



USAID
FROM THE AMERICAN PEOPLE

COVER PHOTO: GRADE 2 STUDENTS, ARTIBONITE, HAITI
ANDREW CARMONA, 2017



FINAL EVALUATION REPORT

Haiti Early Reading Program (ERP) Baseline Evaluation

This publication was produced for review by the United States Agency for International Development by Social Impact, Inc.

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



Haiti Early Reading Program (ERP)

Baseline Evaluation

USAID/Haiti
AID-521-C-17-00002
March 16, 2018

This publication was produced at the request of the United States Agency for International Development (USAID). It was prepared independently by Andrew Carmona, Basab Dasgupta, Amanda Stek, Joanie Cohen-Mitchell, and Euphonise Loiseau on behalf of Social Impact, Inc. Team member qualifications can be found in Annex G of this report.

CONTENTS

ACRONYMS	IV
ACKNOWLEDGEMENTS	V
ABSTRACT	VI
EXECUTIVE SUMMARY	VIII
INTRODUCTION	I
METHODOLOGY	5
FINDINGS	16
CONCLUSIONS	54
ANNEX A: FULL LISTING OF REFERENCES IN REPORT	58
ANNEX B: DATA COLLECTION TOOLS	60
ANNEX C: EVALUATION METHODOLOGY	116
ANNEX D: ADDITIONAL FINDINGS	123
ANNEX F: EVALUATION STATEMENT OF WORK	127
ANNEX G: QUALIFICATIONS OF TEAM MEMBERS	136
ANNEX H: EVALUATION DISSEMINATION PLAN TEMPLATE	139
ANNEX I: DATA QUALITY ASSURANCE MEASURES	141
ANNEX J: INTER-RATER RELIABILITY SCORES OF ENUMERATORS	142
ANNEX K: CONSIDERATIONS FOR FUTURE DATA COLLECTION	144
ANNEX L: DISCLOSURE OF ANY CONFLICTS OF INTEREST	145

TABLE OF TABLES

Table 1: Partners and their SOWs for the ERP Impact Evaluation.....	3
Table 2: List of ERP implementers and evaluation sample schools.....	6
Table 3: Planned qualitative sample size at baseline.....	9
Table 4: EGRA subtasks	14
Table 5: Study limitations.....	15
Table 6: School sample realized by school type, rural/urban, and department.....	17
Table 7: Final EGRA sample realized, by grade and gender	18
Table 8: Final sample realized for qualitative data collection	18
Table 9: Covariate balance statistics after CEM matching.....	19
Table 10: EGRA scores (for all), with p-values.....	21
Table 11: Proportion of zeros in subtasks across treatment groups.....	23
Table 12: Updated MDES, by EGRA sub-tasks and Grade.....	25
Table 13: Basic school-level information across schools	26

Table 14: School level resources and activities, by treatment and comparison	27
Table 15: EGRA subtask scores from other Haiti EGRA studies	31
Table 16: Distribution of ORF scores across grades and by CWPM categories	32
Table 17: Distribution of CLPM scores across grades	33
Table 18: Distribution of CLSPM scores across grades	33
Table 19: Regression models showing the effects of various factors on ORF.....	40

TABLE OF FIGURES

Figure 1: Sampled schools of the ERP Baseline Evaluation.....	vii
Figure 2: Sample size calculations and power analysis for the ERP Impact Evaluation.....	9
Figure 3: Decile distribution of timed learning test scores (percent of students in each category).....	22
Figure 4: Quartile distribution of non-zero scores in timed outcomes (for all students)	24
Figure 5: School Director experience	27
Figure 6: Teacher characteristics, per school	28
Figure 7: Descriptive statistics for classroom-level variables.....	29
Figure 8: Language spoken by students at home, by school type and location.....	29
Figure 9: Extent of punitive measures (as reported by students), by school type (percentage)	30
Figure 10: Percent of students who eat in the morning before school.....	30
Figure 11: Average scores (in words) in timed subtasks, by grade and gender.....	34
Figure 12: Proportion of zero scores in timed subtasks, by grade and gender	34
Figure 13: Timed learning outcomes non-zero scores, by gender	35
Figure 14: Average scores in non-timed subtasks, by grade and gender.....	36
Figure 15: Proportion of zeros in non-timed subtasks, by grade and gender	36
Figure 16: Non-timed learning outcomes, non-zero scores, by gender.....	37
Figure 17: Timed learning outcomes, by grade and implementer	38
Figure 18: Non-timed learning outcomes, by grade and implementer	38
Figure 19: Timed learning outcomes, by grade and type of school.....	39
Figure 20: Non-timed learning outcomes, by grade and type of school	39
Figure 21: Decomposition of drivers that explain ORF performance, overall and by gender	44
Figure 22: Classroom-level variables on classroom management and phonics approaches.....	49
Figure 23: Classroom-level variables on classroom materials and other topics	50

