reducing malnutrition and improving child survival.[16] Quality school meal programs and improved health have demonstrated a positive effect on school attendance. Anecdotal evidence from other CRS programs suggests that parents belonging to village-based savings groups are more likely to pay school fees on time, thus improving enrollment and attendance.

The Theory of Change, presented graphically in the results framework, has been assessed during the evaluations.

[13] Bundy, Donald. 2011. Rethinking School Health: A Key Component of Education for All. Directions in Development--Human Development. World Bank.
[14] University Research Co. Reading Within REACH (2018) Towards the Design and Implementation of Comprehensive Primary Grade Literacy and Numeracy Programs. USAID.
[15] Rassas, Ariza-Nino, & Peterson. N.D. The McGovern-Dole International Food for Education and Child Nutrition. Program Health Interventions and their Educational and Health Outcomes in Developing Countries: A Systematic Review and Meta-Analysis.
[16] Black et al., 2013. Executive Summary of *The Lancet* Maternal and Child Nutrition Series.

Critical Assumptions

Critical Assumption 1: Security will remain stable in project areas. Dankpen prefecture in Kara borders Ghana and experiences patterns of displaced persons due to border tensions. Additionally, the Northern border region is adjacent to Burkina Faso, where extremist groups have carried out attacks; however, this has not affected Togo. CRS will monitor developments through its monitoring systems and the partner SMCs and alert USDA of any real or potential impact on project implementation. CRS MGD programs in Mali and Burkina Faso face similar security concerns. CRS Togo applied learning from their previous experiences to the Togo context.

Since the start of the STARS project, Critical Assumption 1 has changed as a growing number of incursions and attacks by armed non-state actors have occurred in the Savanne region, particularly in FY23, thereby limiting project implementation and access to the region. The increased risk in the region changed the evaluation team's planned activities, including limiting entry into the region, conducting remote interviews, and or not including the Kpendjal region in the final evaluation. The ongoing insecurity in the region resulted in the closure of at least three schools in Kpendjal and the displacement of several others, resulting in the dispersing of students to project and non-project schools, impacting the number of students and meals. As a result of the ongoing insecurity in Kpendjal, CRS withdrew most of its activities in the new MGD project.

Critical Assumption 2: UNICEF carried out the operations agreed upon during the proposal and startup, including implementing continued sanitation activities and new WASH, as well as protection, school governance, and community engagement activities.

Strategic Objectives

The STARS project centers around the two USDA McGovern-Dole strategic objectives (SOs):

- SO 1: School-aged children in the Savanes and Kara regions have improved literacy
- SO 2: Communities in the Savanes and Kara regions have increased use of improved health, nutrition, and dietary practices

Both SOs are supported according to the STARS Project Results Framework (Annex C).

Under the project's first SO, STARS implemented several school-based activities to improve school-aged children's literacy in 138 intervention schools. CRS recognized teachers' critical role in students' learning and focused on literacy training for teachers, school directors, and inspectors. These efforts were bolstered by providing quality teaching materials for use in the classroom.

As the heart of the McGovern-Dole project, daily school lunches were provided through community-operated canteens at all intervention schools to encourage students' attendance and attentiveness. Food preparers and school administrators received training on proper food preparation, storage, and sanitation practices.

The project's second SO sought to increase the use of health, nutrition, and dietary practices by promoting health, nutrition, and personal hygiene initiatives within the schools and communities. As such, CRS worked to improve school water and sanitation facilities, enabling students to put proper health behavior into practice. The project built and repaired gender-segregated latrines per national standards, and new wells were built at schools currently without access to water. CRS also distributed take-home rations to pregnant and lactating women and to children under two years of age who participated in CRS's community-based maternal and child nutrition activities.

To achieve these ambitious goals and promote local and national sustainability, the STARS team consistently worked alongside local communities, organization partners, and the Government of Togo ministries, departments, and agencies, including the Ministries of Education, Health, Agriculture, and Grassroots Development.

1.4. Purpose of the Evaluation

CRS contracted School-to-School International (STS) as the independent external evaluator for the STARS project. In addition to the final evaluation conducted in November 2024 and outlined in this report, the project's evaluation plan includes a baseline evaluation completed in November 2020 and a midterm evaluation conducted in November 2022.

The final evaluation measures progress on SO 1 and SO 2. This report generates data for comparative analysis and helps CRS validate the project's strategies and assumptions. Results illustrate the project's successes and can provide direction for the focus of further interventions.

2. Evaluation Design and Methodology

2.1. Evaluation Design

The endline evaluation was the third stage of the evaluation plan. It sought to 1) assess if the project achieved planned results and 2) identify implementation strengths, challenges, opportunities, good practices, lessons learned, and replicability for the GoT, USDA, CRS, and other project stakeholders, in accordance with CRS, DAC, and the USDA Monitoring and Evaluation Policy.

The evaluation's methodology, selection of the evaluation team, and key audience for the evaluation are described in the Evaluation Design and Methodology section. Table 2 presents the proposed timeline.

Table 2: Proposed Final Evaluation Timeline

Activity	Dates
Submit revised ToR to USDA	September 2024
Refine and re-validate survey tools	October2024
Collect final evaluation data	November 2024
Final report submitted to USDA	May 2025
Share study findings with stakeholders	To be scheduled

Evaluation Timeline Shifts

Under the original terms of reference, the baseline evaluation was planned for the end of the 2019-2020 academic year with grade 2 students (*cours préparatoire 2*, CP2) in the spring of 2020. However, the COVID-19 pandemic interrupted the baseline evaluation after STS completed initial activities—tool development and enumerator training—in March 2020. With school closures across Togo in April 2020, data collection was paused until the situation stabilized and schools could reopen.

After months of disruption, baseline evaluation activities were able to resume in October 2020 at the start of the 2020-2021 academic year. This delay required conducting a second round of

enumerator training due to the eight-month gap between the original STS training in Lomé in March 2020 and the new data collection timeline of November 2020.

Due to COVID-19 and the revised data collection timeline, school closures also warranted a shift in the target sample to grade 3 students (*cours élémentaire 1*, CE1). While Indicator #1 measures the "percent of students who, by the end of two grades of primary schooling, demonstrate that they can read and understand the meaning of grade-level text," the baseline evaluation assessed students at the start of CE1 as a proxy for students at the end of CP2 because their exposure to CE1 instruction was minimal at the time of the evaluation. This was then mirrored at midterm and endline to produce valid comparisons between baseline and midterm.

Assessing students at the start of a new academic year as a proxy measure for student learning levels at the end of the prior academic year is common among education evaluations. Further, COVID-19-related school closures in April 2020 meant that students entering CE1 in the 2020-21 school year had not been exposed to the full CP2 curriculum by the start of the new school year.

Ethical Considerations

The CRS Togo team reviewed the study tools before the beginning of data collection to ensure that the study adhered to applicable ethical rules and societal norms. STS and its data collection partner trained all enumerators in child protection policies and procedures. Enumerators obtained affirmative informed consent from all head teachers and classroom teachers to assess the children in their care. All children provided affirmative assent to be assessed and interviewed and could opt out of the assessment or survey anytime.

Furthermore, for data privacy concerns, data collected electronically were stored on a secure, password-protected server, which only STS can access. Respondents were assigned a randomly generated identification code, so no names were recorded in the datasets that included respondents' answers.

2.2. Sampling Methods

A two-stage cluster sampling approach was used for the baseline, midterm, and endline evaluations. Sample sizes were calculated using Equations (6), (19), and (22) for clustered continuous, non-clustered binary, and clustered binary outcomes, respectively, in McConnell and Vera-Hernandez, using the standard 80 percent power and 5 percent significance level.[17] First, 80 schools were randomly selected from the list of 129 intervention schools to serve as clusters. Within each selected school, enumerators sampled the following units for surveys or observations:

- One head teacher or assistant head teacher
- One classroom between grades 1 and 5 to be observed for a classroom observation
- Three parents of students who also have a child under the age of two[18]

For the second sampling stage, enumerators followed a specific procedure to randomly select 20 students to participate in the evaluation—10 boys and 10 girls—from those present in the CE1 classroom at each school on the day of the data collection visit. This number was more than the

[17] McConnell, Brendon, and Marcos Vera-Hernandez. 2015. Going beyond simple sample size calculations: a practitioner's guide. Institute for Fiscal Studies.

[18] Sampled parents were identified and invited by the head teacher. For the midterm and endline evaluations, parents were selected from active participants in STARS activities to ensure they meet the sampling requirements. Enumerators were asked to call back the director the day before the visit and ask for the presence of 3 parents (preferably the mothers) with at least one child aged 6 months to 2 years. The probability of parents being selected for the sample was not explicitly randomized, as selection was based on active participation in STARS activities and availability. Weighting for the survey responses was applied at the school level, accounting for differences in participation and ensuring representativeness across the schools included in the study.

minimum target sample size of 15 students per school to allow for an equal number of boys and girls per school. If a school had more than one CE1 class, enumerators randomly selected one classroom to identify the 20 students. Sample sizes were increased from baseline in response to the level and low variation in baseline scores to get enough statistical power.

The target sample size of 80 schools covered just over half (62.0 percent) of the 129 intervention schools. The sample was designed to be generalizable at the project level. The target and achieved sample numbers are reflected in Table 3.

Table 3: Quantitative Target and Actual Sample Number

Group	Minimum Target Sample	Actual Sample	Response Rate
Schools	80	80	100.0%
Head Teachers	80	80	100.0%
CE1 Students	1600	1,571	98.2%
Classroom observation	80	79	98.8%
Parents	240	239	99.6%

In addition to the sample, STS created a list of replacement schools in case of unforeseen challenges. For each closed or inaccessible school, the study team selected a comparable school from the list of replacement schools to visit.

Quantitative data was supplemented with qualitative data from Focus Group Discussions (FGDs) and Key Informant Interviews (KIIs). The target and achieved sample numbers are reflected in Table 4.

Table 4: Qualitative Target and Actual Sample Numbers

Respondent	Type	Target Sample	Actual Sample	Response Rate	Pending
School Director	KII	6	6	100.0%	
Teacher	FGD	6	6	100.0%	
Mayor	KII	6	6	100.0%	
COGEP	FGD	3	6	200.0%	
Lead Mothers / SILC	FGD	3	5	166.7%	
CRS	FGD	1	1	100.0%	
CIMAS	FGD	1	0	0.0%	CIMAS unavailable for data collection
MoE, prefecture or canton level, incl. pedagogical advisors & inspectors	KII	3-6	2	66.7%	
Local Partners	KII	2	2	100.0%	

Data Collection Tools

To collect the qualitative data, Focus Group Discussions (FGDs) and Key Informant Interviews (KIIs) were conducted as follows:

- Two teacher FGDs and two COGEP FGDs were conducted in each prefecture in the selected schools.
- Five FGDs with *Mamans lumières* were conducted, except in Dankpen, where only one village had a committee.

2.3. Data Collection Methods

Data Collection Tools

EGRA

The STARS endline evaluation utilized the same data collection tools as at baseline, which were adapted from comparable contexts. The tools included an Early Grade Reading Assessment (EGRA); a CRS-developed student survey and a classroom observation tool used across CRS McGovern-Dole projects; and surveys for head teachers and parents. Additionally, STARS used qualitative tools, including focus group discussion (FGD) and key informant interviews (KIIs), at endline.

CRS funded a fully contextualized EGRA for the Ministry of Education, yet it was decided to use the same tool for all time points to maintain continuity in the project and allow for the best comparisons between stages. Instead, STS and the CRS Togo team reviewed the tools and made specific revisions before data collection to ensure survey tools were responsive to the STARS performance monitoring plan and were culturally appropriate.

STS administered an endline EGRA to students at the start of CE1 to measure their foundational early grade reading skills. The endline assessment contained five subtasks—Initial Sound Identification, Letter Sound Identification, Oral Reading Fluency, Reading Comprehension, and Listening Comprehension. The baseline and midterm assessments included a sixth subtask: Nonword Reading. The Nonword Reading subtask was removed before endline due to floor effects observed at prior time points and to reduce assessment duration, minimizing student test fatigue. Table 5 provides a summary of the subtasks in order of administration.

Table 5: Early Grade Reading Assessment Subtasks

Subtask	Core Reading Skill	Subtask Description
Initial Sound Identification	Phonemic awareness	The enumerator said 10 common words aloud and asked students to identify the first letter of each word.
Letter Sound Identification	Alphabet knowledge	The enumerator presented students with a grid of 100 letters, or groups of letters, in both uppercase and lowercase in a random order and asked them to say the sound of as many letters as they could in one minute.

Subtask	Core Reading Skill	Subtask Description
Oral reading fluency	Decoding and reading fluency	The enumerator asked students to read a short, grade-appropriate story of 57 words in one minute with accuracy and little effort.
Reading comprehension	Reading comprehension	The enumerator asked students as many as five questions about the passage read in the previous subtask, including four literal questions and one inferential question.
Listening comprehension	Listening comprehension and oral language	The enumerator read aloud a short story of 38 words and asked students five questions about the story, including four literal questions and one inferential question.

Enumerators administered the EGRA to 20 CE1 students at each school on tablets using Tangerine®, an electronic data collection software. Following the EGRA subtasks, enumerators administered a short survey to these same students (Table 6). Results from the EGRA contribute to measuring STARS indicators 1 and 3.

School-based Surveys and Observation Tools

Enumerators collected data with three survey tools and a classroom observation tool at each school to provide a comprehensive picture of a sampled school's environment. The content of these surveys is described in Table 6.

Table 6: School-based Surveys and Observation Checklists

Tool	Types of Information Collected
Student Survey	Students' feelings about school; their teachers' use of quality teaching practices; educational support at home; available water and sanitation resources at school and home; and home socioeconomic factors.
Head Teacher Survey	Enrollment and attendance data; teacher attendance and support information; school administration tools; teaching and learning materials available; and school water, sanitation, and nutrition resources.
Parent Survey	Household demographics; child school absences; knowledge of and use of nutrition, health, and sanitation practices; educational support at home; and dietary practices for children under two years.
Classroom Observation	Presence and use of teaching and learning materials in the classroom; use of quality teaching practices within an observed lesson; evidence of student attentiveness; and the school's physical attributes, including sanitation facilities, water sources, and food preparation and storage areas.

The CRS global education team developed the CRS Global Child Learning tool, which includes a classroom observation and student survey for use across all their McGovern-Dole projects. At baseline, STS had added a few questions to these tools to address the required performance

indicators but kept the core tools consistent. Also, at baseline, STS developed the parent and head teacher surveys with input from the STARS project team to align with the performance indicators and adapted several questions from similar tools from CRS's McGovern-Dole projects in both Benin and Burkina Faso. These same tools were utilized at midterm and endline. Results from the modified CRS Global Child Learning tool contribute to measuring STARS indicators 8, 9, 15, 16, 19, 23, 27, 28, 30, 34, and 36.

Qualitative Data Tools

To collect the qualitative data at endline, STS developed FGD and KII questions and protocols. CRS reviewed and proposed relevant changes to the FGD and KII questionnaires.

FGDs and KIIs were conducted as follows:

- Two teacher FGDs and two COGEP FGDs were conducted in each prefecture in the selected schools.
- Five FGDs with *Mamans lumières* were conducted, except in Dankpen, where only one village had a committee.

Findings from the FGDs and KIIs help inform findings and provide context to the STARS indicators.

Recruitment and Training of Enumerators

STS contracted the data collection firm Innovative Hub for Research in Africa (IHfRA) to manage local aspects of the evaluation, including the selection and hiring of enumerators, training logistics, and the supervision and management of data collection in the field. IHfRA recruited 43 enumerators who were part of the baseline data collection or are familiar with the terrain and spoke one or more languages in common with communities across the survey regions. These individuals participated in the enumerator training on evaluation tools and protocols from November 20-26, 2024, in Kara, Togo.

Before the training, STS designed the training agenda and created supplementary PowerPoint presentations, handouts, and other training resources to support the enumerators' learning. The training covered the STARS project and evaluation design; contents of the EGRA tool and school-based surveys; administration protocols for the Tangerine data collection software and use of tablets; ethical considerations and data quality measures; and the responsibilities of enumerators and supervisors during data collection.

Upon conclusion of the training, STS and IHfRA selected the 36 top-performing enumerators to conduct data collection. The merit-driven assessment processes gave each participant an equal opportunity for fieldwork selection. Participants were assessed with written quizzes and observed evaluations of their performance both within the classroom and in the field. These tests ranged from comprehension of questions during classroom activities to mid-training quizzes. Assessor Accuracy Measure (AAM) tests were also administered and scored.

Field Tests of Data Collection Tools

The training in Kara included one day of field testing at a nearby school, in which all the survey and observation tools were tested. This activity enabled enumerators to practice the administration of tools in a real-life setting while enabling the evaluation team to identify potential challenges and solutions. The need to hire enumerators with the appropriate local language fluency was a lesson

learned during the first baseline enumerator training field test and was applied to hiring enumerators for the midterm and endline evaluations.

School-based Data Collection

The final data collection was conducted in the Savanes and Kara districts from November 28–December 6, 2024. Twelve teams of three enumerators each visited one school per day. Within each team, one enumerator was designated as the supervisor responsible for introducing the teams to the school and conducting the classroom and student sampling for each team.

Data Monitoring and Quality Assurance

STS and IHfRA closely supervised enumerators throughout data collection to ensure data quality. IHfRA had three field coordinators to supervise teams and accompany them during data collection to conduct on-site spot checks and troubleshoot any issues teams encountered in the field. Additionally, STS's Senior Data and Technical Writing Associate monitored the incoming data daily, checking results uploaded to the server for completeness. Communication with the enumerator teams was maintained through a WhatsApp© group comprised of team supervisors, IHfRA, and STS, allowing for broader communication and faster responsiveness when issues arose.

IHfRA's staff ensured that enumerator teams followed data collection procedures and submitted a field report that logged discrepancies in the number and type of data collected and prescribed in the target sample. STS cross-referenced these reports against the uploaded data. Disposition codes were applied to categorize any issues that emerged during the data collection process. These coding and flagging procedures helped to ensure that the nuanced contexts of data collection at the school level were sufficiently cataloged and considered during the data cleaning, analysis, and reporting process.

2.4. Data Analysis Methods

Sample Weighting

The analysis used sampling weights to produce more representative estimates in the sample of students. Random sampling does not acknowledge that some students have a lower probability of being selected when they represent smaller subgroups within the population, so sampling weights enable analysts to account for these differences in probabilities.

STS computed the weights using background data available from each school in the sample populations, including the number of CE1 classrooms and students in each classroom.[19] STS collected this information via the head teacher survey. Weights were applied when analyzing the EGRA and all survey results. STS used a combined school and student weight for all students and applied the school weight to the two school-based surveys: student surveys and head teacher surveys.