ACRONYMS

AIR	American Institutes for Research
ATL	Assistant Team Leader
CEM	Coarsened Exact Matching
CLPM	Correct letter identification per minute
CLSPM	Correct letter sound per minute
CWPM	Correct word per minute
DQA	Data Quality Assurance
EGRA	Early Grade Reading Assessment
EMIS	Education Management Information System
EPT	Éducation Pour Tous
EQ	Evaluation Question
EQUI®	Evaluation Quality, Utilization, and Impact®
ERP	Early Reading Program
ESS	Evaluation and Survey Services
ET	Evaluation Team
FGD	Focus Group Discussion
FONHEP	Haitian Foundation for Private Teaching
ICC	Intra-class Correlation
IDB	Inter-American Development Bank
IE	Impact Evaluation
IFOS	Institut de Formation du Sud
IP	Implementing Partner
IRR	Inter-rater reliability
JICA	Japan International Cooperation Agency
KII	Key Informant Interview
LAC/READS	Latin American and Caribbean READS Program
MDES	Minimum Detectable Effect Size
MENFP	The Ministry of Education and Professional Training
NGO	Non-governmental Organization
ODK	Open Data Kit
ORF	Oral Reading Fluency
PTA	Parent-Teacher Association
QA	Quality Assurance
QED	Quasi-experimental Design
RCT	Randomized Control Trial
SD	Standard Deviation
SI	Social Impact, Inc.
SOW	Scope of Work
SSME	Snapshot of School Management Effectiveness
ToTAL	Tout Timoun Ap Li
UND	University of Notre Dame
UNICEF	The United Nations Children's Fund
USAID	United States Agency for International Development

ACKNOWLEDGEMENTS

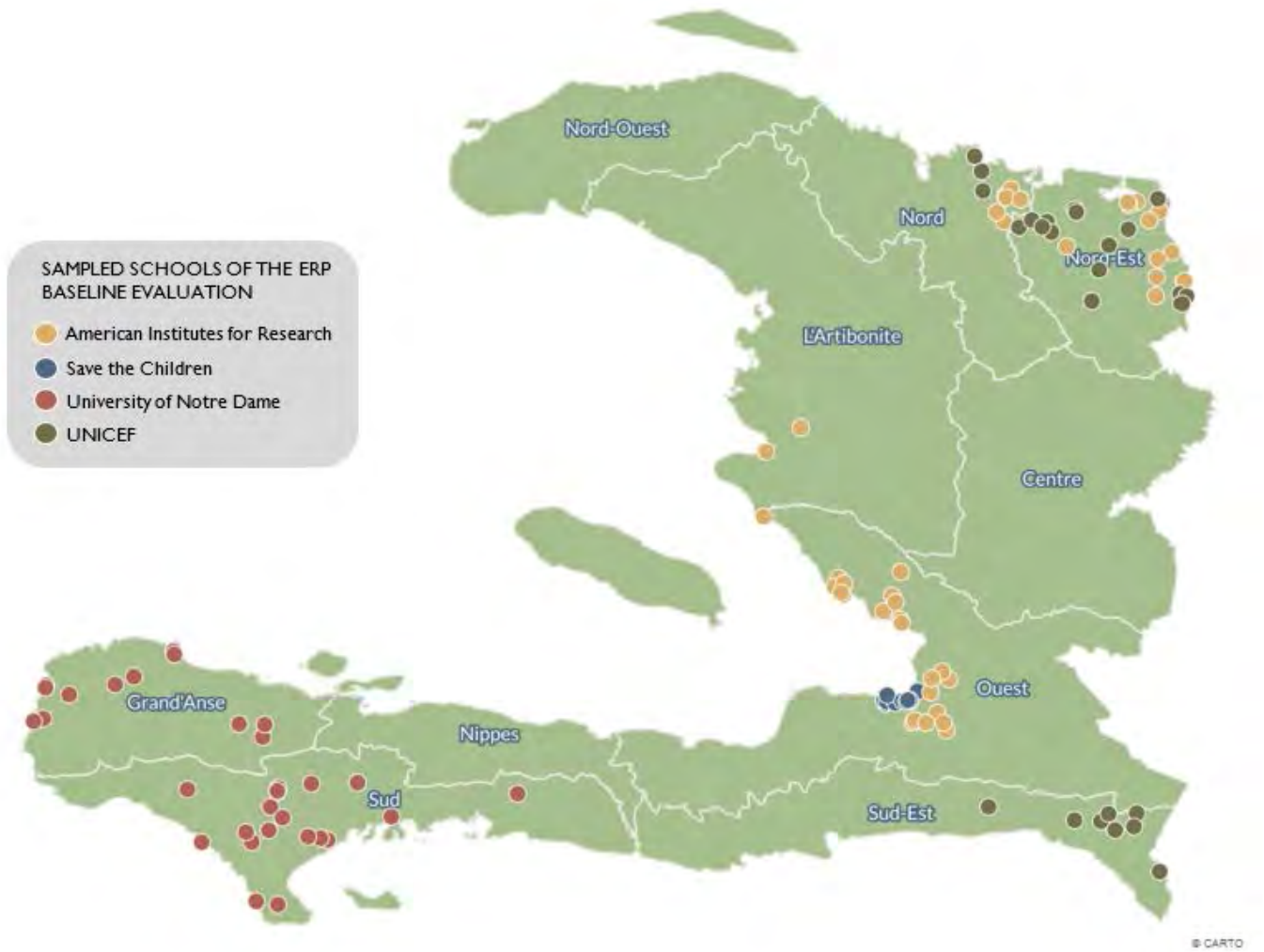
The authors would like to wholeheartedly thank several key individuals who were critical to the successful completion of this baseline report. We thank the Education Team at USAID/Haiti, namely Christine Veverka and Ezechiel Macon, for being flexible, available, and responsive to our various questions and inquiries. We extend very deep thanks to Samuel Emieux Jean, Ulrick Jean Claude, and Charles Tardieu of IFOS, as well as the 122 supervisors and enumerators who carried out data collection for this study. This report would not have been possible without the hard work of these motivated young Haitian men and women who spent five weeks collecting data in the middle of the rainy season across Haiti. The authors would also like to express their appreciation to the ERP partners at Save the Children (Ossemer and James), AIR (Rachelle and Mohammed), UND (TJ, Kate, and Paul), and UNICEF (Mirko, Beatrice, and Vegains) for their assistance in facilitating our work. We also want to acknowledge the public-sector stakeholders at the MENFP who authorized this study and provided valuable input, in particular Harold Narcisse and the Director General, Dr. Meniol Jeune. Finally, we offer our full gratitude to the students, teachers, school directors, parents, and departmental-level MENFP officials who graciously gave their time to take learning tests, participate in observations, and give candid feedback to our team. It is for them and all the students in Haiti that we continue to work hard to ensure that every child everywhere can read.