Generation of Findings

STS generated the following descriptive statistics using the data:

- **Mean scores*:*** Average percentage of items answered correctly on a given EGRA subtask

[19] STS shifted the data analysis software used from SPSS to Stata since the baseline. Stata uses a slightly different formula when applying weights, which may cause minor variations (most often at the decimal level) in results from baseline. To ensure accuracy, baseline results were run alongside the endline in Stata to ensure valid comparisons and statistical evaluation.

- **Zero scores**: Proportion of students who did not answer a single item correctly on a given EGRA subtask
- **Proportions**: Proportion of respondents who replied in a specific way to an item on the EGRA, surveys, or observation tools
- **Means**: Average score on survey and observation items

Analysts computed inferential statistics on subtask mean scores to determine differences in performance between girls and boys. Where detected, statistically significant differences are noted in the findings. To assess statistical significance for continuous variables, such as percentage scores and variables from classroom observations and parent surveys, survey-weighted Ordinary Least Squares (OLS) regression models were employed, incorporating adjustments for sampling weights, clustering, and stratification. For binary categorical variables, survey-weighted logistic regression models (logit) were used to evaluate significance. When testing associations between categorical variables, survey-adjusted chi-square tests were conducted to determine statistical significance.

2.5. Evaluation Limitations

The following limitations should be considered when reviewing the findings of the STARS endline evaluation:

- **Non-contextualized EGRA.** The endline data collection utilized the same tool as at baseline and midterm, an existing French EGRA tool adapted in Djibouti. Therefore, the tool was not created specifically for the Togolese context. CRS funded a fully contextualized EGRA for the Ministry of Education after completing the baseline evaluation. The STARS project later decided to use the same tool for all time points to maintain continuity and allow for the best comparisons between stages.
- **Language of the EGRA tool.** The language of the learning assessment—French—is not the mother of many of the students; instead, their mother tongues include the local languages of Konkomba (Dankpen), Gourma (Kpendjal), and Ngam-gam (Oti-Sud). Based on the Listening Comprehension subtask results, many students struggle with listening comprehension in French and may not have understood the EGRA's instructions or testing content. This known limitation was discussed with CRS at baseline, and it was determined that providing an EGRA tool in all local languages would not be feasible. Many different dialects and mother tongues are spoken across the regions where the project is working. For this reason, CRS Togo decided to use the official language of instruction: French. To balance this limitation, IHfRA primarily contracted enumerators from the study area who have language affinity in these regions. The tools were not formally translated, but enumerators were instructed to provide clarification or support in local languages if necessary.
- **The inherent bias in sampling children and parents present on the day of assessment.** Due to the need to sample students who are physically present, EGRA results may be biased towards students who attend school regularly and may exclude those who are enrolled but do not regularly attend school. Similarly, parents willing to participate may have differed in meaningful ways from those not. Therefore, results are not generalizable beyond the target population. However, they are a good representation of the project target population within the same context. Additionally, the sampling approach remained consistent for all

assessments—baseline, midterm, and endline—so the comparison across time points is valid.

3. Findings

3.1 Performance Indicators

The STARS performance monitoring plan sets out numerous indicators to measure the project's progress. The values in Table 7 represent data from both STS's external evaluation and CRS's internal monitoring data. Census data provided by CRS from all 129 intervention schools are presented in shaded boxes, while the non-shaded boxes show evaluation data collected only from the 77 schools randomly sampled at each evaluation timepoint. STS randomly evaluated 77 schools at baseline, 80 at midterm, and 80 at endline. Because the sample was randomized, participation in an earlier evaluation period did not impact a school's likelihood to participate in future evaluation periods.

Table 7: Updated Indicator Performance Tracking Table for Non-Zero Baseline (2020) Indicators

STARS Indicator No.	Indicator Name	Indicator No.	LOP Target	Baseline			Midterm			Endline		
				Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
1	Percentage of students who, by the end of two grades of schooling, demonstrate that they can read and understand the meaning of grade-level text	McGovern-Dole 1	21%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.1%	0.2%
#	Percent of students who, by the end of two grades of primary schooling, demonstrate that they can correctly identify letter sounds	CRS Custom	9.3%	5.8%	4.8%	5.3%	5.1%	3.6%	4.3%	11. 3%	7.0%	9.2%[20]
8	Percentage of students in target schools identified as attentive during class / instruction	CRS Custom	60%	59.7%			79.4%			69.1%		
9	Average student attendance rate in USDA supported classrooms / schools	McGovern-Dole 2	93%	81.3%	79.0%	80.2%	90.6%	89.2%	89.9%	90.7%	88.9%	89.9%
15	Number of schools with improved sanitation facilities	McGovern-Dole 28	66	57			87			116		

[20] Values for this custom indicator are calculated as the percentage of correct responses on the letter-sound identification subtask.

STARS Indicator No.	Indicator Name	Indicator No.	LOP Target	Baseline			Midterm			Endline		
				Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
16	Number of schools using an improved water source	McGovern-Dole 27	90	70			94			107		
19	Percentage of instructional time lost due to teacher absenteeism	USAID Proposed	52%	9.3%			9.3%			12.7%		
23	Percent of teachers providing quality classroom instruction with USG support	USAID Education Proposed	80%	23.4%			43.1%			70.6%		
27	Percentage of school officials in target schools who demonstrate the use of new and quality supervision and leadership techniques or tools	CRS Custom	10%	6.5%			9.0%			15.3%		
28	Percentage of children 6–23 months receiving a minimum acceptable diet[21]	FFP #BL12	9.3%	20.1%			20.3%			12.4%		

[21] The baseline and midline reported the MDD-7 definition per WHO 2007 (in which children must consume foods from four of seven groups to be considered achieving minimum diversity). These figures have now been updated to the MDD-8 definition per WHO 2021 (in which children must consume foods from five of eight groups to be considered achieving minimum diversity) to facilitate comparison between time points in the STARS evaluations, while also enabling comparison to other in-country data points using the WHO 2021 standard.

STARS Indicator No.	Indicator Name	Indicator No.	LOP Target	Baseline			Midterm			Endline		
				Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
30	Percentage of parents who stated that their children had health-related school absences in the previous month	CRS Custom	30%	15.0%			35.0%			10.2%		
34	Percentage of caregivers who report on spending time on literacy activities with their school-age children in the previous week	CRS Custom	42%	15.8%			26.7%			32.9%		
36	Percentage of community members who promote early childhood practices and support their children's education	CRS Custom	20%	60.1%			61.1%			54.3%		

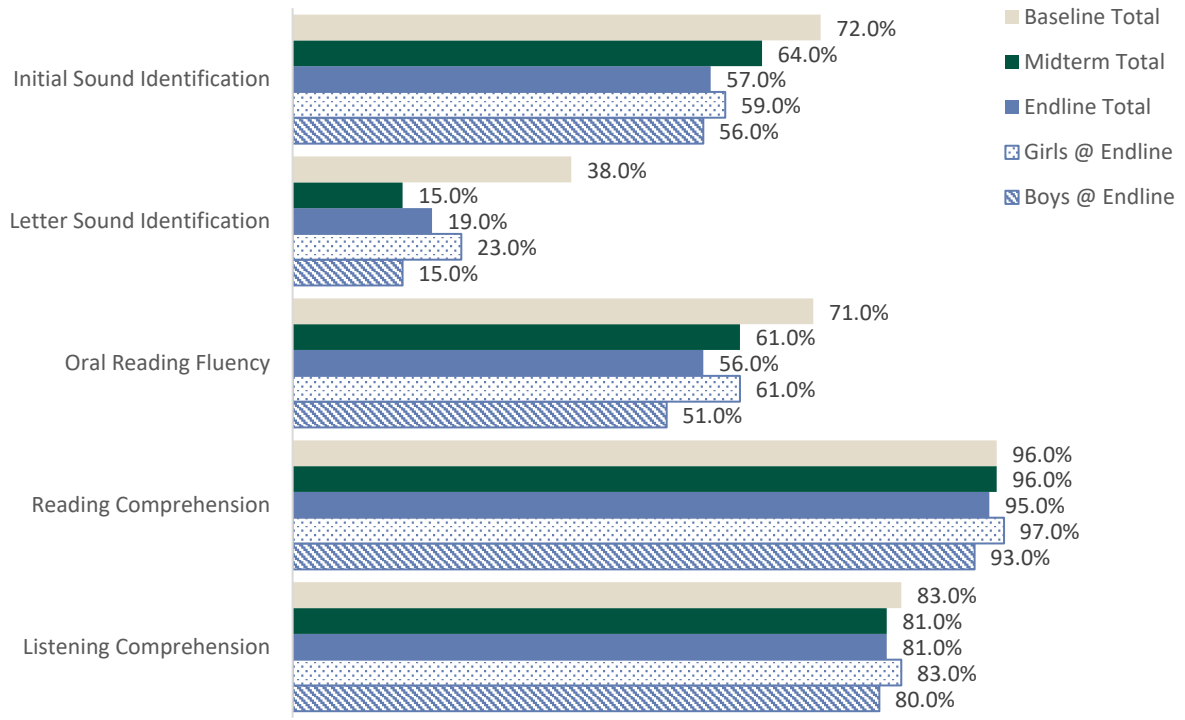
3.2 SO 1: School-Age Children in the Savanes and Kara Regions Have Improved Literacy

The first strategic objective of the STARS project is to improve the literacy of school-aged children in the Savanes and Kara regions. The achievement of this SO is measured through the percentage of students who, at the end of second grade, demonstrate that they can read and understand the meaning of the grade-level text (McGovern-Dole Indicator #1). At the endline, students meet this threshold if they can correctly answer at least three of the five reading comprehension questions or a 60 percent accuracy score. About 0.2 percent of students assessed at endline correctly answered at least three of the five reading comprehension questions.

Zero Scores

The proportion of students who did not answer a single item correctly on each subtask—known as a zero score—is presented in Figure 4. For this study, improvement is seen when a smaller percentage of students receive a zero score. Therefore, zero-score decreases are expected, given the interventions as the project progresses. At endline, most students received zero scores in five out of the six subtasks. The proportion of students with zero scores was lowest on the Letter Sound Identification subtask (19 percent) and highest on the Reading Comprehension subtask (95 percent). Across all subtasks, boys had a lower proportion of zero scores than did girls.

Figure 4: Percentage of Students Receiving Zero Scores, Overall and by Gender



In a sign of improvement, **the proportion of zero scores significantly decreased from baseline to endline among all students on three subtasks:** Initial Sound Identification (13-percentage point

decline), Letter Sound Identification (19-percentage point decline), and Oral Reading Fluency (15-percentage point decline). [22]

- **Initial Sound Identification**. The overall percentage of students who received a zero score significantly declined from 72.0 percent at baseline to 57.0 percent at endline. The proportion of boys with a zero score dropped from 73.2 percent at baseline to 56.0 percent, while the percentage of girls with a zero score decreased from 75.8 percent to 59.0 percent between the two timepoints.
- **Letter Sound Identification.** Like baseline, boys and girls had the lowest proportion of zero scores on this subtask at midterm and endline. At baseline, 38.0 percent of students overall received zero scores on the subtask, including 32.4 percent of boys and 37.3 percent of girls. At endline, only 19.0 percent of the total sampled population received zero scores, including 15.0 percent of boys and 23.0 percent of girls.
- **Oral Reading Fluency.** At baseline, 71.0 percent of students received zero scores on the subtask, including 65.2 percent of boys and 70.3 percent of girls. At the endline, the overall sample of students receiving a zero score decreased to 56.0 percent, including 51.0 percent of boys and 61.0 percent of girls.

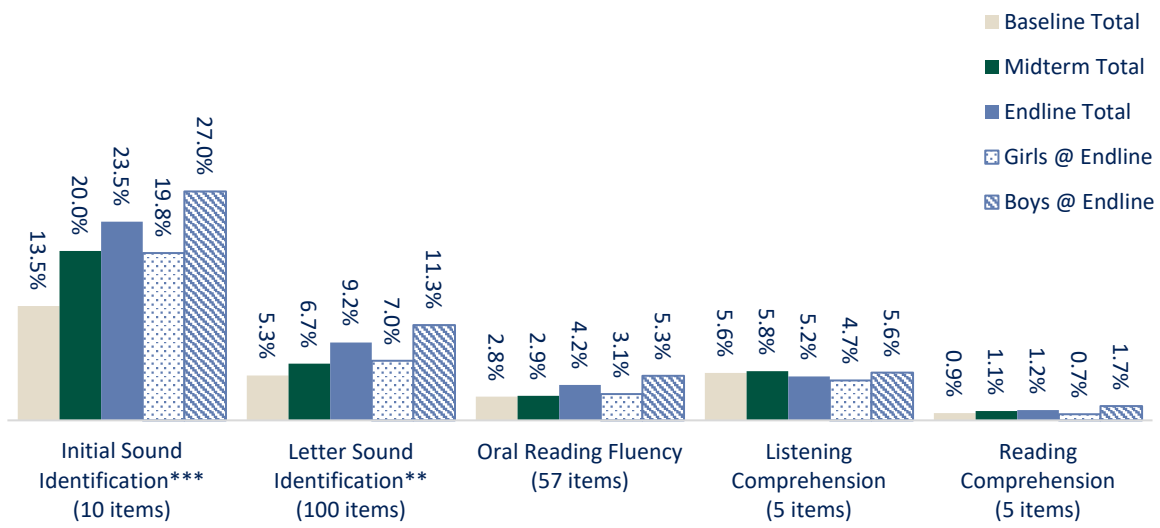
No statistically notable changes were measured on the other subtasks; however, there was a decrease in the proportion of zero scores for Reading Comprehension and Listening Comprehension. Specifically, the proportion of students receiving zero scores decreased—from 96.6 percent to 95.1 percent on Reading Comprehension and 83.0 percent to 81.3 percent on Listening Comprehension. Unfortunately, these changes were minor and not statistically significant. Qualitative data indicates that several respondents perceived overall improvements in pupils' reading ability, but do not provide answers for lack of statistically significant change in reading comprehension or listening comprehension specifically.

Accuracy Scores

When evaluating early-grade reading skills, it is critical to differentiate between accuracy—the ability to correctly identify and decode words—and fluency—the speed and ease with which words are read. Accuracy provides insight into foundational literacy skills, highlighting whether students are developing the essential decoding skills needed for comprehension, while fluency reflects overall reading proficiency and automaticity. In interpreting the EGRA results, the accuracy data reveal important insights. Specifically, low accuracy scores suggest students face difficulties correctly decoding words and indicating gaps in fundamental reading instruction or limited language proficiency. Student performance by subtask in terms of accuracy—as measured by the percentage of correct answers, rather than by the percentage of zero scores—is presented in Figure 5, disaggregated by gender.

[22] Weighted ordinary least squares regressions were used to test the correlational relationship between zero scores at baseline, midterm, and endline with a 95-percent confidence threshold. Significance between the entire sample is indicated with an asterisk on the total bar and disaggregated by gender.

Figure 5: Percentage Accuracy Scores for Literacy Subtasks, Overall and by Gender



Mean Scores

The following section presents mean scores for each EGRA subtask in greater detail, providing a better understanding of students' reading performance to assess students' reading fluency and comprehension by the end of Grade 2. Tests were conducted to determine if the mean scores between boys and girls at endline and from baseline to endline were statistically significant; any statistically significant differences are noted under each table.

Initial Sound Identification

For the Initial Sound Identification subtask, enumerators read aloud 10 common words to students, one at a time. The enumerator then asked students to say the name of the letter corresponding to the word's initial sound. This untimed subtask measures students' awareness of phonemes and their ability to distinguish between multiple phonemes.

Results for the Initial Sound Identification subtask are displayed in Table 8. Students correctly identified the initial sound of two items on average of ten possible items. This was a significant increase from baseline (1.4), with scores improving significantly for both boys and girls. Gender differences persisted at the endline, with boys scoring significantly higher than girls.

Table 8: Initial Sound Identification Mean Scores by Gender (Correct out of 10)

Gender	N	Mean Score at Endline	Percent Correct	Standard Error	Change since Baseline	P value	Significance
Boys	792	2.7	27%	0.2	Increase	0.001	*
Girls	779	1.9	19%	0.2	Increase	0.014	*
Total	1,571	2.34	23.4%	0.2	Increase	0.000	*

Note: This table presents the weighted means and percentage scores. Standard errors are from mean scores and clustered at the school level. Significance testing is conducted with regression analyses on percentage scores (^Y (Pct Score)=β0+β1(Midline)+β2(Endline)+ϵ). The asterisk (*) indicates a significant difference between baseline and endline observations at the 95 percent threshold. Empty cells indicate that the difference between observations was not statistically significant.

Letter Sound Identification

In the Letter Sound Identification subtask, enumerators presented students with a grid of 100 uppercase and lowercase letters and asked students to say the sound of as many letters as they could in one minute. This untimed subtask measures students' knowledge of letters of the alphabet and their ability to recognize each letter's graphemic features.

Endline results for the Letter Sound Identification subtask are presented in Table 9. On average, students named 9.2 letters correctly out of 100. Overall, the Letter Sound Identification scores showed a statistically significant increase from baseline to endline. The difference between boys' baseline and endline scores was statistically significant. Gender differences persisted at the endline, with boys scoring significantly higher than girls.

Table 9: Letter Sound Identification Mean Scores by Gender (Correct out of 100)

Gender	N	Mean Score at Endline	Percent Correct	Standard Error	Change since Baseline	P value	Significance
Boys	792	11.2	11.2%	0.9	Increase	0.000	*
Girls	779	7.0	7.0%	0.5	Increase	0.386	
Total	1,571	9.2	9.2%	0.7	Increase	0.002	*

Note: This table presents the weighted means and percentage scores. Standard errors are from mean scores and clustered at the school level. Significance testing is conducted with regression analyses on percentage scores. The asterisk (*) indicates a significant difference between baseline and endline observations at the 95 percent threshold. Empty cells indicate that the difference between observations was not statistically significant.

Oral Reading Fluency and Reading Comprehension

For the Oral Reading Fluency and Reading Comprehension subtasks, enumerators presented students with a short story of 57 words and asked students to read as much of the story aloud as they could in one minute. After students finished the story, enumerators read aloud as many as five comprehension questions, four direct and one inferential, to students to test their understanding of the story. The number of comprehension questions asked was linked to how much of the story students were able to read in one minute; in other words, students were not asked questions about parts of the story they did not read. These two subtasks measure decoding, reading fluency, and reading comprehension.

Endline results for the Oral Reading Fluency subtask are presented in Table 10. Students correctly read 2.4 words per minute (CWPM) on average, which was an increase from 1.6 CWPM at baseline. This difference was not statistically significant; however, boys showed a statistically significant increase from baseline. Additionally, boys had statistically significantly higher mean scores at endline than girls.

Table 10: Oral Reading Fluency Mean Scores by Gender

Gender	N	Mean Score at Endline	Percent Correct	Standard Error	Change since Baseline	P value	Significance
Boys	792	2.9	5.6%	0.4	Increase	0.011	*
Girls	779	1.7	3.0%	0.2	Increase	1.000	
Total	1,571	2.4	4.2%	0.2	Increase	0.151	

Note: This table presents the weighted means and percentage scores. Standard errors are from mean scores and clustered at the school level. Significance testing is conducted with regression analyses on percentage scores. The asterisk (*) indicates a significant difference between baseline and endline observations at the 95 percent threshold. Empty cells indicate that the difference between observations was not statistically significant.

Endline mean scores for the untimed Reading Comprehension subtask are presented in Table 11. Overall, students were able to answer 0.06 questions correctly at the endline. Although the score was higher than at baseline (0.04 questions), the difference was not statistically significant. Additionally, boys answered 0.08 questions correctly on average and girls only 0.04, but this difference was not statistically significant.

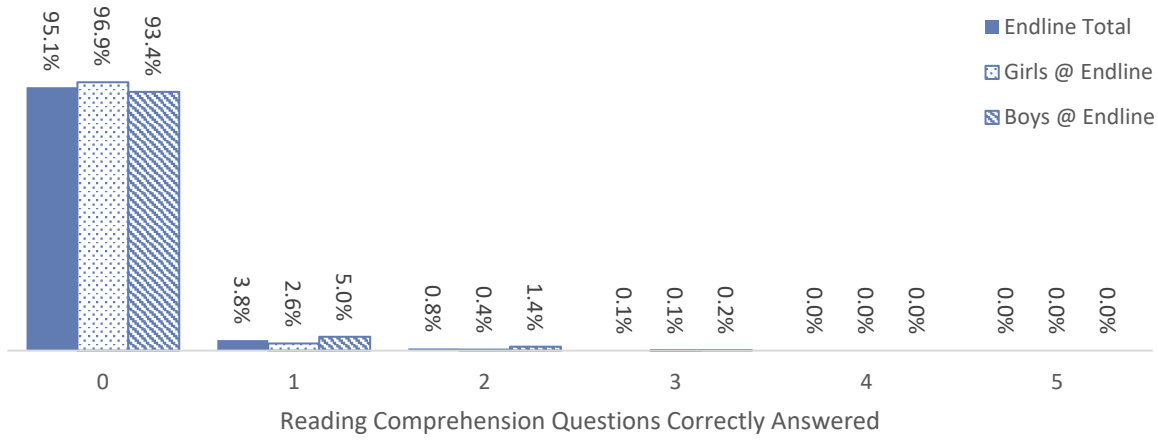
Table 11: Reading Comprehension Mean Scores by Gender (Correct out of 5)

Gender	N	Mean Score at Endline	Percent Correct	Standard Error	Change since Baseline	P value	Significance
Boys	792	0.08	1.7%	0.5	Increase	0.689	
Girls	779	0.04	0.7%	0.2	Decrease	1.000	
Total	1,571	0.06	1.2%	0.3	Increase	0.972	

Note: This table presents the weighted means and percentage scores. Standard errors are from mean scores and clustered at the school level. Significance testing is conducted with regression analyses on percentage scores. The asterisk (*) indicates a significant difference between baseline and endline observations at the 95 percent threshold. Empty cells indicate that the difference between observations was not statistically significant.

The distribution of students who answered reading comprehension questions correctly is detailed in Figure 6. No students answered more than three questions correctly, the threshold set for reading comprehension.

Figure 6: Distribution of Correct Reading Comprehension Questions, Overall and by Gender



Listening Comprehension

The Listening Comprehension subtask consisted of a short story of 38 words that the enumerator read aloud to students. The enumerator then asked students five comprehension questions related to the story—four direct and one inferential. Listening Comprehension measures students' overall oral language comprehension and vocabulary. The Listening Comprehension subtask complements the reading passage and comprehension subtasks, enabling a better understanding of whether reading comprehension difficulties result from reading skills or bigger language comprehension issues.

Endline results for the untimed Listening Comprehension subtask are presented in Table 12. Out of a possible five questions, students correctly answered 0.25 questions on average. At the endline, there was no statistically significant difference from the baseline overall, and there were no score differences between boys' and girls' scores at the endline. The distribution of students who answered listening comprehension questions correctly is detailed in Table 13.

Table 12: Listening Comprehension Mean Scores by Gender (Correct out of 5)

Gender	N	Mean Score at Endline	Percent Correct	Standard Error	Change since Baseline	P value	Significance
Boys	792	0.28	5.6%	0.04	Increase	1.000	
Girls	779	0.23	4.7%	0.04	Decrease	0.373	
Total	1571	0.25	5.2%	0.03	Decrease	1.000	

Note: This table presents the weighted means and percentage scores. Standard errors are from mean scores and clustered at the school level. Significance testing is conducted with regression analyses on percentage scores. The asterisk (*) indicates a significant difference between baseline and endline observations at the 95 percent threshold. Empty cells indicate that the difference between observations was not statistically significant.

Table 13: Distribution of Correct Listening Comprehension Questions by Gender

Number of Questions Correct	Boys	Girls	Total
0	80.18%	82.52%	81.32%
1	14.13%	12.67%	13.41%
2	3.91%	3.7%	3.81%
3	0.9%	0.9%	0.93%
4	0.89%	0.15%	0.52%
5	0.0%	0.0%	0.0%

IR 1.1. Improved Quality of Literacy Instruction

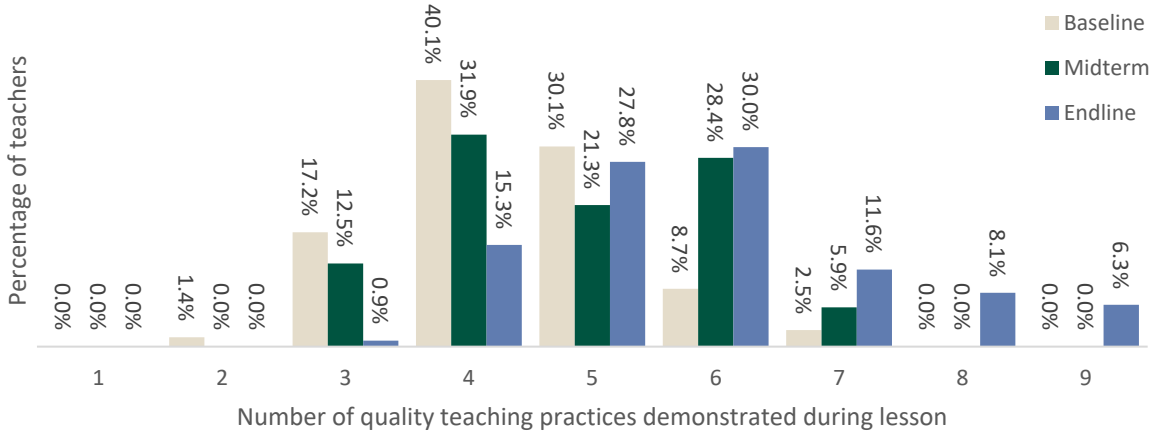
Enumerators used a classroom observation tool to measure the quality of classroom literacy instruction in 80 project schools.[23] Observers observed a one-hour classroom lesson (grades 1 - 6) and recorded activities linked to quality literacy instruction. As defined by the CRS standard pilot classroom observation tool, **70.6 percent of observed teachers met the threshold at the endline**, scoring at least five out of nine on the quality instruction index.[24,25] This is a statistically significant increase from midterm (43.1 percent of observed teachers) and baseline (0.0 percent)[26].

The range of teachers' composite scores of overall quality literacy instruction at baseline, midterm, and endline are shown in Figure 7. For example, at midterm, no teachers demonstrated eight or all

[23] It is important to note that the observation itself, having an observer in the classroom, could bias the results. Specifically, that instructors might engage in different, potentially more rigorous, behaviors when being observed. Steps were taken to try to mitigate this, primarily randomly selecting the classroom to be observed on the day of the observation.
[24] This total reflects the weighted total; unweighted total is 70%.
[25] The classroom observations observed both math and literacy activities; only items relevant to literacy were used to calculate the score. In cases where an item was skipped, the item score was treated as zero. Each question was equally weighted. This means that all activities were given a possible score of 1. While some items were treated as a binary yes or no (e.g., "did the instructor speak French?"), several questions used ordinal response items, asking the enumerator to rate the quality of an activity. In this case each question received a total possible score of 1, with each rating incrementally increasing in value from 0 (e.g., 1-4 will be transferred to .25, .5, .75, 1 respectively).
[26] Survey-weighted regression (OLS regression) was used for statistical testing.

nine quality teaching practices; however, by endline, some teachers demonstrated these highest levels.[27]

Figure 7: Range of Teachers' Quality Literacy Scores at Baseline, Midterm, and Endline



An individual analysis of each component was run to understand further the improvement in the quality of literacy instruction measures. In doing so, the analysis tests what specifically has improved since midline. Aspects of this composite that significantly improved since midline were: learning opportunities to support the development of literacy skills (letter identification, phonetics), that the teacher referred to a lesson plan to structure their literacy teaching, that the teacher spoke French, teachers reading books to help children listen and speak, learning opportunities that allow children to engage in gross motor skills activities, and learning activities that promote free choice or open play.

IR 1.1.1. More Consistent Teacher Attendance

Enumerators asked the head teacher at each sampled school (n=80) a series of questions about teacher attendance, including the number of teachers in the official school records, the number of teachers present on the day of endline data collection, and the average number of hours per school day teachers are estimated to be teaching.[28] These individual questions were used to calculate the percent of instructional time lost due to teacher absenteeism, as seen in Table 14. It is estimated that, across 80 schools, 311 hours of teaching time were lost due to teacher absenteeism, or 12.7 percent. This proportion was higher than the baseline (9.9 percent). Qualitative data did not explain this increase; to the contrary, respondents consistently described strategies such as mentoring cells and regular presence checks by COJEP that have reduced teacher absences.

Table 14: Instructional Time Lost Due to Teacher Absenteeism at Endline

Sampled Schools	80
Teachers Enrolled (total hours)	2,446 hours
Teachers Present (total hours)	2,135 hours

[27] Following recent best practices, this analysis calculated weights at the school level rather than applying student level weights to classroom analyses. Retrospective analysis on baseline records reports small deviations from results presented at baseline. Figure 5 reports results for baseline, midterm, and endline from this updated analysis.

[28] In cases where records of teaching time were abnormally high (over 13 hours, as high as 60 hours), average time responses were reverted to the mean. In one case where one of the records were missing, the case was dropped. Results presented in Table 15are unweighted.

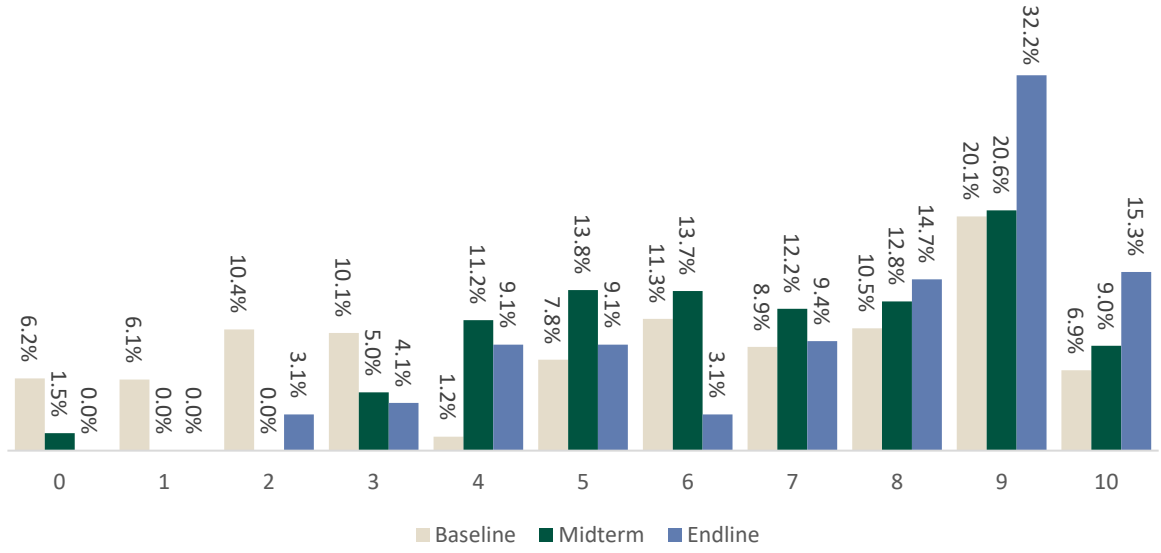
Estimated Hours Lost	311 hours
Estimated Percentage Lost	12.7%

IR 1.1.5. Increased Skills and Knowledge of School Administrators

Enumerators asked the head teacher at each sampled school (n=79) questions about the school's management tools. These tools, provided by the STARS project, included a record of daily teacher attendance, a teacher task list, visual teaching aids and teaching materials, an inventory book, and school records. Out of 10 possible items, a head teacher used quality supervision techniques and tools if an enumerator observed or was shown all 10 items.

The frequency of observed quality supervision tools at all timepoints is illustrated in Figure 8. The number of observed quality supervision tools significantly increased from baseline to endline.[29] For instance, 41.8 percent of school officials used five or fewer tools at baseline, while at endline 47.5 percent used nine or more.

Figure 8: Frequency of School Officials Using Quality Supervision Tools



IR. 1.2. Improved attentiveness

As part of classroom observations, enumerators rated students' level of engagement during the lesson. Children are considered engaged throughout the observation if they are paying attention, watching the teacher, concentrating on the lesson or work, and participating in activities. Engagement levels fall into one of four categories:

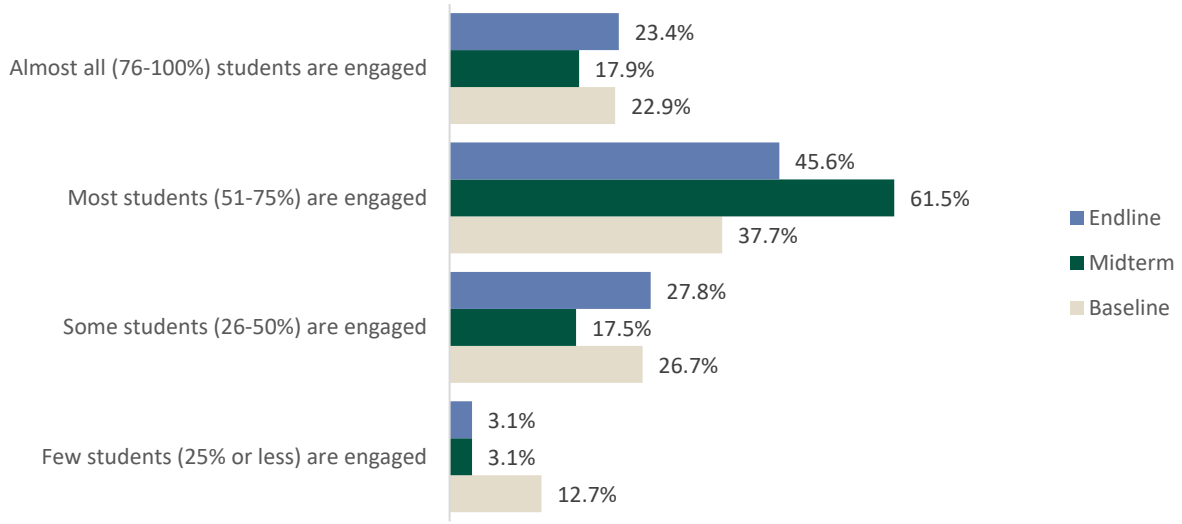
1. Few children (25 percent or less) are engaged for most of the observation
2. Some children (26 percent to 50 percent) are engaged for most of the observation
3. Most children (51 percent to 75 percent) are engaged for most of the observation
4. Almost all children (76 percent to 100 percent) are engaged for most of the observation

If a majority of students were engaged for most of the observation—categories 3 and 4—the classroom was considered "attentive." By this measure, student attentiveness significantly decreased from midterm to endline (Figure 9). At the endline, 69.1 percent of observed classrooms were "attentive," compared with 79.4 percent at midterm. This was primarily driven by a lower

[29] Chi-squared test, p-value (0.00), weighted sample.

proportion of observed classrooms, with 51 to 75 percent of engaged students at endline compared to midterm. Compared to the baseline measure of 59.7 percent of classrooms, endline attentiveness is only a marginally insignificant improvement (p<0.1).

Figure 9: Proportion of Attentive Classrooms Observed



IR 1.2.1. Reduced Short-Term Hunger

To evaluate the effect of CRS interventions on hunger, the endline analysis investigates two measures: minimum dietary diversity (MDD) and minimum meal frequency (MMF). The data for these measures was collected from parents, not students. These two measures are then used to calculate the minimum acceptable diet (MAD). MAD refers to the proportion of children 6–23 months of age who received the MDD and MMF the previous day.[30]

MDD refers to the proportion of children 6–23 months of age who receive food from five or more food groups. Parents were asked, *"Now I want you to take a minute and think about all the food prepared yesterday for your youngest child over 6 months of age and under 2 years of age who eats solid foods. Did you give..?:"*[31] Observations were coded as meeting the MDD requirements if parents listed five or more food groups (Table 15). The proportion of children 6-23 months of age meeting this definition statistically significantly declined between midterm (30.5 percent) and endline (21.5 percent). This appears driven by reductions in the food groups of grains, roots, tubers, plantains, pulses, nuts, and seeds. The difference between endline and baseline (26.1 percent) was also statistically significant, driven by reductions in grains, roots, tubers, and plantains; dairy products; and eggs.

MMF refers to the proportion of breastfed and non-breastfed children 6–23 months of age who receive solid, semi-solid, or soft foods (but also including milk feeds for non-breastfed children) the minimum number of times or more.[32] Significantly more parents indicated that they breastfed their infant at midterm (98.2 percent) and endline (96.9 percent) than baseline (90.0 percent). However,

[30] With the exception of at least two milk feeds for non-breastfed children, as the data collection tools did not capture this frequency.
[31] Participants were asked to consider only one child during this portion of the survey. Answers were analyzed into the following 8 food groups: 1. breast milk; 2. grains, roots, tubers and plantains; 3. pulses (beans, peas, lentils), nuts and seeds; 4. dairy products (milk, infant formula, yogurt, cheese); 5. flesh foods (meat, fish, poultry, organ meats); 6. eggs; 7. vitamin-A rich fruits and vegetables; and 8. other fruits and vegetables.
[32] The minimum number of times is considered: 2 times for breastfed infants 6–8 months, 3 times for breastfed children 9–23 months, 4 times for non-breastfed children 6–23 months.

the average number of times parents reported their child eating solid, semi-solid, or soft foods decreased at the endline (2.5 times) versus both baseline (3.2 times) and midterm (3.3 times).

The percentage of children between 6–23 months receiving a MAD was calculated by combining the frequency and diversity of children's diets. At the endline, 12.4 percent of parents had children who met the MAD threshold, a significant decrease compared to both baseline (20.1 percent) and midterm (20.3 percent). Qualitative data included no reference to children's diets getting worse at any point. Rather, multiple respondents described improvements in nutrition and school feeding, while no respondent reported diets deteriorating.