ABSTRACT

In Fall 2017, Social Impact, through its Haiti-based Evaluation Survey Services (ESS) project, conducted a baseline evaluation of the Early Reading Program (ERP), a USAID-funded intervention focused on improving reading outcomes for Grade 1 and 2 students, teaching quality, and capacity of the Ministry of Education in Haiti to implement and scale up reading programs. The ESS team designed and conducted the baseline for a pre-post impact evaluation (IE) design with statistical matching, administering Early Grade Reading Assessment (EGRA) learning tests, a student questionnaire, a school director survey, a classroom observation tool, focus group discussions with students and parents, and key informant interviews with teachers and MENFP officials across seven departments (Artibonite, Grand'Anse, Nord, Nord-Est, Ouest, Sud, and Sud-Est). The team concluded that students are generally unable to read in Creole. Further, many important factors at the household-level (eating before school, parents' involvement in a student's education, access to books), school-level (director's experience, number of formally trained¹ teachers, existence of extracurricular activities, helping students to pay their tuition) and classroom-level (teachers' use of certain phonics approaches) were associated with higher learning scores. Teachers generally managed their classrooms well (assessed by observing students' behavior in the classroom, utilization of classroom time between start and end times of class, use of varying materials and strategies to teach), though corporal punishment was observed in 11-13 percent of classrooms. Teachers consistently lacked sufficient books and other educational materials and did not generally use best practice phonics approaches for reading. Qualitative data suggested that the MENFP lacks the capacity and resources to support schools and teachers with funding, materials, or training.

¹ A formally trained teacher is one that has a professional degree from a teaching college.

Figure 1: Sampled schools of the ERP Baseline Evaluation



EXECUTIVE SUMMARY

CONTEXT

According to 2015 estimates,² 61 percent of Haitian adults (64 percent of males and 57 percent of females) are literate. Education in the country is plagued by constraints that range from low funding to inadequate teacher training (CIA World Fact Book, 2018). The Ministry of Education and Professional Training (MENFP) is aware of students' poor performance in reading and has worked with the United States Agency for International Development (USAID) in recent years to implement reading interventions to address this situation.

BACKGROUND

In 2017, USAID recognized a need to scrutinize their ongoing investments being made to improve reading outcomes in Haiti. USAID determined that the best way to evaluate these ongoing activities was through an independent impact assessment. The impact evaluation (IE), titled the Early Reading Project, will assess a two-year time period of intervention from multiple different ongoing activities and implementing partners already working to support early grade reading instruction in Grade 1 and Grade 2 of primary school in approximately 658 public, private, community, and semi-private schools in the departments of Artibonite, Ouest, Grand'Anse, Sud, Sud-Est, Nord, and Nord-Est. Broadly speaking, and working across these early grade reading activities in seven different departments, USAID has recognized three common objectives: (1) Improve MENFP's capacity to implement and eventually scale up early grade reading reform; (2) Improve teacher ability to teach early grade reading; and (3) Improve students' early grade reading outcomes.

Partners participating in the Early Reading Project IE include Save the Children (STC), American Institutes for Research (AIR), the University of Notre Dame (UND), and the UN Children's Fund (UNICEF). All of these partners use, or will use, the early grade reading and writing curriculum materials co-developed by the Ministry of Education and USAID. The French version of these materials is called "Je parle bien français." The Creole version of these materials is called "M'ap li net ale". Materials should be used concurrently in all schools. The curriculum employs a scripted model of instruction for teachers and is phonics-based. In first grade, the French portion is completely oral, focusing on building children's vocabulary and knowledge of the French language, while the Creole portion focuses on explicit reading and writing instruction for children, so they are first learning to read in a language they speak and understand. Across partners, the intervention is employed in two different models: "core" and "enhanced." Both models include the implementation of the aforementioned co-designed (USAID and MENFP) national curriculum and teacher training and coaching in early grade reading and writing, but the enhanced model differs in other key extracurricular interventions. The "enhanced" group also contains some UNICEF-supported schools that employ a "pre-Kindergarten" reading model. Baseline data for the Early Reading Project IE was collected in November and December of 2017. Midline data will be collected during the 2018-2019 school year. Endline data will be collected in the spring of 2020.

² CIA World Factbook, updated 1 January 2018. Obtained on 2 March 2018.

PURPOSE AND EVALUATION QUESTIONS

USAID requested that Social Impact (SI)'s Haiti Evaluation and Survey Services (ESS) to design and implement an IE of their ongoing early grade reading and writing investments. The evaluation is intended to measure whether early grade reading and writing interventions have resulted in students in the target schools demonstrating greater skills in reading than their peers of the same grades in non-targeted, similar schools. The evaluation findings will inform USAID/Haiti's future programming in early grade reading in Haitian schools. The results of the evaluation are also intended to inform Government of Haiti policy-level decisions around implementation of reading curriculum and teacher training on a national scale. The lessons learned and recommendations documented in the final endline evaluation report may also inform future programming for other donors' activities related to early grade reading in Haiti. The evaluation's five high-level evaluation questions (EQ) look specifically at the outcomes of (1) reading fluency among Grade 1 and 2 students, (2) teacher quality, and (3) MENFP capacity to implement and scale up reading programs. The questions are as follows:

- Q1. What is the impact, if any, of the “core” ERP intervention on students' reading outcomes in the first and second grades? Some sub-questions the evaluation team will consider in addressing this question are:
- Q2. What is the relative impact, if any, on student learning, of the “enhanced” model carried out by UND and UNICEF? Do students exposed to this model acquire stronger reading skills than their peers in control and other ERP schools? What aspects of the model impact student learning the most?
- Q3. What is the impact, if any, on student learning, of the “pre-reading instruction” model piloted by UNICEF in the pre-K and K levels of 120 of the ERP program schools? Do students exposed to the “pre-reading instruction” model acquire stronger reading skills than their peers in control and other ERP schools?
- Q4. What is the impact of the ERP intervention, if any, on teachers' classroom practices in early grade reading instruction?
- Q5. What is the impact, if any, of the ERP program on the Ministry's capacity to implement and eventually scale up early grade reading reform? Are the human, material, and financial management capacities at the MENFP sufficiently improved because of the ERP intervention that the Ministry could successfully expand the ERP program nationwide?

DESIGN

Haiti ESS followed a pre-post IE design with statistical matching. The evaluation team (ET) collected and analyzed data on equal numbers of girl and boy students in Grades 1 and 2 from participating schools and compared their performance to a control group. To successfully measure the three outcomes of interest, the ET administered Early Grade Reading Assessment (EGRA) learning tests supplemented by student questionnaires, school director surveys, classroom observations, and qualitative data collected through focus group discussions (FGDs) and key informant interviews (KIs). At baseline, the ET administered the EGRA in Haitian Creole with one short EGRA subtask on oral French comprehension.

A sample of 130 schools was randomly sampled proportionally from “core” and “enhanced” implementing partner schools to form the core and enhanced treatment samples, with 65 schools in each sample. At baseline, the ET also identified 65 suitable comparison schools from non-treated schools. After an EGRA validation workshop and school listing activity conducted in October 2017, the ET collected baseline data in November-December 2017. A midline will be conducted in April-May 2019 and an endline in April-May 2020.

FINDINGS AND CONCLUSIONS

GENERAL

The absence of an EMIS system in Haiti hampers educational research and program implementation. The availability of school-level data is essential for conducting rigorous research and for better understanding gaps in educational access throughout the country. Without regular, authoritative census data on schools nationwide, implementers, evaluators, and public-sector actors at the MENFP will be limited in their ability to conduct high quality programming and respond to the needs of communities.

READING FLUENCY

Disproportionate numbers of zero scores have led to a situation where EGRA scores at baseline are partially imbalanced across treatment status. We found that students in Grade 2 in the core treatment group outperform their counterparts from the enhanced and control groups significantly in almost all subtasks (timed and untimed). Variation in EGRA scores at baseline was also noted across genders, school types, and implementers. Among implementers, better reading performance was found among Save the Children schools, which are all located in the Port-au-Prince urban area. Girls score significantly better than boys across all tests (and in both grades), and public and rural schools had lower oral reading fluency scores relative to private, community, religious, and urban schools.