Table 15: Minimum Dietary Diversity

Measure	Baseline	Midterm	Endline
MDD: Children 6–23 months of age who receive foods from 5 or more food groups of 8	26.1%	30.5%	21.5% [a, b]
MMF: Was the child breastfed yesterday, during the day, or at night?	90.0%	98.2%	96.9% [a]
MMF: How many times did a child eat solid, semi-solid, or soft foods other than liquids yesterday during the day or at night? (average # of times)	3.2	3.3	2.5 [a, b]
Children 6-23 months of age who receive Minimum Acceptable Diet (MDD & MMF)	20.1%	20.3%	12.4% [a, b]

Note: An "a" indicates statistically significant difference from baseline; a "b" indicates statistically significant difference from midline.

IR 1.3. Improved Student Attendance

School enrollment figures were collected as part of the head teacher survey, while enumerators recorded school attendance by counting the number of students in class as part of classroom and school observations. These student attendance and enrollment measures were used to determine the project schools' average student attendance rate. Overall, there are significant improvements from the baseline in average attendance rates in total for both girls and boys. Specifically, boys' attendance increased from 85.7 percent at baseline to 90.7 percent at endline, and girls' attendance similarly increased from 83.5 percent to 88.9 percent. These differences, as well as the overall difference across genders, are statistically significant.

Table 16: Average Attendance Rate by School

Gender	Baseline (N=77)	Midterm (N=80)	Endline (N=80)
Boy	85.7%	88.0%	90.7%*
Girl	83.5%	86.3%	88.9%*
Total	84.7%	87.2%	89.9%*

Note: This table presents the unweighted attendance data from baseline to endline. An asterisk (*) indicates a significant difference between baseline and endline observations at the 95 percent threshold.

IR 1.3.2. Reduced Health-Related Absences

For the parent survey, which was administered to three parents at each school, respondents were asked about student absences over the past month and the cause. As shown in Table 17, at the endline, only 11.2 percent of parents stated that at least one of their children missed school in the past month, a statistically significant decrease from midterm (37.1 percent) and baseline (20.7 percent). Of those who stated that their child missed school, 90.5 percent noted that it was because

of illness, equivalent to 10.2 percent of all parents reporting that their child missed school over the past month due to illness (11.2 percent x 90.5 percent = 10.2 percent). This is a significant decrease from midterm (34.9 percent) and baseline (14.9 percent).

Table 17: Parent Responses to Reasons for Child Absence

Stage	*Have any of your children missed school in the past month?*		*Of the children who missed school, was it because of illness?*		Percentage of parents who state their children had health-related school absences in the previous month
Baseline	20.7%	x	71.8%	=	14.9%
Midterm	37.1%	x	94.1%	=	34.9%
Endline	11.2% [a, b]	x	90.5% [a]	=	10.2% [a, b]

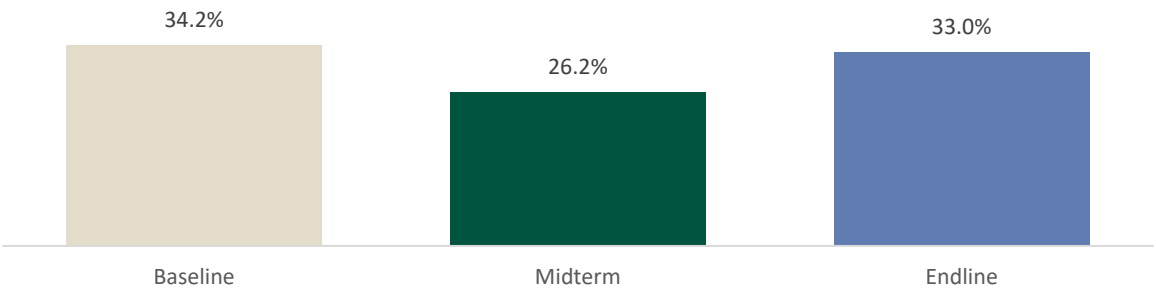
Note: An "a" indicates a statistically significant difference from baseline; a "b" indicates statistically significant difference from midline.

IR 1.3.5. Increased Community Understanding of the Benefits of Education

Percentage of caregivers spending time on literacy activities with their children in the previous week

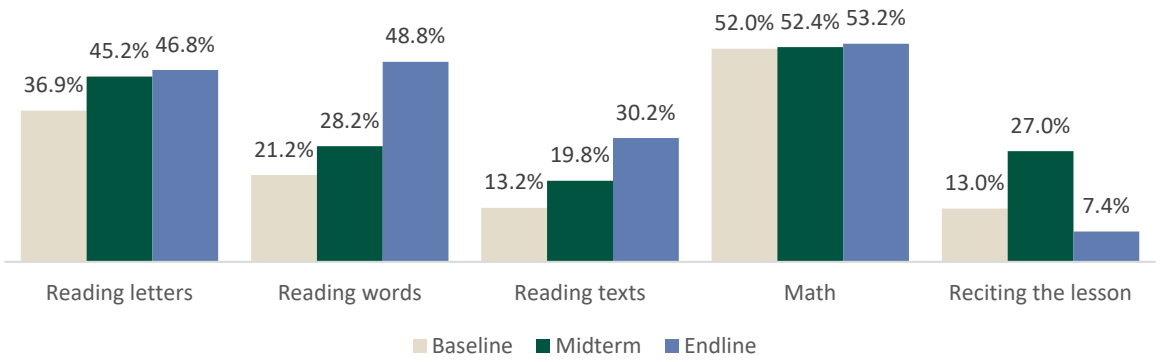
Enumerators asked parents and caregivers whether they supported their children's learning and engaged in literacy activities at home. STS first examined the percentage of caregivers who reported spending time on literacy activities with their school-age children the previous week. The proportion of parents who supported their children's learning and engaged in literacy activities at home by helping them with their homework in the last week increased significantly from midterm (26.2 percent) to endline (33.0 percent), as shown in Figure 10. The endline value is statistically like the baseline value (34.2 percent).

Figure 10: Percentage of Parents who Answered "Yes" to "Did You Help Your Children with Homework in the Past Week?"



Among the parents who reported helping their children with their homework in the past week, the endline showed a trend toward more reading words and texts and less reciting of the lesson, while reading letters and math levels remained similar to midterm (see Figure 11).

Figure 11: Percentage of Parents who helped with types of Literacy Activity in the Past Week



Percentage of community members who promote early childhood practices and support their children's education

For the broader indicator of the percentage of community members who promote early childhood practices and support their children's education, STS looked across the entire sample of parents and caregivers by calculating the percentage of parents who participated in broader at-home education activities *beyond* the past week. These activities included the following four things:

1. Telling stories to children
2. Having children read aloud to parents
3. Asking children what they learned in school
4. Helping children with their homework or having another family member help with homework

Just over half of respondents—54.3 percent—reported having participated in three or more of these education activities with their child or children at home, as shown in Figure 12. This is a statistically insignificant change from midterm (60.3 percent) and baseline (49.4 percent).

Figure 12: Parents Practicing at least 3 of 4 Education Activities with Their Children

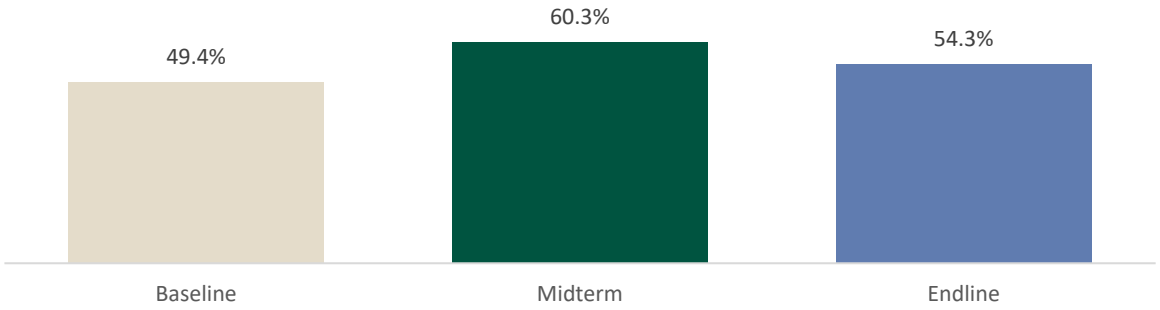
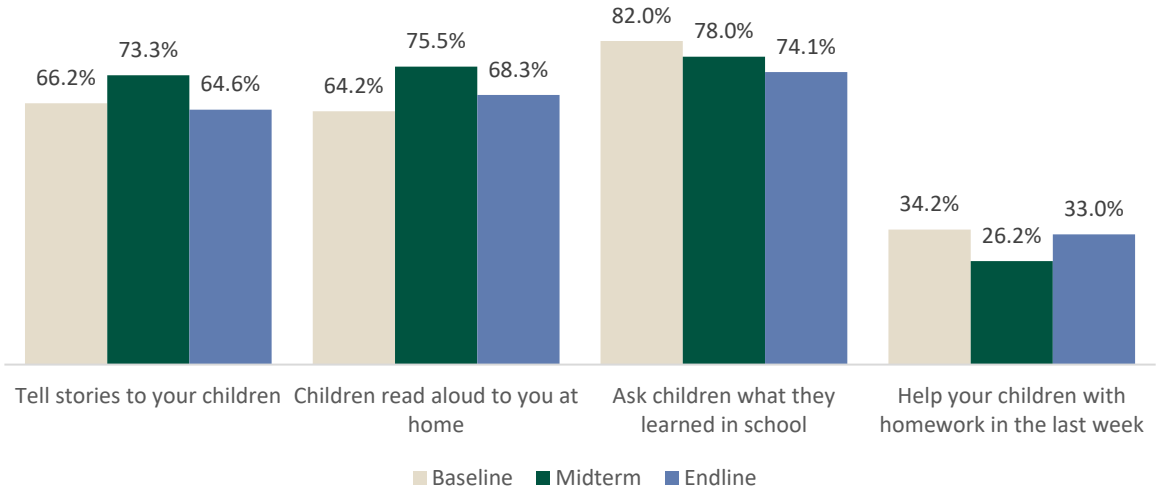


Figure 13 shows that the composition of educational activities largely reverted to baseline patterns, except for a significant decrease in asking children what they learned in school.

Figure 13: Parents Practicing Specific Education Activities with their Children



Correlational Analysis: What Drives Literacy Outcomes

Additional analyses using the student survey and parent survey were conducted to investigate drivers of literacy outcomes in the endline sample. The analysis utilized variables such as whether teachers help students improve academically, the average school attendance rate, whether parents or guardians speak French, and the availability of a latrine, books, and electricity at home, as well as whether students feel that what they learn at school applies to their daily lives. Weighted ordinary least squares regressions were performed on each subtask reported in Table 18. The level to which a student agrees that their teacher helps them is significantly correlated with higher literacy scores across all subtasks. Alternatively, results mostly suggest that these observational measures related to a student's living situation are not correlated with literacy outcomes.

Similar to midterm, the most substantial relationship seen in the endline sample is between student perception of teacher helpfulness. Students were asked, "Do your teachers help you do better in school? If yes, how often does the teacher help you?"[33] Students who reported that their teachers frequently helped them do better at school were more likely to have higher scores on all the subtasks. Notably, however, the school attendance rate and literacy subtasks results indicate a weak negative correlation, suggesting that higher Listening Comprehension values are associated with slightly lower school attendance rates.

Considering whether a student's parent speaks French, analysis shows significant relationship to letter sound score increases. Having a latrine at home, a proxy for economic status, does not correlate with literacy outcomes. Having books at home is not correlated with any subtasks. Having electricity at home, among the endline sample, did not correlate with the subtasks. Lastly, when looking at whether students believe the information they are learning in school is helpful in their daily lives, higher levels of perceived usefulness are related only to higher scores on Listening Comprehension.

[33] Answer outcomes: The teachers do not help you. The teachers help you sometimes. The teachers help you most of the time. The teachers help you all the time.

Table 18: Regression Analyses on the Relationship between Observational Data and Literacy Outcomes[34]

	Initial Sound	Letter Sound	Nonword Reading	Oral Reading Fluency	Reading Comprehension	Listening Comprehension
Do your teachers help you do better in school?	Y (Positive)	Y (Positive)	Y (Positive)	Y (Positive)	Y (Positive)	Y (Positive)
Average School Attendance Rate	NS	NS	NS	NS	NS	Y (Negative)
Do your parents / guardians speak French?	NS	Y (Positive)	NS	NS	NS	NS
At your home, is there a latrine?	NS	NS	NS	NS	NS	NS
At your home, are there books?	NS	NS	NS	NS	NS	NS
At your home, is there electricity?	NS	NS	NS	NS	NS	NS
Does what you learn at school help you in your daily life?	NS	NS	NS	NS	NS	Y (Positive)

Note: Y indicates a significant relationship; NS indicates that no statistical signification relationship was found.

In addition to the correlations described above, qualitative respondents pointed to several complementary factors driving the observed improvements in children's reading skills, including the provision of dedicated reading materials, enhanced teacher training, supervised follow-up, reading competitions, and incentives, and structured, hands-on instructional modules. Teachers and parents highlighted that supplying each pupil with their own leveled reading books ("*livrets*") and workbooks transformed the learning dynamic. As one head teacher explained, having individual booklets meant that: "*The books and work documents have improved the kids. Especially in reading. [...] For those*

[34] OLS regressions were conducted between students' observational data and literacy outcomes by gender. Y indicates that, yes, they are significantly correlated at the 95 percent confidence threshold. N indicates that, no, they are not correlated at the 95 percent confidence threshold.

who know how to read, when the child comes back with the booklet at the end of the term, you can already see your child's work [...]." With printed texts in hand, children could practice daily, and teachers could track progress one-on-one.

After initial workshops, the Ministry of Education and inspection teams began conducting "suivi sanctionné," tying teachers' application of new methods to their school's performance rating: "…if the ministry says, well, you're going to grade them. Those who have implemented it well... the school is performing well. If they haven't, the school isn't performing…" This accountability mechanism encouraged teachers to embed the new reading and writing pedagogies in every class. Regular "école performante" contests—complete with prizes for the top two readers—fostered friendly rivalry and kept pupils motivated, as one teacher reported: "there's always been a competition [...] we tell them that whoever is always first or second can have this or that gift. So that motivates them."

Finally, structured, hands-on instructional approaches with teachers practicing lesson simulations and peer reviews before rolling them out in class were cited as improving literacy. For example, a teacher reported "In the first year, I learn to read and write. Second year, I'm learning to speak [...] we simulated a teacher giving a lesson and a director coming along and interviewing him at the end of his performance."

Special Study: Body Mass Index and Literacy Performance

In line with the Evaluation Plan, at the endline, enumerators were able to collect students' weight and height, in addition to the learning assessment data and observational data. Valid age and BMI data was available for 1,503 students with an average age of 8.8 years. Body mass index (BMI) was calculated using the height and weight measurements and then compared to the body mass index-for-age scale. Children who are considered underweight have a BMI-for-age under the fifth percentile.[35] The fifth percentile for BMI-for-age is listed in Table 19. Underweight was calculated based on CDC percentile ranges for children's age rounded down to the nearest year of age. Children under age 9 were considered underweight if their BMI was below 13.5. Children aged 9 years 0 months to 9 years 11 months were considered underweight if their BMI was below 13.75. Children aged 10 or over were considered underweight if their BMI was below 14.

Table 19: BMI-for-age Underweight by Age and Gender[36]

Gender	Age	5th Percentile BMI
Girls	8	13.5
	9	13.8
	10	14.0
Boys	8	13.6
	9	13.8
	10	14.0

On average, BMI scores for students measured at endline were not considered underweight. Girls, on average, had an average BMI of 14.6 (listed in Table 20), which is higher than the highest threshold of the fifth-percentile cutoff of age 10 girls, which is a BMI of 14. Boys had an average BMI

[35] Developed by the National Center for Health Statistics in collaboration with the National Center for Chronic Disease Prevention and Health Promotion (2000).
[36] Source: For Girls https://www.cdc.gov/growthcharts/data/set2/chart-16.pdf; For Boys https://www.bcm.edu/cnrc-apps/bodycomp/cdcBMIboys.pdf

of 14.9 (listed in Table 20), higher than the highest fifth (5th) percentile cutoff for age 10 boys: a BMI of 14. Boys' BMI scores were significantly higher than girls' BMI scores.

Table 20: Student Weight, Height, and BMI by Gender[37]

	Average Weight	Average Height	Average BMI
Girls	24.9 kg	130.1 cm	14.6
Boys	25.3 kg	129.7 cm	14.9
Total[38]	25.2 kg	130.0 cm	14.8

However, some students in this study fell below this 5th percentile threshold. A minority (12.2 percent) of students sampled had BMIs below the threshold for their age—183 students, including 99 girls and 84 boys. However, at 12.2 percent, this minority is still more than twice what would be expected in the broader population (5.0 percent). In addition, about 4.3 percent of students had BMI values below 10. These extremely low values were excluded from further analysis, as they are considered biologically implausible and likely reflect data entry error. Given the overall low reading performance observed across the sample, it is unlikely that nutritional status would emerge as a major predictor of literacy outcomes in this context. Therefore, the BMI analysis below is treated as exploratory, and the associated findings should be interpreted with caution regarding both data limitations and contextual factors such as school quality.

BMI was not correlated with students' literacy outcomes at the endline. Weighted ordinary least squares regressions were performed on each subtask disaggregated by gender (Table 21). As indicated by "NS," no regression analyses resulted in a relationship with a statistical probability exceeding the 95 percent confidence level.

Table 21: Regression Analyses on the Relationship between BMI and Literacy Outcomes[39]

	Initial Sound	Letter Sound	Nonword Reading	Oral Reading Fluency	Reading Comprehension	Listening Comprehension
BMI (Girls)	NS	NS	NS	NS	NS	NS
BMI (Boys)	NS	NS	NS	NS	NS	NS
BMI (Total)	NS	NS	NS	NS	NS	NS

Note: NS indicates that no statistical signification relationship was found.

3.3 SO 2: Communities in the Savanes and Kara Regions have increased use of improved health, nutrition, and dietary practices

Sanitation facilities have significantly improved from baseline to endline at the 79 sample schools. Findings on sanitation facilities at the 79 sampled schools are presented in Table 22. At endline, the number of schools with no toilets available dropped from 29 to 15. The number of schools with pit latrines or buckets stayed the same (38 at all timepoints). The number of schools with composting toilets increased from 10 to 26 from baseline to endline.