Surveyed students have Creole language skills but do not yet have reading skills. Students in both Grades 1 and 2 scored highest in Creole letter dictation and listening comprehension, but lowest in Creole word dictation and reading comprehension. This result was to be expected in Grade 1 and somewhat in Grade 2 as well, as students begin their formal education, but students in higher grades in primary school (Grade 3 and up) should begin to show higher reading scores. This is in line with other EGRA studies in Haiti that found students were not able to read at sufficient speed (with comprehension to understand text) until Grade 3 or later.³

The home environment plays a critical role in a student's reading scores. At the student-level, students who eat before school, have parents involved in their education, spend time reading, have access to a library, and have books at home have higher oral reading fluency scores than students who do not. Qualitative data showed that parents are involved and invested in their children's education (for both girls and boys). However, they are mostly unaware of PTA activities even though school directors report that nearly three in four schools have active PTAs. Most parents in FGDs noted that school fees are a major barrier to school attendance, although they are willing to go to great lengths to pay for their children's education. While at baseline girls reported taking on a disproportionate amount of housework outside of school, a shift in community attitudes seems to suggest more girls are being allowed to go to school than before.

³ EDDATA II: Education Data for Decision Making. The Early Grade Reading Assessment (EGRA) in Haiti. June 2014. Prepared by RTI International. Contract Number EHC-E-00-04-00004-00. www.eddataglobal.org. The three studies reviewed in the EDDATA paper showed the average third grade Haitian students were able to correctly read between 23-39 correct words per minute (cwpm) in Creole; 35-60 cwpm is considered the minimum rate necessary to comprehend a basic story.

School-level factors explain the largest amount of variation in oral reading fluency for girls and boys. Regression and decomposition analysis found that a host of school-level variables have significant positive associations with reading scores. School management assisting with students' school fees, the experience level of school directors, and number of formally trained female teachers at a school are all associated with positive oral reading fluency scores. Notably, while a minority of schools report having extracurricular programs and services, having these programs was associated with higher oral reading fluency. Furthermore, students were associated with better reading fluency when the average number of hours taught (per day at school) was found to be higher or when the school had a library. As the student-to-teacher ratio increased (more students per teacher), student reading scores decreased. Girls appeared to benefit most from more and better-trained female teachers at a school.

In the classroom, use of certain phonics approaches works better than others, and seating arrangement matters. About half of teachers used a phonics approach while teaching reading, and those teachers who practiced blending sounds and decoding syllables saw better results in reading scores than those who simply asked students to repeat words aloud. Treatment of girls and boys in the classroom was gender equitable; however, regression analysis showed that seating boys and girls together was associated with significant adverse effects on reading.

TEACHING QUALITY

Classroom management is generally good in terms of observed students' behavior in the classroom, utilization of classroom time; however, corporal punishment was observed in some classrooms. Availability, access to, and use of educational materials in the classroom are highly variable. Grade 1 and 2 teachers generally manage student behavior well, start and end class on time, and use a variety of teaching techniques. However, students openly claim that teachers practice corporal punishment, and the practice of corporal punishment was observed in 11-13 percent of classrooms. Teachers generally practice gender equitable treatment of students, with no major statistical differences between how girls and boys are treated in the classroom. Classroom observations noted an inconsistent use of books and a paucity of basic items such as pencils. Qualitative data from MENFP, teachers, and parents suggest a persistent scarcity of books and materials.

Use of pedagogical reading approaches in class is not consistent. Only about half of teachers use a recognized phonics approach during teacher observation. As reported in teacher KIIs, teachers with more years of experience were more likely to be able to specify good phonics practices. Reading aloud to students was the most practiced phonics approach, though it did not necessarily correlate to improved reading scores.

Low levels of teacher formalization probably contribute to low reading scores. Only about a quarter of teachers in sampled schools were formalized (professionally trained) and on-the-job teacher training opportunities are nonexistent or occur very rarely. Most teachers reported teaching for years at a time without being offered training. Many teachers have no formal training. MENFP acknowledges that it lacks the resources to hold regular trainings. This deficit could pose a significant barrier to improvements in reading, given the regression analysis finding that formalized teachers with more experience are associated with higher oral reading fluency scores, especially for girls.

MENFP CAPACITY TO IMPLEMENT AND SCALE UP READING PROGRAMS

Departmental-level MENFP lacks the financial capacity to adequately support schools. Quantitative and qualitative data combine to paint a picture in which schools do not receive books and other educational materials, teachers are sometimes delayed payment, and MENFP-supported teacher

training is not offered on a regular basis. Families must take on the burden of supporting education for their children, and when they cannot, children go to school without necessary supplies or are held back. At baseline, MENFP offices relied heavily on NGOs to support and sustain schools with materials, funds, and teacher training.

Departmental-level MENFP staff appear unaware of policies or plans through which deficiencies in the educational system can be improved. Officials were able to cite problems, such as non-compliance by teachers of the unofficial language policy, high levels of unqualified teachers, and the lack of teacher training, but they could not specify mechanisms or actions through which these issues could be resolved.

MENFP officials carry out monitoring visits often but many do not appear to be deeply knowledgeable about phonics/reading pedagogy. Nearly three in four schools sampled had received an MENFP visit in the past year. However, many MENFP officials interviewed could not discuss reading pedagogy in depth. This lack of knowledge could have implications for policy development and quality oversight of reading programs in the future.

GENDER

In general, findings around gender equity in the household and classroom were positive when it came to education. For instance, parents informed that girls are sent to school at a higher rate than in the past and are treated equally well (or punished equally) compared to boys in the classroom. In fact, qualitative FGDs with parents revealed that parents are allowing girls to go to school more than they have in the past, signaling a shift in attitudes towards girls' education. Some discrepancies remain, however. For instance, more girls than boys are required to do household chores.

Perhaps the most important finding about gender was that girls' reading scores are significantly higher across all measures when compared to those of boys. This finding is not entirely surprising. Studies have shown that girls do better than boys in most subjects, probably due to the fact that girls study more, have more positive perceptions of school and behave better,⁴ according to findings supported by KIIs and FGDs held at baseline. An interesting finding of the regression analysis showed that when a school has more and better-trained female teachers, girls have better reading scores. The finding suggests that female teachers may act as role models for girls, enabling and encouraging them to attend and participate in school.

⁴ Voyer, Daniel, and Susan D. Voyer. "Gender differences in scholastic achievement: A meta-analysis." *Psychological bulletin* 140.4 (2014): 1174.

INTRODUCTION

CONTEXT

According to 2015 estimates, 61 percent of adults in Haiti (64 percent of males and 57 percent of females) are literate. Nationwide literacy rates in Haiti have risen over the last two decades, but when compared globally, Haiti still ranks near the bottom at 141 out of 161 countries. Haiti's literacy rate is well below the regional Latin America and Caribbean average of 90 percent.⁵ Education in Haiti is plagued with constraints that range from low funding to inadequate teacher training. Between 2000 and 2010, the national budget line item for education declined from 17 percent to 10 percent, putting Haiti in the bottom quartile of countries ranked by public spending on education.⁶ The 2010 Haiti earthquake exacerbated these problems, displacing students and damaging schools throughout the country.

The Ministry of Education and Professional Training (MENFP) is aware of students' poor outcomes in reading as well as the system-wide challenges that must be addressed to improve those outcomes. According to the national curriculum, Haitian primary students are expected to study reading in Creole for 90 hours per academic year (Grades 1 and 2) and French for 30 hours (Grade 2 only), or for 40 minutes, three out of five days of the week each school year.⁷ However, reading instruction in the early grades in Haitian schools is often cursory, under-resourced, and ineffective. As a result, children score poorly on standardized reading assessments. Few properly designed instructional materials for teaching reading in the early years are available and teachers are rarely trained in the techniques proven to be most effective in teaching reading. Formative evaluation is rare and lacking in rigor; consequently, students who are struggling to learn to read are often not identified in time for appropriate intervention. System-wide, Haiti practices "automatic promotion" of students in the early primary years, and students who have not learned to read are passed up from grade to grade. Eventually, illiterate students either fail or drop out of school.⁸

In recent years, MENFP has worked with USAID to implement interventions that will improve the likelihood that children will learn to read in Haitian schools. From 2011-2016, with support from USAID, MENFP implemented two reading interventions: the ToTAL and the An n ALE programs. These interventions supported 297 schools in improving reading instruction. Support included scripted reading lesson materials, which did not always arrive on time to be used appropriately in the classroom, and intervention-specific teacher training, which was not implemented sufficiently to improve teacher effectiveness in teaching reading.

EGRA AND ITS USE IN HAITI

The Early Grade Reading Assessment (EGRA) is an assessment that is administered orally to students in the early phases of primary school. EGRA can be used to assess Grade 1-3 student reading skills and identifies gaps or shortcomings that need to be addressed for future student success in reading. Core subtasks that make up the EGRA include letter reading, phonemic awareness, word reading, and story

⁵ CIA World Factbook, updated 1 January 2018. Obtained on 2 March 2018. <https://www.cia.gov/library/publications/the-world-factbook/geos/ha.html>

⁶ Suzata, Eriko. 2011. Education in Haiti: An Overview of Trends, Issues, and Plans. World Innovative Summit for Education.

⁷ EDDATA II: Education Data for Decision Making. The Early Grade Reading Assessment (EGRA) in Haiti. June 2014. Prepared by RTI International. Contract Number EHC-E-00-04-00004-00. www.eddataglobal.org.

⁸ USAID. Haiti ESS ERP Impact Evaluation Scope of Work. October 2017.

reading with comprehension questions. Other EGRA subtasks may include pre-literacy skills, basic vocabulary, listening comprehension, and sentence or non-word writing.⁹ Oral reading fluency, when coupled with comprehension, can be considered a measure for basic reading competencies. Overall, EGRA provides a methodology for quickly assessing a variety of early literacy skills. In general, EGRA identifies strengths and weaknesses in student reading which in turn indicate the quality of learning taking place in schools.¹⁰

EGRA has been used in Haiti to evaluate both reading programs and the general state of literacy. As part of the USAID EQUIP2 project implemented by Save the Children in 2007, EGRA effectively revealed a high degree of variability in reading scores among private schools in Haiti; the EGRA found no evidence that private or public schools were doing a better job of teaching reading. A study conducted by RTI in 2008-2009 used the EGRA to assess reading in Grades 1-3 in the Nippes and Artibonite departments. The assessment found that while there were no differences between male and female test-takers, there was a clear pattern of higher literacy among students who possessed books at home.