[37] Weight, height, and BMI data in this table is weighted. BMI was calculated using the following formula: [weight (kg) / height (cm) / height (cm)] x 10,000. Students whose height was recorded under 70 cm were dropped with the assumption of data collection error.
[38] There is a statistically significant difference between the BMI scores of girls and boys (p=.018).
[39] OLS regressions were conducted between students' BMI and literacy outcomes by gender. Y indicates that, yes, they are significantly correlated at the 95 percent confidence threshold. N indicates that, no, they are not correlated at the 95 percent confidence threshold.

Table 22: Sanitation Facilities at Sampled Schools, Availability of Toilets

	Baseline		Midline		Endline	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
No toilets available (only in the bush or in the fields)	29	37.7%	24	29.4%	15	21.1%
The toilets are pit latrines or buckets	38	49.4%	38	46.9%	38	41.2%
The toilets are composting toilets	10	13.0%	18	23.7%	26	37.7%
Total	77		80		79	

Note: Percentages reflect weighted totals, and frequencies depict unweighted totals. Due to weights, the percentage totals may not equal 100. Regression analysis was used to determine statistical significance between baseline and midterm using weights.

Additionally, the functionality of sanitation facilities at the observed schools has decreased from baseline to endline. As shown in Table 23, about 84.3 percent of toilets were functional at the endline, compared to 93.8 percent.[40]

Table 23: Sanitation Facilities at Sampled Schools, Functionality of Toilets

	Baseline		Midline*		Endline	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Not functional	3	6.3%	7	13.1%	11	15.6%
Functional	45	93.8%	49	86.9%	53	84.3%
Total	48		56		63	

Note: Percentages reflect weighted totals, and frequencies depict unweighted totals. Due to weights, the percentage totals may not equal 100. Regression analysis was used to determine statistical significance between baseline and midterm using weights. Enumerators were asked to verify the source and indicate if it was functional.

More functional toilets were available at endline. The expansion of latrines throughout the project more than offsets the decline in the percentages of functionality. As shown in Table 24, 53 schools had acceptable and functional sanitation services at endline, compared to only 45 at baseline—an eight-percentage point increase.

Table 24: Sanitation Facilities at Sampled Schools, Availability of Functional of Toilets

	Baseline		Midline*		Endline	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Toilets not available or not functional	32	41.6%	31	38.8%	26	32.9%
Toilets available and functional	45	58.4%	49	61.2%	53	67.1%
Total	77		80		79	

[40] Schools observed with no toilets available are excluded from this further analysis.

Handwashing systems have also improved in sampled schools since baseline. As detailed in Table 25, the percentage of schools with running water or a hand pour system and soap increased from 36.4 percent at baseline to 48.7 percent at midterm and decreased at endline to 40.6 percent. Further, the proportion of schools with no handwashing option during the observation decreased by 0.5 percentage points from baseline to endline.

Table 25: Handwashing Facilities at Sampled Schools

	Baseline		Midline		Endline	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
No handwashing station at the school	24	31.2%	21	24.1%	27	30.7%
Shared basin or bucket (handwashing is done in water; water does not flow or is not poured)	8	10.4%	11	12.2%	6	8.2%
Hand pouring system with used water separated from water to clean hands but without soap	17	22.1%	12	15.0%	17	20.5%
There is running water OR a hand pour system (with the wastewater separated from the clean water for washing hands) AND soap	28	36.4%	36	48.7%	29	40.6%
Total	77	-	80	-	79	

Note: Percentages reflect weighted totals, and frequencies depict unweighted totals. Due to weights, the percentage totals may not equal 100. Regression analysis was used to determine statistical significance between baseline and midterm using weights.

There was no significant change in the level of accessibility in handwashing stations in sampled schools since baseline. Of the 52 schools that did have some form of handwashing station at endline, the percentage of those where handwashing was not accessible to the youngest children or children with disabilities decreased from 7.6 percent at baseline to 5.0 percent at endline. However, this difference was not statistically significant (shown in Table 26).

Table 26: Accessibility of Handwashing Facilities at Sampled Schools

Handwashing Facilities	Baseline		Midline		Endline	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Not accessible to the youngest	4	7.6%	8	15.6%	3	5.0%

Handwashing Facilities	Baseline		Midline		Endline	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
children or children with disabilities						
Accessible to the youngest children OR children with disabilities	6	11.2%	2	3.3%	13	22.3%
Accessible to the youngest children AND children with disabilities	43	81.1%	49	81.1%	36	72.6%
Total	53	-	59	-	52	

Note: Percentages reflect weighted totals, and frequencies depict unweighted totals. Due to weights, the percentage totals may not equal 100. Regression analysis was used to determine statistical significance between baseline and midterm using weights.

IR 2.5: Number of schools using an improved water source

The proportion of schools with an improved water source increased nearly 13 percentage points from baseline to endline—from 33.8 percent to 47.9 percent. As shown in Table 27, the increase in proportion was statistically significant. The proportion of schools with no water available decreased significantly at the endline, from 58.4 percent at baseline to 33.7 percent at the endline.

Table 27: Water Sources at Sampled Schools

	Baseline		Midline		Endline*	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
No water available at school. Water, if present, is provided by parents, children, or staff	45	58.4%	42	54.3%	25	33.7%
Available water is: Unprotected inground well / spring, untreated rainwater, surface water	6	7.8%	0	0.0%	7	7.9%
Available water is a cart with a small tank / drum or a protected spring	0	0.00%	1	1.6%	11	10.5%
The available source of sanitary water is running water, a public tap, treated	26	33.8%	37	44.1%	36	47.9%

	Baseline		Midline		Endline*	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
rainwater, a protected dug well, or bottled water						
Total	77	-	80	-	79	

Note: Percentages reflect weighted totals, and frequencies depict unweighted totals. Due to weights, the percentage totals may not equal 100. Regression analysis was used to determine statistical significance between baseline and midterm using weights.

The functionality of water sources significantly improved from baseline to endline. Of the 54 schools with water sources observed at the endline, 100 percent were functioning, as shown in Table 28. This was a significant improvement from the baseline.

Table 28: Status of Water Source

	Baseline		Midline*		Endline*	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Not Functional	8	25.0%	2	4.1%	0	0.0%
Functional	24	75.0%	36	95.9%	54	100.0%
Total	32	-	38	-	54	

Note: Percentages reflect weighted totals, and frequencies depict unweighted totals. Due to weights, the percentage totals may not equal 100. Regression analysis was used to determine statistical significance between baseline and midterm using weights. Enumerators were asked to verify the source and indicate if it was functional.

Sustainability

Qualitative data collection revealed strong sentiments held by respondents that many of the project's achievements were being sustained locally due to a combination of community ownership, institutional linkages, and gradual handover to state structures. Four main themes predominate: community-driven continuity, infrastructure maintenance, institutional embedding, and remaining challenges with recommendations for deepening sustainability.

A recurring point was how beneficiary groups have appropriated activities so fully that they "continue, and will continue until 1000 years," as one community member put it. Likewise, the "mères lumière" trained in nutrition, hygiene and child-feeding practices continue to meet and sensitize their peers autonomously: "[...] despite the fact that the project has come to an end, there are a good number of women who continue to raise awareness among their peers [...] so that these good practices in the community can be promoted and kept up every day."

Saving-and-credit groups were viewed as bolstering financial resilience by mobilizing internal loans and supporting local education and health expenditures. As one community member noted, "Now people save, women save to meet their children's needs at school, for health care." Multiple categories of respondent reported that such peer-to-peer mechanisms have become self-reproducing, requiring minimal ongoing outside support.

Many respondents highlighted that durable, high-quality infrastructure, such as canteen kitchens built to international hygiene standards and storage facilities have fostered enduring practices in sanitation and school feeding. Community committees were also reported as remaining actively engaged in upkeep. One mayor emphasized: "When you see the infrastructures built and people continue to maintain them, even though the funding has ended, they continue to maintain them."

Many respondents called for deeper collaboration with decentralized state structures to avoid “repetition of past mistakes” and secure national recognition of successful approaches. One manager observed: “Given the lessons learned [...] intensify collaboration with the decentralized structures of the State and even see the State involve them more [...] for the success of the project.” The continuity of hot-meal provision through the follow-on project until 2028 exemplifies this institutional bridge, according to a leader in one of the local communities.

4. Evaluation Questions

This section responds to the evaluation questions investigated and documented in the endline evaluation drawn from the findings.

Question 4 in the Learning Agenda’s Health Evidence Gaps section: “What systems of community health care governance are the most effective at sustaining the delivery of health interventions through school meal programs?”

As reported by qualitative data collection respondents, stakeholders’ and volunteers’ willingness to participate without financial expectations was a key strength of community care governance. Multiple respondents reported that community members and project volunteers continued to engage in project activities "without receiving anything in return." For example, in discussing the “mères lumière,” one facilitator explained that the initial framing clarified that the “motivations” offered were not cash payments. As a result, "in the community, without incentives, the work will be done. That’s what was done with those lumières." Similarly, when probing COJEP volunteers, one interviewer asked why they remained engaged. The response highlighted alignment with an existing government decree that envisioned these committees as unpaid, "essential for the functioning of schools," and that sustaining this approach would ensure permanence beyond the project’s lifespan:"...if we align ourselves behind this strategy [...], it could be good, and it must be sustainable." Finally, in a direct question on readiness to work without reward, a COJEP member unambiguously stated: "Are you ready to work without expecting anything in return? Yes, for the good of the whole community."

Additionally, teachers and Head Teachers have shown themselves as effective actors in their roles as educators. At the endline, they were observed in higher numbers to be using quality teaching practices and supervision tools.

Students are still experiencing high rates of school absences due to illness, as reported by parents. While illness cannot always be prevented, the project could consider including handwashing and hygiene education programs and materials to be shared with parents and schools. This could increase the effectiveness of sustaining the delivery of health interventions.

The role of parents as actors in promoting and sustaining the delivery of health interventions could be further strengthened. More parents could be encouraged to participate in additional educational activities at home with their children.

Question 5 in the Learning Agenda’s Education/Literacy Evidence Gaps section: “What are the differences in educational outcomes from school meal programs between malnourished or undernourished children and those who are not?”

Looking at the results of the special study, there is no correlation between BMI and any of the learning outcomes measured during this evaluation. Regression analyses were performed to study

the correlation between BMI and subtask scores, and no relationship was found between increased BMI and higher learning outcomes.

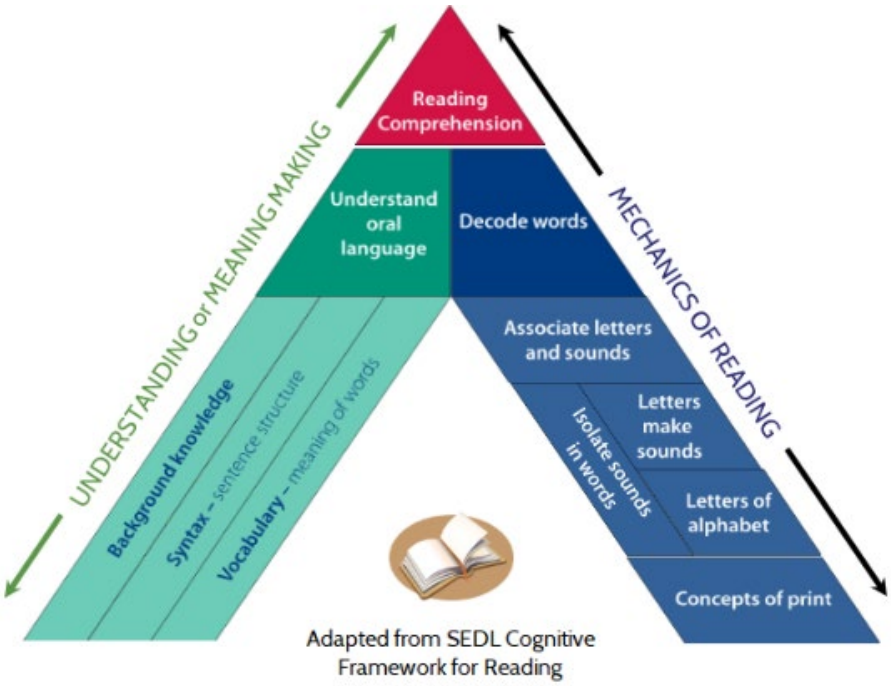
As suggested at midterm, the comparison between malnourished and undernourished is potentially misaligned with the project context. As the BMI numbers suggest, only a small portion of students fall into the under-nourished category, let alone present as malnourished.

5. Analysis of Findings and Conclusions

Strategic Objective 1: School-Age Children in the Savanes and Kara Regions Have Improved Literacy

By comparing the results of this endline evaluation to the baseline study, the STARS project's impact on students' progression in their fundamental reading skills has been examined, as measured by the EGRA subtasks. Using the Southwest Educational Development Laboratory's (SEDL) Cognitive Framework for Reading, it is possible to map EGRA subtasks to reading skills as follows:[41]

Figure 14: Reading Skills Framework with EGRA Subtask Mapping



Understanding or Meaning Making	Mechanics of Reading	Reading Comprehension
• Listening Comprehension	• Initial Sound Identification • Letter Sound Identification • Nonword Reading • Oral reading fluency	• Reading Comprehension

Students completed a Listening Comprehension subtask, which assessed students' basic understanding or meaning-making abilities in French. For this subtask, which consisted of five

[41] Sebastian Wren, The Cognitive Foundations of Learning to Read: A Framework. Southwest Educational Development Laboratory, 2001. https://sedl.org/reading/framework/framework.pdf

questions about a story read aloud in French, students answered only 0.25 questions correctly on average. This indicates that the CE1 student population has a very limited ability to understand French. Furthermore, listening comprehension among sampled students has not significantly improved since baseline. Listening comprehension is often considered a forerunner in measuring children's overall language comprehension and vocabulary, which is critical to reading. Without this basic and fundamental language skill, children cannot easily transition to applying their knowledge of letter sounds (phonemic awareness) to decoding.

Four EGRA subtasks speak to students' abilities with the mechanics of reading. Students must master these necessary building blocks to progress to reading comprehension. Literacy and reading instruction in the early grades—including those targeted by the STARS project—often focus on these skills. On average, students correctly responded to about two out of 10 items on the Initial Sound Identification subtask, a significant increase from baseline but still ultimately low. On the Letter Sound Identification subtask, students correctly identified 9.2 letters out of 100 in one minute, on average, which was a significant increase from baseline. At the endline, students read 2.4 words on average, an increase from 1.6 at baseline. This difference was not statistically significant; however, boys showed a statistically significant increase from baseline. Grade 3 students at the endline have improved on some of the essential lower-level reading skills, but they had considerable opportunity to improve these skills in these areas, especially considering the large proportion of zero scores on these subtasks.

The final subtask—Reading Comprehension—measures students' ability to utilize the mechanics of reading, demonstrate fluency, and understand a passage's meaning. It is the most advanced EGRA subtask, as it measures the goal of literacy: comprehension. Similar to their baseline and midterm performance, grade 3 students scored low on Reading Comprehension at the endline. On average, students did not correctly answer a single reading comprehension question, with the average number of questions correctly answered only 0.06. Reading Comprehension did not significantly improve from baseline to endline.

The proportion of students unable to provide a single correct response on each subtask was often high. On the Initial Sound Identification subtask, 57.5 percent of students did not correctly respond to even one of the five items. Still, there was a significant decrease from the baseline. The Letter Sound Identification subtask had the lowest proportion of students with a zero score, with only 19 percent of students not being able to correctly identify at least one letter sound in one minute. This was a significant decrease since baseline. These significant changes are mechanical reflections of the changes in mean scores on these lower-level literacy skills. When presented with a reading passage, 56 percent of students did not read a single word. Linked to the reading passage subtask, the reading comprehension questions also had a high number of zero scores, as 95 percent of students did not correctly answer a single reading comprehension question. In terms of Listening Comprehension, 81 percent of students did not answer a single question correctly.

The project's largest impact can be seen in lower-level reading skills, the best practices used by school personnel, and the improvement of school sanitation facilities. The literacy findings suggest that the project interventions have had the greatest impact on lower-level literacy skills and are beginning to show effects on the higher-order skill of word reading. Both boys and girls were significantly less likely to receive zero scores—to not answer a single item correctly on a subtask—on the Letter Sound Identification, Initial Sound Identification, and Oral Reading Fluency subtasks. Boys have also shown a significant increase in mean scores on Oral Reading Fluency since baseline—the

proportion of teachers demonstrating quality teaching practices during lessons improved from baseline to endline. Further, quality supervision tools used at schools increased from baseline to endline. In parallel, there was also an observed change in student behavior with higher attendance levels. Lastly, school infrastructure significantly improved with greater latrines and running water access.

The evaluation findings demonstrate significant improvements in sanitation and water facilities across the 80 sampled schools, reflecting the positive outcomes of the project's efforts for strategic objective 2. The availability of sanitation facilities has notably increased, with the number of schools without toilets decreasing from 29 at baseline to 15 at endline. Additionally, the number of schools with composting toilets grew from 10 to 26, indicating successful infrastructure upgrades. The overall availability and functionality of sanitation facilities have improved, with 67.1 percent of schools having functional and available toilets at the endline, a notable increase from 58.4 percent at baseline.

Strategic Objective 2: Communities in the Savanes and Kara Regions Have Improved Use of Health, Nutrition and Dietary Practices

Water sources in sampled schools also saw marked improvements. The proportion of schools with an improved water source increased from 33.8 percent at baseline to 47.9 percent at endline, with the percentage of schools lacking water dropping significantly from 58.4 percent to 33.7 percent. Furthermore, all 54 water sources observed at endline were functional, a significant improvement from baseline when only 75 percent of sources were operational. These findings highlight the successful expansion and maintenance of water infrastructure, although ongoing efforts to sustain water and sanitation services are essential.

Handwashing practices have also been promoted within the schools, with qualitative data indicating that the introduction of latrines helped encourage handwashing. However, challenges remain, particularly in schools lacking clean, running water and sufficient handwashing stations. At endline, 40.6 percent of schools had handwashing systems with running water and soap, a slight decrease from midline. Nonetheless, the proportion of schools without handwashing stations decreased slightly, indicating progress. Despite improvements in accessibility, some children, particularly the youngest or those with disabilities, still face challenges in accessing handwashing stations, although the proportion of these cases decreased from 7.6 percent at baseline to 5.0 percent at endline.

In conclusion, the project has successfully contributed to improving sanitation facilities, water sources, and handwashing practices in the sampled schools. These advancements have positively impacted sanitary practices among school-aged children. However, further attention is needed to address the functionality of sanitation facilities and the accessibility of handwashing stations, particularly in schools with limited water access. Continued investment in these areas is crucial for sustaining the health benefits observed and for ensuring that all students have access to clean and functional sanitation and water resources.

6. Recommendations

STS proposes the following recommendations for CRS for project implementation.

Examine existing student and teacher French language abilities.

Overall student performance, particularly in Listening Comprehension, indicates that students have a limited ability to understand spoken French, affecting their reading outcomes. Food for Education

projects may want to consider undertaking more targeted research into the reasons for this gap in comprehension, and future programming should take into account the body of research that students learn to read better when they start in their mother language before transitioning to a second language.[42] Additionally, future data collection should consider strategies to ensure students understand what is being asked of them for any survey or assessment where the questions are in French.