Use of EGRA in the USAID PHARE project, implemented by AIR in 2009, demonstrated no significant differences in reading scores between public and private schools or between the sexes. The EGRA assessment did, however, reveal large differences in reading scores between geographic departments. In 2013, Concern International noted high levels of illiteracy in the higher grades and expanded its Saut d'Eau project to include an EGRA assessment for Grades 1-6. This example proves that EGRA allows flexibility and adaptability and facilitates tailored approaches to specific local contexts. The 2011-2016 implementation of the USAID-funded ToTAL project is the most recent use of EGRA in Haiti. The ToTAL project linked EGRA results to results from a Snapshot of School Management Effectiveness (SSME) survey, which allowed for analyzing reading outcomes in a more holistic way. In linking SSME data to EGRA results, ToTAL staff sought to help local stakeholders by identifying helpful take-away points that could be applied at the classroom, school, and district levels.

PROJECT BACKGROUND

THE EARLY READING PROJECT PROGRAM

In 2017, USAID recognized a need to scrutinize their ongoing investments being made to improve reading outcomes in Haiti. USAID determined that the best way to assess these ongoing activities was through an independent impact evaluation (IE). The IE, titled the Early Reading Project, will analyze a two-year time period of intervention from multiple different ongoing activities and implementing partners already working to support early grade reading instruction in Grade 1 and Grade 2 of primary school in approximately 658 public, private, community, and semi-private schools in the departments of Artibonite, Ouest, Grand'Anse, Sud, Sud-Est, Nord, and Nord-Est. Broadly speaking, and working across these early grade reading activities in seven different departments, USAID has recognized three common

⁹ Core subtasks or common instrument components of the present EGRA testing tool consist of listening comprehension, letter identification, familiar or non-word reading, oral reading fluency, and initial or final sound identification. Other components that do not fall under core subtasks are dictation and phoneme segmentation. Earlier EGRA tools included tests for pre-literacy skills such as picture-based subtask (as those in the Peabody Picture Vocabulary Test (PPVT)), or pictures to identify knowledge of common vocabulary such as words for body parts like hand, head, or toe. Print assessments (a student's ability to indicate where on a page to begin reading or where to read next, for example) have also been eliminated from the current EGRA (Early Grade Reading Assessment Toolkit, Second Edition, 2016).

¹⁰ RTI International. 2015. Early Grade Reading Assessment (EGRA) Toolkit, Second Edition. Washington, DC: United States Agency for International Development.

objectives: (1) Improve MENFP's capacity to implement and eventually scale up early grade reading reform; (2) Improve teacher ability to teach early grade reading; and (3) Improve students' early grade reading outcomes.

Partners participating in the Early Reading Project IE include Save the Children (STC), American Institutes for Research (AIR), the University of Notre Dame (UND), and the UN Children's Fund (UNICEF). All of these partners use, or will use, the early grade reading and writing curriculum materials co-developed by the Ministry of Education and USAID. The French version of these materials is called "Je parle bien français." The Creole version of these materials is called "M'ap li net ale". Materials should be used concurrently in all schools. The curriculum employs a scripted model of instruction for teachers and is phonics-based. In first grade, the French portion is completely oral, focusing on building children's vocabulary and knowledge of the French language, while the Creole portion focuses on explicit reading and writing instruction for children so they are first learning to read in a language they speak and understand. Across partners, the intervention is employed in two different models: "core" and "enhanced." Both models include the implementation of the aforementioned co-designed (USAID and MENFP) national curriculum and teacher training and coaching in early grade reading and writing, but the enhanced model differs in other key extracurricular interventions (Table 1). Baseline data for the Early Reading Project IE was collected in November and December of 2017. Midline data will be collected during the 2018-2019 school year. Endline data will be collected in the spring of 2020.

Table 1: Partners and their SOWs for the ERP Impact Evaluation

Partner	Subcontractor	Model	Year 1 Schools	Year 2 Schools	Department(s)
American Institutes for Research (AIR)	FONHEP	Core	150	150	Nord, Nord-Est, Ouest, Artibonite
Save the Children	N/A	Core	39	39	Ouest
University of Notre Dame (UND)	Catholic Relief Services (CRS); CEEC	Core and Enhanced: School libraries and summer camps	160	160	Grand-Anse, Sud
United Nations Children's Fund (UNICEF)	CARE; PLAN	Core and Enhanced: Reading clubs and PTAs; Pre-Scolaire program	106	106	Nord, Nord-Est, Sud-East

DEVELOPMENT HYPOTHESIS

The evaluation team has synthesized the development objectives and hypothesis as follows:

DEVELOPMENT OBJECTIVES:

1. Increase literacy through the provision of core skills and organizational reforms to MENFP (for example, improved coordination, consistent and effective capacity building, improved leverage among education sector actors and networks, and creation of strong feedback loops between classroom and program designers) to assist learners in 670 schools in mastering basic literacy skills.
2. Increase proportion of students who can read and demonstrate reading comprehension by the end of the primary grades as defined by the Haitian curriculum, standards, or as agreed-upon by national experts.
3. Increase proportion of students who demonstrate improvement in reading and understanding of grade level texts by the end of two grades of primary schooling.