Increase students' exposure to French in all settings to increase literacy levels.

EGRA results also show that students have improved their lower-level reading skills and are beginning to show signs of higher-order reading skills since the STARS project began. Additional training, more time, and introducing a bilingual pilot may be necessary to continue enhancing literacy outcomes. Food for Education projects may want to consider interventions aimed at increasing literacy need to center around increasing instructional time during the day devoted to reading in school. Importantly, this reading needs to be done in French:

- One strategy to increase reading time during the day would be to engage with parents and guardians to encourage reading in French at home. For households who are fluent in French, co-reading should be integrated into daily home habits. The project might support this by distributing reading materials to students' families. In households where parents or guardians are uncomfortable using French, dual language materials, including both French and local language translation, could be created to support reading in the home. A potential missing actor that could be brought in to increase students' reading exposure is a sibling; they may have more fluency with French and could be encouraged to participate in co-reading.
- Another recommendation is to encourage teachers to collaborate across subjects to incorporate reading into other subjects, such as mathematics. For example, word problems written in French would help increase students' instructional time reading during the day.
- A large component of reading fluency and comprehension is vocabulary. Teacher trainings, materials, and instructional time should prioritize vocabulary in French. Materials could be developed in local languages and French to support this development within the classroom and if provided to families at home.

Bilingual Education Programs and TaRL

For students who are not fully proficient in the language of instruction, implementing bilingual education programs could help bridge the gap between their native language (L1) and the language of instruction (L2). This approach would support improved comprehension and contribute to better learning outcomes.

Based on the correlational analysis, which identified that students who received more support from their teachers performed better, it is recommended to consider implementing a Teaching at the Right Level (TaRL) approach in future programs. This approach can effectively target struggling learners by tailoring instruction to their specific learning needs, thereby improving their French literacy levels. Additionally, follow-on interventions could include the introduction of reading clubs to encourage regular reading practice and foster a love for learning. Incorporating local language

[42] See, for example: UNESCO. (2008). *Improving the Quality of Mother Tongue-Based Literacy and Learning: Case Studies from Asia, Africa, and South America*. Paris: UNESCO.
Dutcher, N. (2005). *Expanding Educational Opportunity in Linguistically Diverse Societies*. Washington, DC: World Bank.
Heugh, K. (2006). *Theory and Practice – Language Education Models in Africa: Research, Design, Decision-Making, and Outcomes*. In Optimizing Learning and Education in Africa – The Language Factor (pp. 1–34). Paris: ADEA/UNESCO.

into Early Childhood Education (ECE) emergent literacy programs may also enhance school readiness, particularly for students who are not yet proficient in French. These interventions align with CRS's education sector goals and would further support the development of foundational literacy skills in students.

Further improvements in school water and sanitation sources are warranted.

At the endline, upgrades of water facilities remain necessary, although notable improvements in school facilities were observed. Project interventions could make an impact by improving handwashing facilities and improving their accessibility.

Collaborate more deeply with decentralized state structures.

Qualitative data collection revealed multiple respondents who called for deeper collaboration with decentralized state structures to avoid "repetition of past mistakes" and secure national recognition of successful approaches. One manager observed: "Given the lessons learned [...] intensify collaboration with the decentralized structures of the State and even see the State involve them more [...] for the success of the project."

Future S02 interventions should focus on nutrition for school-age children targeted by school feeding programs.

Food for Education projects should investigate why minimum dietary diversity (MDD) has fallen between baseline and endline and why minimum meal frequency (MMF) declined during the same period. Future interventions should consider educational and food provision components that target these dynamics. Further projects should consider the costs and benefits of investment in nutrition activities for children 6-23 months to determine whether that age range should be a project priority.

Examine gender constraints within target communities.

Girls' underperformance compared with boys deserves further exploration and may warrant a specific focus within the project to address the underlying causes of these gender disparities. When comparing baseline to endline, these gender gaps in learning outcomes appear to be either remaining stagnant or even growing. Future interventions, such as gender responsive pedagogy, should focus on resources to help close this gap—while accommodating the complexity of the fact that many boys are also behind the average girl.

Annexes

Annex A: Bibliography

McConnell, Brendon, and Marcos Vera-Hernandez. Going beyond simple sample size calculations: a practitioner's guide. Institute for Fiscal Studies. 2015.

UNICEF. Analyse du secteur de l'éducation de la République togolaise, Des défis pour un enseignement de qualité pour tous, République togolaise, IIPE-Pôle de Dakar - UNESCO, 2019.

United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects 2019, Online Edition. Rev. 1, 2019.

United States Department of Agriculture. "McGovern-Dole Food for Education Program," accessed January 20, 2021, https://www.fas.usda.gov/programs/mcgovern-dole-food-education-program

WHO/UNICEF. Joint Monitoring Programme for Water Supply, Sanitation and Hygiene (JMP). Last update: June 2019. https://data.unicef.org/resources/dataset/sowc-2019-statistical-tables/

World Bank. Togo Country Overview. Last update: November 2020. https://www.worldbank.org/en/country/togo/overview

Wren, Sebastian. The Cognitive Foundations of Learning to Read: A Framework. Southwest Educational Development Laboratory, 2001. https://sedl.org/reading/framework/framework.pdf

Annex B: Updated Indicator Performance Tracking Table

No.	Results framework statement	RF	Activity	Performance Indicator	Standard or CRS Custom	Baseline	Midterm	Endline
1	School-Age Children in the Savanes and Kara Regions Have Improved Literacy (SO 1)	SO1	Raising awareness on importance of education (Activity 12)	Percent of students who, by the end of two grades of primary schooling, demonstrate that they can read and understand the meaning of grade level text	Standard #1	0%	0%	0.2%
2	School-Age Children in the Savanes and Kara Regions Have Improved Literacy (SO 1)	SO1	Provide school meals (Activity 11)	Number of individuals benefiting indirectly from USDA-funded interventions	Standard #31	0	105,196	109,127
3	Communities in the Savanes and Kara Regions Have Increased Use of Improved Health, Nutrition and Dietary Practices (SO 2)	IR 2.1	Raise awareness on health, nutrition, and WASH (Activity 12)	Number of individuals who demonstrate use of new child health and nutrition practices as a result of USDA assistance	Standard #19	0%	0	6,610
4	Communities in the Savanes and Kara Regions Have Increased Use of Improved Health, Nutrition and Dietary Practices (SO 2)	SO2	Training: Food preparation and storage practices (Activity 15)	Number of individuals who demonstrate use of new safe food preparation and storage practices as a result of USDA assistance	Standard #20	0%	796	1,161
5	Communities in the Savanes and Kara Regions Have Increased Use of Improved Health, Nutrition and Dietary Practices (SO 2)	SO2	Provide school meals (Activity 11)	Number of individuals participating in USDA food security programs	Standard #30	0	78,430	77,123
6	Communities in the Savanes and Kara Regions Have Increased Use of Improved Health, Nutrition and Dietary Practices (SO 2)	SO2	Provide school meals (Activity 11)	Number of schools reached as a result of USDA assistance	Standard #32	0	138	138
7	Improved Quality of Literacy Instruction (IR 1.1)	IR 1.1	Training: Teachers (Activity 18)	Percent of teachers providing quality classroom instruction with USG support	USAID Education Proposed	0%	43.1%	70.6%

No.	Results framework statement	RF	Activity	Performance Indicator	Standard or CRS Custom	Baseline	Midterm	Endline
8	IR 1.2 Improved Attentiveness	IR 1.2	Provide school meals (Activity 11)	Percent of students in target schools identified as attentive during class/instruction	Custom	59.7%	74.9%	69.1%
9	Improved Student Attendance (IR 1.3)	IR 1.3	Take home rations (Activity 14)	Average student attendance rate in USDA supported classrooms/schools	Standard #2	80.2%	89.9%	89.9%
10	Increased Knowledge of Safe Food Prep and Storage Practices (IR 2.2)	IR 2.2	Training: Food preparation and storage practices (Activity 15)	Number of individuals trained in safe food preparation and storage as a result of USDA assistance	Standard #22	0	1,102	0
11	Improved Knowledge of Health and Hygiene Practices (IR 2.1)	2.1	Raise awareness on health, nutrition, and WASH (Activity 12)	Number of individuals trained in child health and nutrition as a result of USDA assistance	Standard #23	0	0	9,382
12	Increased Knowledge of Nutrition (IR 2.3)	IR 2.3	Raise awareness on health, nutrition, and WASH (Activity 12)	Number of children under five (0-59 months) reached with nutrition-specific interventions through USDA-supported programs	Standard #24	0	10,662	20,916
13	Increased Knowledge of Nutrition (IR 2.3)	IR 2.3	Raise awareness on health, nutrition, and WASH (Activity 12)	Number of pregnant women reached with nutrition-specific interventions through USDA-supported programs	Standard #26	0	3,688	1,447
14	Increased Access to Clean Water and Sanitation Services (IR 2.4)	IR 2.4	Raise awareness on health, nutrition, and WASH (Activity 12)	Number of children under two (0-23 months) reached with community-level nutrition interventions through USDA-supported programs	Standard #25	0	8,253	8,133
15	Increased Access to Clean Water and Sanitation	IR 2.4	Building/ Rehab:	Number of schools with improved sanitation facilities	Standard #28	57	87	116

No.	Results framework statement	RF	Activity	Performance Indicator	Standard or CRS Custom	Baseline	Midterm	Endline
	Services (IR 2.4)		Latrines (Activity 2)					
16	Increased Access to Clean Water and Sanitation Services (IR 2.4)	IR 2.5	Building/ Rehab: Wells and water stations/ systems (Activity 4)	Number of schools using an improved water source	Standard #27	70	94	107
17	Increased Access to Clean Water and Sanitation Services (IR 2.4)	IR 2.5	Building/ Rehab: Latrines (Activity 2)	Percent of health and nutrition infrastructure, constructed as a result of USDA assistance, maintained by communities/local authorities	Custom	0%	100%	100%
18	Increased Access to Requisite Food Prep and Storage Tools and Equipment (IR 2.6)	IR 2.6	Building/ Rehab: Kitchens (Activity 1)	Number of Schools receiving energy saving stoves	Custom	0	3	42
19	More Consistent Teacher Attendance (Sub-IR 1.1.1)	Sub-IR 1.1.1	Promote teacher attendance (Activity 10)	Percent of instructional time lost due to teacher absenteeism	USAID Education Proposed	9.3%	9.3%	12.7%
20	More Consistent Teacher Attendance (Sub-IR 1.1.1)	Sub-IR 1.1.1	Promote teacher attendance (Activity 10)	Number of schools implementing the use of school score cards	Custom	0%	0	138
21	Better Access to School Supplies and Materials (Sub-IR 1.1.2)	Sub-IR 1.1.2	Distribution School supplies and materials (Activity 6)	Number of teaching and learning materials provided as a result of USDA assistance	Standard #3	0	83,289	77,411
22	Increased Skills and Knowledge of Teachers (Sub-IR 1.1.4)	Sub-IR 1.1.4	Training: Teachers (Activity 18)	Number of teachers/educators/teaching assistants in target schools who demonstrate use of new and quality teaching techniques or tools as a result of USDA assistance	Standard #4	0%	313	460
23	Increased Skills and Knowledge of Teachers (Sub-IR 1.1.4)	Sub-IR 1.1.4	Training: Teachers (Activity 18)	Percentage of teachers/educators/teaching assistants in target schools who demonstrate the use of new and quality teaching techniques or tools as a result of USDA assistance	Custom	23.4%	43.1%	70.6%

No.	Results framework statement	RF	Activity	Performance Indicator	Standard or CRS Custom	Baseline	Midterm	Endline
24	Increased Skills and Knowledge of Teachers (Sub-IR 1.1.4)	Sub-IR 1.1.4	Training: Teachers (Activity 18)	Number of teachers/educators/teaching assistants trained or certified as a result of USDA assistance	Standard #5	0	421	0
25	Increased Skills and Knowledge of School Administrators (Sub-IR 1.1.5)	Sub-IR 1.1.5	Training: School admins (Activity 17)	Number of school administrators and officials in target schools who demonstrate the use of new techniques or tools as a result of USDA assistance	Standard #6	0%	31	50
26	Increased Skills and Knowledge of School Administrators (Sub-IR 1.1.5)	Sub-IR 1.1.5	Training: School admins (Activity 17)	Number of school administrators and officials trained or certified as a result of USDA assistance	Standard #7	0	146	0
27	Increased Skills and Knowledge of School Administrators (Sub-IR 1.1.5)	Sub-IR 1.1.5	Training: School admins (Activity 17)	Percent of school officials in target schools who demonstrate the use of new and quality supervision and leadership techniques or tools	Custom	6.5%	9.0%	50%
28	Reduced Short-Term Hunger (Sub-IR 1.2.1)	Sub-IR 1.2.1	Take home rations (Activity 14)	Percent of children 6–23 months receiving a minimum acceptable diet	FFP #BL12	20.1%	20.3%	12.4%
29	Increased Economic and Cultural Incentives (Sub-IR 1.3.1)	Sub-IR 1.3.1	Provide school meals (Activity 11)	Number of school-age children receiving daily school meals (breakfast, snack, lunch) as a result of USDA assistance	Standard #17	0	50,805	53,161
30	Reduced Health-Related Absences (Sub-IR 1.3.2)	Sub-IR 1.3.2	Raise awareness on health, nutrition, and WASH (Activity 12)	Percent of parents who state their children had health-related school absences in the previous month	Custom	15.0%	34.94%	10.2%
31	Improved School Infrastructure (Sub-IR 1.3.3)	Sub-IR 1.3.3	Building/ Rehab: Kitchens (Activity 1)	Number of educational facilities (i.e., school buildings, classrooms, improved water sources, and latrines) rehabilitated/constructed as a result of USDA assistance	Standard #8	0	224	0
31	Improved School Infrastructure (Sub-IR 1.3.3)	Sub-IR 1.3.3	Building/ Rehab: Kitchens (Activity 1)	Number of educational facilities (i.e., school buildings, classrooms, improved water sources, and latrines) rehabilitated/constructed as a result of USDA assistance **[Warehouses]**	Standard #8	0	119	0
31	Improved School Infrastructure (Sub-IR 1.3.3)	Sub-IR 1.3.3	Building/ Rehab:	Number of educational facilities (i.e., school buildings, classrooms, improved water sources, and latrines)	Standard #8	0	76	0

No.	Results framework statement	RF	Activity	Performance Indicator	Standard or CRS Custom	Baseline	Midterm	Endline
			Kitchens (Activity 1)	rehabilitated/constructed as a result of USDA assistance **[Kitchens, cook areas]**				
31	Improved School Infrastructure (Sub-IR 1.3.3)	Sub-IR 1.3.3	Building/ Rehab: Kitchens (Activity 1)	Number of educational facilities (i.e., school buildings, classrooms, improved water sources, and latrines) rehabilitated/constructed as a result of USDA assistance **[Latrines]**	Standard #8	0	16	0
31	Improved School Infrastructure (Sub-IR 1.3.3)	Sub-IR 1.3.3	Building/ Rehab: Kitchens (Activity 1)	Number of educational facilities (i.e., school buildings, classrooms, improved water sources, and latrines) rehabilitated/constructed as a result of USDA assistance **[Improved water sources]**	Standard #8	0	13	0
32	Increased Student Enrollment (Sub-IR 1.3.4)	Sub-IR 1.3.4	Raising awareness on the importance of education (Activity 13)	Number of students enrolled in school receiving USDA assistance	Standard #9	0%	50,805	53,161
33	Increased Student Enrollment (Sub-IR 1.3.4)	Sub-IR 1.3.4	Raising awareness on the importance of education (Activity 13)	Number of schools that held an enrollment campaign.	Custom	0	138	135
34	Increased Community Understanding of the Benefits of Education (Sub-IR 1.3.5)	Sub-IR 1.3.5	Establish activities to promote literacy (Activity 7)	Percent of caregivers who report spending time on literacy activities with their school-age children in the previous week	Custom	15.8%	26.26%	32.9%
35	Increased Community Understanding of the Benefits of Education (Sub-IR 1.3.5)	Sub-IR 1.3.5	Raising awareness on the importance of education (Activity 13)	Number School Management Committee (SMC) and Parent Teacher Association (APE) members, and Mother Leaders trained in activities to promote literacy	Custom	0	1,443	1,289
36	Increased Community Understanding of the Benefits of Education (Sub-IR 1.3.5)	Sub-IR 1.3.5	Establish activities to promote literacy (Activity 7)	Percent of community members who promote early childhood practices and support their children's education	Custom	60.1%	61.1%	54.3%

No.	Results framework statement	RF	Activity	Performance Indicator	Standard or CRS Custom	Baseline	Midterm	Endline
37	Increased Access to Food (Output 1.2.1.1, 1.3.1.1)	Output 1.2.1.1, 1.3.1.1	Take home rations (Activity 14)	Quantity of take-home rations provided (in metric tons) as a result of USDA assistance	Standard #14	0	230	213,769
38	Increased Access to Food (Output 1.2.1.1, 1.3.1.1)	Output 1.2.1.1, 1.3.1.1	Take home rations (Activity 14)	Number of individuals receiving take-home rations as a result of USDA assistance	Standard #15	0	12,214	47,792
39	Increased Access to Food (Output 1.2.1.1, 1.3.1.1)	Output 1.2.1.1, 1.3.1.1	Provide school meals (Activity 11)	Number of daily school meals (breakfast, snack, lunch) provided to school-age children as a result of USDA assistance	Standard #16	0	7,754,804	5,947,171
40	Increased Access to Food (Output 1.2.1.1, 1.3.1.1)	Output 1.2.1.1, 1.3.1.1	Provide school meals (Activity 11)	Number of social assistance beneficiaries participating in productive safety net as a result of USDA assistance	Standard #18	0	63,019	53,964
41	Increased Access to Food (Output 1.2.1.1, 1.3.1.1)	Output 1.2.1.1, 1.3.1.1	Form savings and lending groups (Activity 9)	Number of individuals participating in group-based savings, micro-finance, or lending programs with USDA assistance	FFPr Standard #6	0	2,664	10,357
42	Increased Capacity of Government Institutions (FR 1.4.1)	FR 1.4.1	Capacity Building: Local, regional, national level (Activity 5)	Number of members of the interministerial steering committee conducting monitoring visits to targeted schools	Custom	0	5	9
43	Improved Policy and Regulatory Framework (FR 1.4.2) Improved Policy and Regulatory Framework (FR 2.7.2)	FR 1.4.2/ 2.7.2	Capacity Building: Local, regional, national level (Activity 5)	Number of policies, regulations, or administrative procedures in each of the following stages of development as a result of USDA assistance	Standard #10	0	3	2
44	Increased Government Support (FR 1.4.3) Increased Government Support (FR 2.7.3)	FR 1.4.3/ 2.7.3	Capacity Building: Local, regional, national level (Activity 5)	Value of new USG commitments, and new public and private sector investments leveraged by USDA to support food security and nutrition	Standard #11	0	$748,801	$481,920
45	Increased Government Support (FR 1.4.3) Increased Government Support (FR 2.7.3)	FR 1.4.3/ 2.7.3	Capacity Building: Local, regional,	Value of new USG commitments, and new public and private sector investments leveraged by USDA to support food security and nutrition **[Host Government amount]**	Standard #11	0	$251,492	$160,796

No.	Results framework statement	RF	Activity	Performance Indicator	Standard or CRS Custom	Baseline	Midterm	Endline
			national level (Activity 5)					
46	Increased Government Support (FR 1.4.3) Increased Government Support (FR 2.7.3)	FR 1.4.3/ 2.7.3	Capacity Building: Local, regional, national level (Activity 5)	Value of new USG commitments, and new public and private sector investments leveraged by USDA to support food security and nutrition **[Private]**	Standard #11	0	$135,937	$1,478
47	Increased Government Support (FR 1.4.3) Increased Government Support (FR 2.7.3)	FR 1.4.3/ 2.7.3	Capacity Building: Local, regional, national level (Activity 5)	Value of new USG commitments, and new public and private sector investments leveraged by USDA to support food security and nutrition **[Another public sector]**	Standard #11	0	$361,372	$319,646
48	Increased Engagement of Local Organizations and Community Groups (FR 1.4.4)	FR 1.4.4	Training: Parent-Teacher Associations (Activity 16)	Number of Parent Teacher Associations (APE) or similar school governance structure supported as a result of USDA assistance	Standard #13	0	138	138
49	Increased Engagement of Local Organizations and Community Groups (FR 1.4.4) Increased Access to Food (Output 1.2.1.1, 1.3.1.1)	FR 1.4.4/ Output 1.2.1.1, 1.3.1.1	Form savings and lending groups (Activity 9)	Number of public private partnerships formed as a result of USDA assistance	Standard #12	0	107	127
50	School-Age Children in the Savanes and Kara Regions Have Improved Literacy (SO 1)	SO1	Raising awareness on importance of education (Activity 12)	Percent of students who, by the end of two grades of primary schooling, demonstrate that they can correctly identify letter sounds	Custom	5.3%	4.3%	9.2%

Annex C: Results Framework for STARS Project

Figure 15: Strategic objective 1 (SO1)

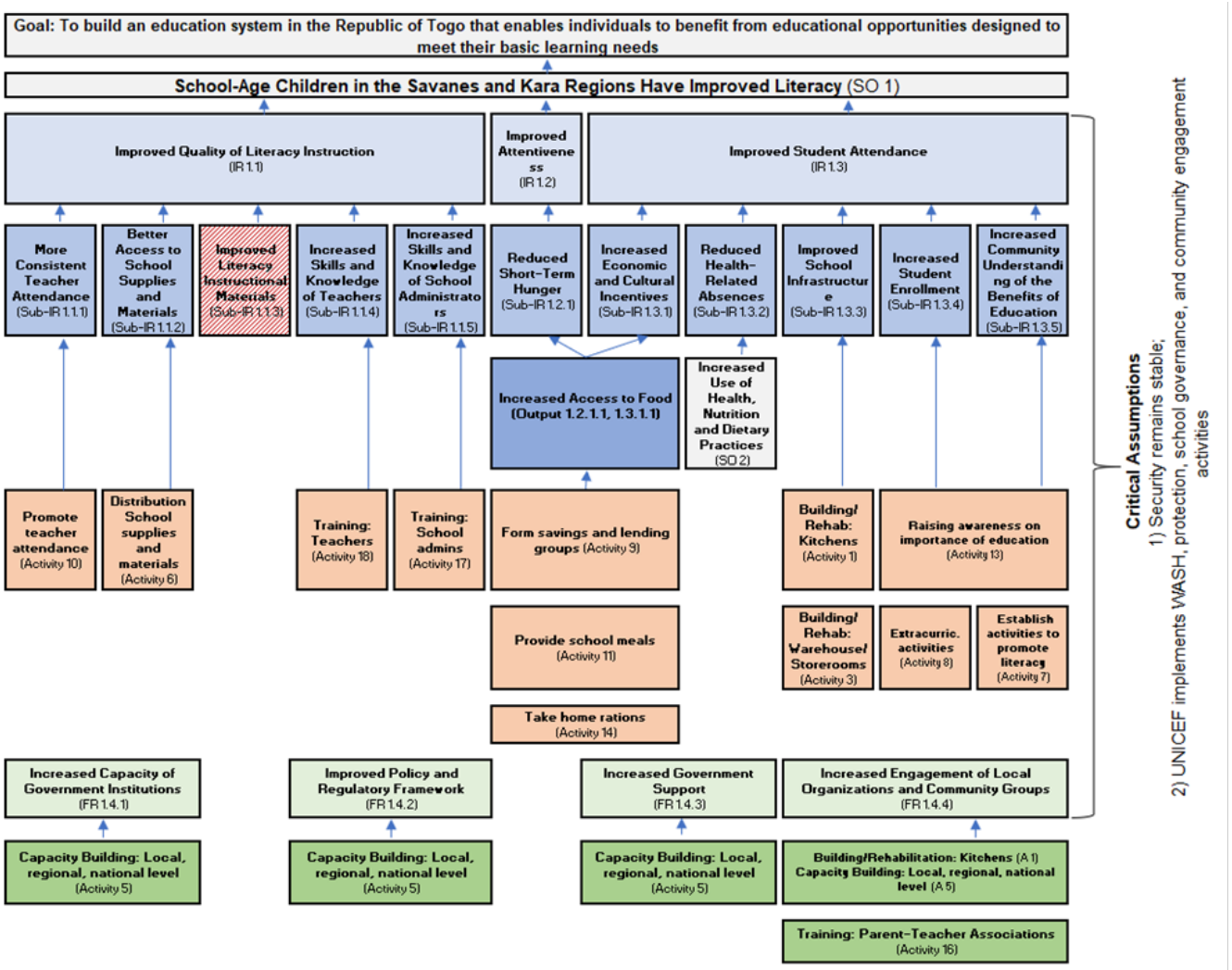
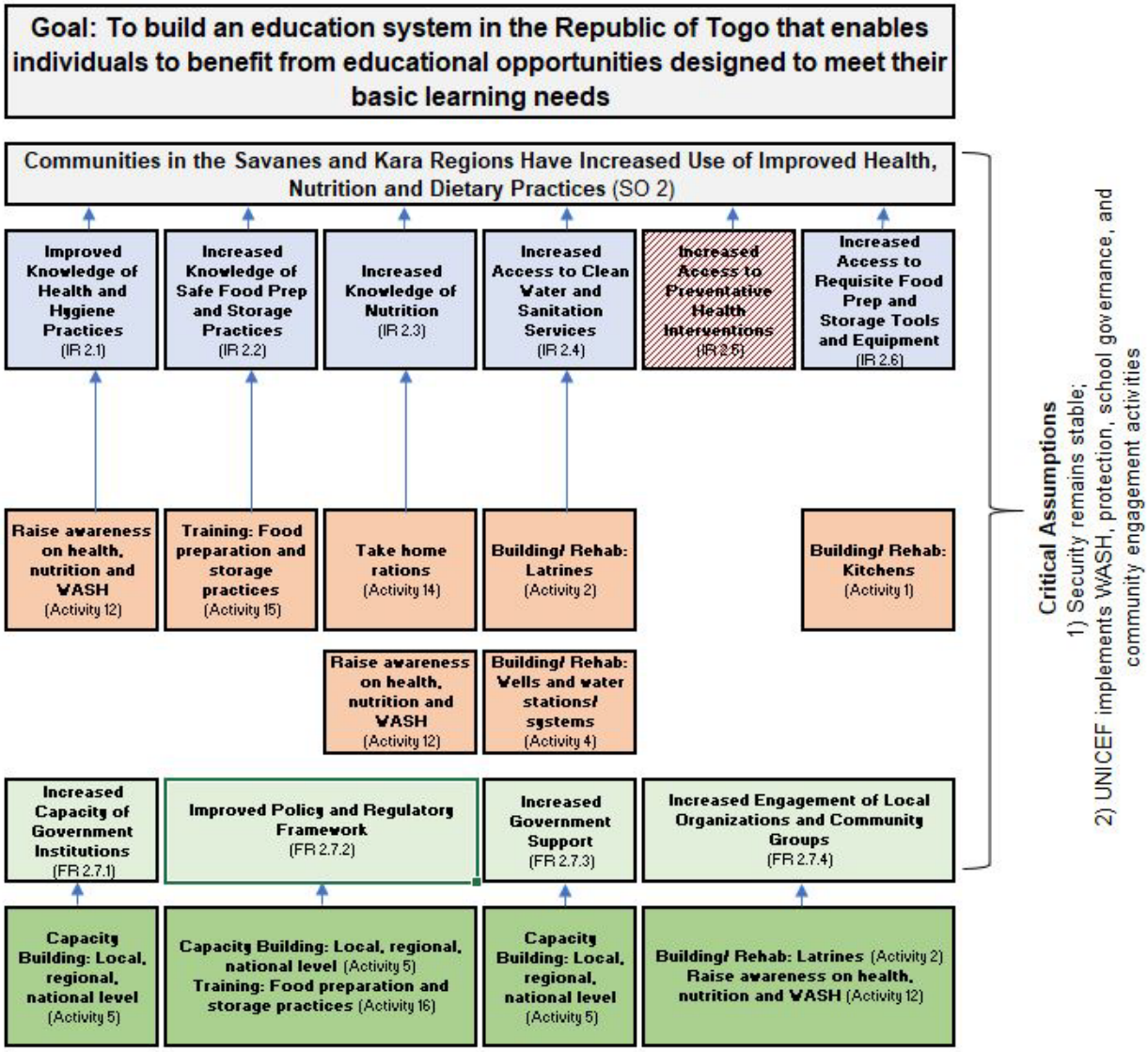


Figure 16: Strategic objective 2 (SO2)



Annex D: Terms of Reference for the Evaluation

TERMS OF REFERENCE (TOR)

Baseline, Midterm and Final Evaluation

Republic of Togo McGovern Dole FY20-FY24

1. Purpose and Overview:

The purpose of these Terms of Reference (TOR) is to outline the conditions and responsibilities of the external evaluator who will undertake the baseline, midterm evaluation and final evaluation of the *Santé, Transformation et Apprentissage pour une Réussite Scolaire* (STARS)[43] project, a USDA-funded McGovern-Dole International Food for Education project in the Republic of Togo.

Please note these ToR and its annexes are subject to donor approval and thus may change before contract signing.

Note these ToR rely heavily on Annex 1. Evaluation Plan for the STARS project; specific relevant sections are outlined below. The external evaluator should be very familiar with Annex 1, and Annex 2. Indicator Performance Tracking Table (IPTT), in addition to the USDA's Food Assistance Indicators and Definitions and its Monitoring and Evaluation Policy. Finally, the external evaluator should also be very familiar with Annex 5, the project's Performance Monitoring Plan (PMP).

The midterm evaluation will be conducted by the same firm who carried out the baseline evaluation, School to School (STS). STS will still be allowed to carry out the midterm evaluation due to the high quality of work they did during the project baseline survey.

2. Project Background:

Section 2 of Annex 1 provides an overview of the STARS project.

3. Evaluation Purpose, Scope, Approach, and Methodology:

Please note that Section 3 of Annex 1 provides an overview of evaluation activities including stakeholders, anticipated data collection tools, the STARS Results Framework, and sample size requirements. Section 8 of Annex 1 describes special studies for which the external evaluator will be responsible.

Information in this section, and in Annex 1, outline the standards expected of the external evaluator during data collection and analysis. Justified deviations from these standards, after consultation with CRS, are possible.

COVID-19 Precautions: CRS will require the external evaluator to propose and implement a satisfactory plan to mitigate the spread of COVID-19 during the data collection phase of the baseline evaluation (and further evaluations, if need be). This COVID-19 plan needs to include contingencies for study design, trainings, data collection, analysis and reporting, and budget implications.

Example of contingency measures in Togo to avoid spread of COVID-19 are:

- working in a well-ventilated room
- sensitization of participants on anti-COVID 19 measures before the beginning of all training

[43] In English: "Health, Transformation and Learning for School Success"

- physical distancing of at least 1 meter between participants during working sessions and training on the field
- systematic wearing of masks during working sessions and training on the field
- hand washing using water and hydroalcoholic gel by all participants during working sessions and training in the field
- No gathering of persons more 50 persons

Anti-COVID-19 training modules are available at CRS to help STS establish the mitigation measures, in case of need.

Data Collection Tools. The selected external evaluator, STS, will work with CRS to update the baseline evaluation tools, keeping in mind the project's PMP. These tools will be completed by additional ones developed by STS to address the evaluation questions related to DAC criteria of relevance, effectiveness, efficiency, impact, and sustainability.

Use random samples and document any sample bias due to non-random sampling. Representative samples should always be selected randomly, ideally from a list or using a random walk, etc. However, often due to resource constraints, sample selection bias does occur. This frequently happens due to security constraints that prevent study teams from reaching an off-limits area or when the rosters from which individuals or clusters are randomly selected are outdated, and it would prove too costly or impossible to locate those randomly selected. In this case, in the limitations section of the evaluation report, describe any sources of bias as best as possible.

For example, if students are not present in school on the day of evaluation, how do absent students differ from those present? Does a t-test of means show that the proportion of key groups (gender, ethnicity, geographic area)[44] in the sample is the same as those that were not included? If not, how might the sample be biased? How else might students not present that day be different? Might they not perform as well on literacy tests, etc., because they might frequently miss school?

Check for statistical differences in outcome-level indicators over time. The mid-term and final evaluations should, at minimum, check for statistical differences between baseline and respective report values. This will be via a t-test; however, a preferred general specification would be:

$$Outcome_{its} = Intercept + Midterm_t + Final_t + Female_i + Strata_s + \varepsilon_{its}$$

where

- $Outcome_{its}$ is the outcome indicator of interest for individual *i* at time *t* (baseline, midterm, or final) in strata *s*
- $Midterm_t$is a binary variable taking the value 1 if the data was collected during the midterm evaluation, and zero otherwise
- $Final_t$is a binary variable taking the value 1 if the data was collected during the final evaluation, and zero otherwise (only relevant at final evaluation)
- $Female_i$ is a binary variable taking the value 1 if individual *i* is female, and zero otherwise
- $Strata_s$ is a vector of binary variables for each stratum (excluding one to avoid the dummy variable trap)
- $\varepsilon_{its}$is the error-term that should be clustered at the cluster-level during analysis

Ideally, a table with each indicator of interest could be presented per row, with the coefficient (or marginal value when using probit/ logit models) and standard errors for the midterm, final, and

[44] The analyst may not have much information about students not present. However, based on student names and school locations, they might at least have this information.

female indicators in columns. It is not necessary to present marginal values per stratum. The specification can be adapted if the outcome indicator is not at the individual level, not stratified, or not clustered.

Sample weights. Sample weights should always be used when providing unconditional descriptive statistics (means or totals) for the underlying population. However, results from regression analyses, would ideally report unweighted and weighted results, and where there are differences, include a discussion of the underlying reasons. For example, observations from a school that has 90 second-graders vs. 30 will carry 3 times the weight; if there are heterogenous project effects for large vs. small schools (e.g. larger schools have a higher teacher/ student ratio; perhaps this lack of student attention results in poorer educational outcomes, etc.) then the conditional means might be different for weighted vs. unweighted analyses (Solon, Haider, and Wooldridge 2015).

Clustered or stratified samples and regression analysis. When reporting weighted conditional means from regression analyses, weighted values should use the appropriate weighted counterpart (e.g., weighted least squares, weighted maximum likelihood, etc.).

Additionally, because observations within a cluster are likely correlated, standard errors should always be clustered at the cluster-level (Cameron and Miller 2015). Statistical packages have functions for this; the appropriate function will vary depending on the method of analysis.
Control for any sample stratification in regression analyses by using binary variables for each stratum (excluding one to avoid the dummy variable trap).

Population Proportional to Size (PPS) cluster selection may not be appropriate. PPS is a quantitative sample selection methodology commonly used to account for the size of clusters when selecting them in the first stage of evaluation studies, in which every person in every cluster has an equal probability of being selected into the sample. If, in the second stage, a simple random sample is used to select each individual among all individuals in the cluster, then the sample is "self-weighting" and no sample weights need be applied at the analysis stage.
Analysts of data collected via a PPS-selected sample should understand that if the sample was stratified, or if a simple random sample was not used in the second stage, then the sample is not self-weighting and sample weights must be used.

At the analysis stage, the Hansen-Hurwitz or Horvitz-Thompson estimators should be used to estimate the sample mean, and variance in any regression models (Hansen and Hurwitz 1942, Horvitz and Thompson 1952).

When using PPS, the measure of size should be accurate, otherwise it will over- or underestimate the sample variance, as compared to simple random selection of clusters (Thomsen, Tesfu, and Binder 1986), despite using the estimators described above. Even if baseline measures of size are accurate, if using a repeated cross-section (schools are commonly maintained across all three evaluation points) when evaluating in the same clusters at midterm or final evaluation and the "size" of the clusters changes notably over time (likely to occur, enrollment is expected to increase as a result of project activities), the same issue of mis-estimating the sample variance will occur.

For all these reasons, using PPS is likely too complex and not appropriate for these evaluations, and therefore not recommended. In lieu of PPS, clusters and individuals can be selected via a random sample, and sample weights used in analysis.

Project indicators. The project Indicators Table below (table1) is the updated version, considering the values of the indicators obtained at the baseline evaluation. Only the nineteen indicators marked with

a or c in Table 1 will be collected during the midterm evaluation. These indicators include the ones who had been measured during the baseline (indicators related to project activities with zero values before the baseline because the activities did not start) and others with non-zero values before baseline and for which the values will be updated after the midterm evaluation, due to the implementation of the project activities. All individual-level data must be disaggregated by gender.

Table 29: STARS project indicators

Performance Indicator	USDA Standard/ CRS Custom	Baseline
1. Percent of students who, by the end of two grades of primary schooling demonstrate that they can read and understand the meaning of grade level text [a]	Standard #1	0%
2. Number of individuals benefiting indirectly from USDA-funded interventions [b]	Standard #31	0
3. Number of individuals who demonstrate the use of new child health and nutrition practices as a result of USDA assistance [c]	Standard #19	0
4. Number of individuals who demonstrate the use of new safe food preparation and storage practices as a result of USDA assistance [c]	Standard #20	0
5. Number of individuals participating in USDA food security programs [b]	Standard #30	0
6. Number of schools reached as a result of USDA assistance [b]	Standard #32	0
7. Percent of teachers providing quality classroom instruction with USG support [c]	USAID Ed Supp-10	0%
8. Percent of students in target schools identified as attentive during class/instruction [c]	Custom	60%
9. Average student attendance rate in USDA supported classrooms/schools [c]	Standard #2	80.2%
10. Number of individuals trained in safe food preparation and storage as a result of USDA assistance [b]	Standard #22	0
11. Number of individuals trained in child health and nutrition as a result of USDA assistance [b]	Standard #23	0
12. Number of children under five (0-59 months) reached with nutrition-specific interventions through USDA-supported programs	Standard #24	0
13. Number of pregnant women reached with nutrition-specific interventions through USDA-supported programs [b]	Standard #26	0
14. Number of children under two (0-23 months) reached with community-level nutrition interventions through USDA-supported programs [b]	Standard #25	0
15. Number of schools with improved sanitation facilities [c]	Standard #28	57
16. Number of schools using an improved water source [c]	Standard #27	70
17. Percent of health and nutrition infrastructure, constructed as a result of USDA assistance, maintained by communities/local authorities [c]	Custom	0%
18. Number of Schools receiving energy saving stoves [b]	Custom	0
19. Percent of instructional time lost due to teacher absenteeism [c]	USAID Ed Supp-11	9.3%
20. Number of schools implementing the use of school score cards [c]	Custom	0
21. Number of teaching and learning materials provided as a result of USDA assistance [b]	Standard #3	0
22. Number of teachers/educators/teaching assistants in target schools who demonstrate the use of new and quality teaching techniques or tools as a result of USDA assistance [c]	Standard #4	0
24. Number of teachers/educators/teaching assistants trained or certified as a result of USDA assistance [b]	Standard #5	0
25. Number of school administrators and officials in target schools who demonstrate use of new techniques or tools as a result of USDA assistance [c]	Standard #6	0
26. Number of school administrators and officials trained or certified as a result of USDA assistance [b]	Standard #7	0
27. Percent of school officials in target schools who demonstrate the use of new and quality supervision and leadership techniques or tools [c]	Custom	0%
28. Percent of children 6–23 months receiving a minimum acceptable diet [c, d]	FFP #BL12	17%
29. Number of school-age children receiving daily school meals (breakfast, snack, lunch) as a result of USDA assistance [b]	Standard #17	0

Performance Indicator	USDA Standard/ CRS Custom	Baseline
30. Percent of parents who state their children had health-related school absences in the previous month [c]	Custom	15%
31. Number of educational facilities (i.e., school buildings, classrooms, improved water sources, and latrines) rehabilitated/constructed as a result of USDA assistance [b]	Standard #8	0
32. Number of students enrolled in school receiving USDA assistance [c]	Standard #9	0
33. Number of schools that held an enrollment campaign [b]	Custom	0
34. Percent of caregivers who report on spending time on literacy activities with their school-age children in the previous week [c]	Custom	15.8%
35. Number School Management Committee (SMC) and Parent Teacher Association (APE) members, and Mother Leaders trained on activities to promote literacy [b]	Custom	0
36. Percent of community members who practice promoted early childhood practices and support their children's education [c]	Custom	60%
37. Quantity of take-home rations provided (in metric tons) as a result of USDA assistance [b]	Standard #14	0
38. Number of individuals receiving take-home rations as a result of USDA assistance [b]	Standard #15	0
39. Number of daily school meals (breakfast, snack, lunch) provided to school-age children as a result of USDA assistance [b]	Standard #16	0
40. Number of social assistance beneficiaries participating in productive safety net as a result of USDA assistance [b]	Standard #18	0
41. Number of individuals participating in group-based savings, micro-finance or lending programs with USDA assistance [b, e]	FFPr Standard #6	0
42. Number of members of the interministerial steering committee conducting monitoring visits to targeted schools [b]	Custom	0
44. Number of policies, regulations, or administrative procedures in each of the following stages of development as a result of USDA assistance [b]	Standard #10	0
45. Value of new USG commitments, and new public and private sector investments leveraged by USDA to support food security and nutrition [b]	Standard #11	0
46. Number of Parent Teacher Associations (APE) or similar school governance structure supported as a result of USDA assistance [b]	Standard #13	0
47. Number of public private partnerships formed as a result of USDA assistance [b]	Standard #12	0
48. Percent of students who, by the end of two grades of primary schooling, demonstrate that they can correctly identify letter sounds [Custom]	Custom	5.3

[a] Collected by only external evaluator

[b] Collected only by CRS; triangulated by external evaluator

[c] Collected by external evaluator; triangulated with CRS annual report data

[d] USAID Food for Peace standard indicator

[e] USDA Food for Progress standard indicator

4. Deliverables:

The evaluator is expected to follow American Evaluation Association's Guiding Principles for Evaluators (http://www.eval.org/p/cm/ld/fid=51). Dependent upon participants in the evaluation, the evaluator should specify steps that will be taken to ensure informed consent, confidentiality, and protection of minors. The evaluator should specify steps taken to safeguard data collected and data management procedures to be used in the evaluation. There will be a data rights clause in the signed contract, and the external evaluator should obtain permission from CRS before sharing the final evaluation report with any external party, including posting it to their organization's website.

All deliverables should be completed in English (and data collection tools must also be in French), be free of typos or grammatical errors, and be a polished document ready for submission to USDA. This means the document contains no factual errors or inaccuracies and citations are properly used.

Deliverables for baseline, midterm, and final include the following:

- Work plan (including evaluator responsibilities for identifying, interviewing, contracting, training, and overseeing a balanced team of male and female enumerators and enumerator supervisors).
- Sampling plan, including if the sample sizes will differ from Annex 1, approved by CRS.
- Instruments, data collection manual, and training materials for enumerators (i.e., focus group guides, key informant interview guide, observation checklist), approved by CRS.
- Quality Assurance Plan (including training of enumerators and weekly check-ins during data collection, approved by CRS.
- Conduct interview with USDA (it is expected USDA will facilitate this exercise by providing the contact person and the means of interview)
- Data sets with accompanying codebook/data dictionary (original paper and/or electronic as well as final, clean electronic data sets with syntax).
    - If the evaluator provides .dta, .do, .sps, or .sav files, they must also provide open-source file versions (.txt, .csv, .doc, etc.)
    - If part of a longitudinal design, an identifier file that links respondent PII with ID numbers in the data file(s)
    - Deidentified transcripts of selected interviews and focus groups and/or data files of coded sections of text from interviews and focus groups
- Draft Report with one round of edits from CRS and another subsequent round from USDA
- Final Report with the following sections:
    - Executive summary 2 to 3 pages (including brief introduction of program evaluated, key evaluation questions, findings, and conclusions)
    - Background
    - Evaluation questions
    - Evaluation design including assumptions and limitations
    - Methodology
    - Findings
    - Conclusions, lessons learned and effective practices (if any)
    - Recommendations (should be clear, concise, relevant, specific, and practical, following directly from findings and conclusions established in the report)
    - Annex with original scope of work (marked for redaction from final web version)
    - Annex with final data collection instruments
    - Annex with description of team members' qualifications and their positionality
    - Annex with additional methodological discussion/ robustness checks as needed
    - Annex with updated IPTT
- Final reports must not contain any propriety or personally identifiable information (PII). PII is any information that directly or indirectly identifies an individual. This information can be used on its own or with other information to identify, contact or locate a single person, or to identify an individual in a specific situation. This may include, for example, a name, national ID number, address, birthplace, etc. PII includes both direct and indirect identifiers that, when taken together, could allow for identification of an individual (such as a village name, gender, age, name, and/ or facial image)."
    - In addition, final reports should not allow for the identification of individual schools or communities. Any list of schools or communities provided should be included as in the report annex, so that it can be easily removed before submitting it to USDA for external sharing.

- Final reports must be compliant with Section 508 of the United States Access Board which requires that information and services be accessible to persons with disability. (See https//section 508.gov/create).
- A two to four-page outward-facing summary document, with easily accessible graphics, highlighting the project's key successes, for sharing with a larger audience
- Presentation of final evaluation to stakeholders. This can occur before or after reporting submission to USDA, as long as any key feedback is incorporated into the final version of the report (that USDA posts to the Development Experience Clearinghouse). This can be done via an additional annex, if the report is in its final stages before this presentation is conducted.
- A webinar of key findings and lessons learned for CRS globally and USDA (if requested).

In addition, at baseline only, a 10-page preliminary report, suitable for presentation to USDA, 6 weeks after the end of data collection. The report will only contain:

- An IPTT for the indicators with non-zero baseline values, including relevant disaggregates.
- Enough information about the methodology to engender confidence in the data quality. This should include a list of the data collection tools, number and gender of people interviewed, any information about stratification, and any data limitations. Whenever possible, the preliminary report should simply refer to the approved ToR and/ or Evaluation Plan, rather than incorporate the information.
- Annex with description of team members' qualifications and their positionality.

5. Items provided to the external evaluator by CRS:

- Use of CRS CommCare software license, if desired. Evaluators are free to use their preferred data collection platform.
- Tablets for data collection.
- Scales and stadiometers for anthropometric data collection as described in Special Study 3.
- All Annexes to this ToR.

6. Main Evaluation Questions and Timetables:

Sections 4 – 6 of Annex 1 outlines the timelines of the baseline, midterm, and final evaluations and present anticipated evaluation questions.

7. Evaluator Qualifications:

Team must have the following qualifications

a) Advanced Degree in social sciences with strong knowledge of statistics/ demography.
b) Knowledge and experience in survey and sampling design.
c) Experience managing complex and multi-sectoral evaluations.
d) Knowledge of performance evaluations, especially in the education sector.
e) Knowledge of the education sector; basic education in the development context; school feeding programs especially in West Africa, preferably Togo.
f) Demonstrated experience in conducting evaluation surveys of similar nature, preferably for USDA-funded projects.
g) Good verbal and written communication skills in English and French.
h) Willingness to work in remote areas without electricity and running water.

8. Evaluation team, management, and coordination:

Section 9 of Annex 1 broadly describes evaluation management. In addition, please see Table 2 below

Table 30: Evaluation team members

Team Member	CRS Staff or hired independently by the evaluation firm	Main Roles and Responsibilities
External evaluator	Hired independently	Preside over the conduct of the entire evaluation, from methodology and tool development to training in the use of the tool to field testing, data collection, entry and analysis and report writing.
Enumerators/data collectors	Hired independently by the evaluation firm	Receive training and undertake data collection in the field.
Data Collection Supervisors	Hired independently by the evaluation firm	Receive training in data collection and supervise data collectors daily for the duration of the data collection exercise.
Data entry clerks	Hired independently by the evaluation firm	Receive training in data entry and enter data collected from the field.
Data Entry Supervisors	Hired independently by the evaluation firm	Receive training in data entry and supervise data entry clerks throughout the data entry exercise.
CRS Togo Country Manager, CRS Benin/ Togo MEAL Coordinator	CRS Staff	Supports the entire evaluation process ensuring compliance on the part of the evaluation firm
CRS MEAL Advisors in Central Africa and Baltimore	CRS Staff	Supports the entire evaluation process ensuring compliance on the part of the evaluation firm.

9. Structure of Proposal and Submission Guidelines

CRS published a request for bids (financial and technical proposals) for the conduct of the baseline, midterm, and final evaluation of the STARS project to both domestically and internationally. Applicants were supposed to meet the qualifications stipulated in this ToR. The bid evaluation process was managed by the Togo CRS Procurement Officer and the Central Africa Regional Technical Advisor (RTA) for MEAL and followed the standard rules and procedures for the competitive and transparent procurement of consultancy services. The successful evaluator, STS, was contracted to execute the baseline, midterm, and final evaluation. However, retention of the evaluator to proceed with the midterm and/or final evaluation was dependent on satisfactory performance of the baseline evaluation. CRS was to re-launch the selection process for the midterm and final evaluation where the baseline consultant(s) does not meet expectations.

Key criteria that will be considered during the bid evaluation process will include the following:

1. Bidders must submit a technical proposal including a detailed description of the study design and methodology for the baseline.
2. Bidders must submit a detailed financial proposal for the baseline, midterm, and final evaluation, and special studies, not exceeding $450,000 for the three data collection points.
    a. Please list a separate line item for Special Study 3 in Annex 1.
3. Bidders should submit a detailed work plan showing clearly how they wish to accomplish the study.
4. Profile of the bidders including relevant knowledge and experience to undertake the assignment

5. Bidders should have stated their relevant qualification and demonstrate relevant experience in the project area and experience in evaluating education programs.
6. Delivery timeline

The proposal should contain no more than a total of 25 pages of which; technical proposal 20 pages and financial proposal 5 pages. See table 31 below.

Following the above criteria, STS won the contract to conduct baseline, midterm, and final evaluations. Their report of the baseline was accepted and approved by both CRS and USDA. Consequently, STS will conduct the midterm evaluation.

Table 31: Proposal layout and number of pages

Proposal content layout	Maximum pages
Technical Proposal	**20**
Expression of interest	1
Table of content	1
Introduction and background	1 ½
Qualification and profile of team members	2 ½
Evaluation methodology	5
Evaluation questions	2 ½
Work plan and deliverables	2 ½
Technical reference for the firm	4
Financial Proposal	**5**
Summary	1
Detailed budget	3
Budget explanatory notes	1
Total	**25**

Sealed bids must be delivered in electronic and/or hard copy to:
The CRS-Togo Office
01 BP 173 Hedzanawoe-Derriere Sito Aeroport
Lomé, Togo
Email: togo@global.crs.org

The proposals must be submitted **no later 23 October 2019 at midnight GMT**.

Bids for multiple awards. CRS currently also has an open bid for its newly awarded McGovern-Dole project in Guinea-Bissau and understands that some bidders may be interested in bidding for both contracts. The process is run separately in each country's program. Applying for both contracts is acceptable, but country programs do consult each other in these processes. Thus, please note the following:

1) Given that timelines overlap, evaluators should clearly demonstrate they have the bandwidth to produce quality evaluations for both countries, either through expected LOE for overlapping staff members; different staff over specified dates; or the use of different study teams altogether.
2) Evaluators that are currently slated to conduct midterm or final evaluations for other CRS country programs during overlapping timeframes should also include clarity around point 1) above.

Table 32: List of annexes (attached as separate documents)

Annex Number	Document
1	STARS Evaluation Plan (Budget Information Redacted)
2	STARS Indicator Performance Tracking Table
3	CRS Report Review Template for USDA Evaluations
4	CRS Standard Tools
5	STARS Performance Monitoring Plan (PMP)

Annex E: Description of Team Members' Qualifications

DR. PARNIKA BHATI

Dr. Bhatia is a cognitive scientist and education researcher with eight years of direct experience in the education sector. Bhatia serves as a technical lead and advisor who designs and facilitates rigorous monitoring, evaluation, and learning processes for partners and organizations like FCDO and USAID. Her areas of expertise include numeracy development, socio-emotional learning, child development, literacy instruction, inclusive education, and community development.

Bhatia has led study design, sampling, monitoring data collection, analysis, and reporting for early grade reading assessments (EGRA) and early grade mathematics assessments (EGMA). She currently serves as a Technical Lead for Strengthening the Teaching of Primary School Mathematics in Malawi and the FCDO's Manahel project. In addition to EGRAs, Bhatia has conducted endline analysis for two projects in the All Children Reading portfolio, examining the impact of EdTech on young students' literacy outcomes in Nepal and Malawi. She also has experience in the adaptation and pilot of the International Social and Emotional Learning Assessment (ISELA) in the Democratic Republic of Congo.

Bhatia's experience extends beyond data collection and analysis and includes capacity strengthening. She has led professional development workshops for teachers in under-served communities in India and developed training modules in data analysis software and best practices in teaching for AmeriCorps Teaching Fellows in the USA. Bhatia has also designed and led a multi-site RCT in public schools in France.

Bhatia holds a Ph.D. in Developmental Cognitive Science from the University of Lyon, France, an Ed.M. in Mind, Brain, and Education from the Harvard University, USA, has taught fourth and fifth graders in India as a Teach for India fellow, and speaks native Hindi, and beginner-level French and Punjabi.

JARRET GUAJARDO

Jarret Guajardo is a monitoring, evaluation, research, & learning (MERL) practitioner with over a decade of experience in teaching, evaluating, and building MERL capacity in international education. At STS, Guajardo is a Principal Researcher, leading quantitative evaluation design and providing high-level technical input, mentorship, and quality assurance on a portfolio of projects. Prior to joining STS, Guajardo led over a dozen impact evaluations and advised dozens more studies across 20 countries on programs in early grade literacy & numeracy, early childhood development, child poverty, health & nutrition, child protection, and anti-trafficking. In his various roles at Save the Children, World Vision, and Innovations for Poverty Action, Guajardo oversaw change management and capacity-strengthening of international NGO country offices, local implementing partners, and donors to improve rigor and realign traditional MERL work toward more strategic learning. Guajardo has lived in India, Cambodia, Lebanon, and Jordan. Guajardo holds a Master of Arts in international relations from the Johns Hopkins School of Advanced International Studies and a Bachelor of Arts in international relations with a minor in Arabic from Stanford University. He speaks English, Spanish, and Arabic.

DREW SCHMENNER

Drew Schmenner has supported STS's portfolio of projects for the last seven years with a variety of skills in editing and writing, monitoring and evaluation, and learning. He specializes in writing and

editing reports and case studies for both qualitative and quantitative studies and serves as a qualitative research advisor on select projects. In previous roles at STS, he coordinated its electronic data capture and data management efforts across its entire project portfolio and supported numerous data collection trainings in the field.

Mr. Schmenner applies the writing and editing expertise he developed for five years as an award-winning newspaper reporter in southern California to STS's technical reports and case studies. He has contributed technical expertise in qualitative analysis and writing for numerous reports, including two performance evaluations of USAID's Reading for Success (RFS) projects in Morocco, two midline studies of Link Education International's Supporting the Transition of Adolescent Girls Through Enhancing Systems (STAGES) project in Ethiopia, and a mid-term evaluation of USAID's Girls' Leadership and Empowerment Through Education (GLEE) project in Mali implemented by Winrock International. Mr. Schmenner is also well-versed in quantitative studies, having co-wrote the EGRA endline report for the USAID Sindh Reading Program.

In the past at STS, Mr. Schmenner has overseen data quality assurance for early grade reading assessments (EGRA) in five countries by monitoring data collected daily and communicating and resolving issues with staff in the field. He has provided expert troubleshooting and internal and external training support to early grade reading data collection software for STS. He has direct experience training data collectors in the field in five countries in key protocols for EGRA data collection and in the functionality of Tangerine, Ona, and SurveyCTO electronic data collection software. In addition, he has cleaned EGRA data in Stata for multiple projects and helped to improve efficiencies in STS's data quality assurance processes in data monitoring and cleaning.

Mr. Schmenner holds an M.A. in International Studies from the University of San Francisco and served as a Peace Corps Volunteer in Niger. He is a native English speaker with basic proficiency in French.