DEVELOPMENT HYPOTHESIS: Effective implementation of USAID-supported early grade reading and writing interventions can lead to improved reading outcomes through:

1. Provision of necessary support and technical guidance to the MENFP;
2. Implementation of the USAID/MENFP early grade reading and writing curriculum in primary schools (even if only in an experimental stage);
3. Establishment of strong feedback loops between all types of schools implementing the program and the design team in Port-au-Prince.

EVALUATION PURPOSE AND AUDIENCE

PURPOSE

USAID requested that SI's Haiti Evaluation Support Services (ESS) design and implement an independent IE of its early grade reading and writing interventions across partners. The purpose of the evaluation is to assess the overall effectiveness of the program on early grade reading among Grade 1 and 2 students. The IE is intended to measure whether the ERP interventions have resulted in students in the target schools demonstrating greater skills in reading than their peers of the same grades in non-targeted, homologous schools. The evaluation findings will inform USAID/Haiti's future programming in early grade reading in Haitian schools. The results of the evaluation are also intended to inform Government of Haiti policy-level decisions around implementation of reading curriculum and teacher training on a national scale. The lessons learned and recommendations documented in the final endline evaluation report may also inform future programming for other donors' activities as they relate to early grade reading in Haiti.

PRIMARY STAKEHOLDERS

The primary stakeholders for this evaluation include the MENFP and USAID/Haiti, the implementing partners evaluated under the Early Reading Project IE; the Latin American and Caribbean (LAC)/Reads program; UNICEF; University of Notre Dame (UND); Save the Children; The United Nations Office for Project Services (UNOPS)); and other donors and entities active in Haiti's education sector, including the World Bank, Japan International Cooperation Agency (JICA), Inter-American Development Bank (IDB), and World Vision, as well as a variety of non-governmental organizations (NGOs) and religious affiliations devoted to improving learning in Haitian primary schools.

OUTCOMES OF INTEREST

READING/LITERACY: In particular, the evaluation team (ET) looked at the following skills: letter-sound mastery (phonetic awareness); familiar word reading; reading connected text/oral reading fluency (ORF); reading comprehension; writing a sentence dictated to the student; and writing a sentence the student generates in response to an open verbal prompt.

TEACHING QUALITY: The ET examined certain aspects of teaching quality, including classroom management, use of phonics approaches, use of educational and teaching materials, gender equity, and corporal punishment. Observing these characteristics also allowed the team to better understand how teachers applied their training in the classroom and whether that training had a beneficial effect on student performance.

MENFP CAPACITY: The ET looked at the activities under the Early Reading Project targeting the MENFP and their contributions to increasing knowledge of and capacity to implement reading programs. The team also examined the existence of financial and human resources necessary for the implementation, continuation, and/or scaling up of early grade reading programs.

EVALUATION QUESTIONS

USAID and the ESS team developed the following evaluation questions (EQ) to guide the IE.

EQ1. What is the impact, if any, of the “core” ERP intervention on students’ reading outcomes in the first and second grades? Some sub-questions the evaluation team will consider in addressing this question are:

- a. How do students exposed to any ERP programming perform when compared to similar students who are not exposed to ERP programming?
- b. What is the marginal effect, if any, of ERP implementer on reading outcomes?
- c. What is the marginal effect, if any, of type of school (semi-private, public, community, religious, etc.) on reading outcomes?
- d. To what extent do environmental or situational factors have an impact on students’ ability to learn to read in the ERP target schools?

EQ2. What is the relative impact, if any, on student learning of the “enhanced” model carried out by UND and UNICEF? Do students exposed to this model acquire stronger reading skills than their peers in control and other ERP schools? What aspects of the model impact student learning the most?

EQ3. What is the impact, if any, on student learning, of the “pre-reading instruction” model piloted by UNICEF in the pre-K and K levels of 120 of the ERP program schools? Do students exposed to the “pre-reading instruction” model acquire stronger reading skills than their peers in control and other ERP schools?

EQ4. What is the impact of the ERP intervention, if any, on teachers’ classroom practices in early grade reading instruction?

- a. Do teachers exposed to the “enhanced” ERP intervention teach reading differently and/or better than their peers in the “core” ERP schools?
- b. Do teachers exposed to the ERP intervention teach reading differently and/or better than their peers in the identical or close-to-identical schools without ERP interventions?
- c. How, if at all, do teachers exposed to the ERP intervention teach reading differently after having received the ERP intervention?

EQ5. What is the impact, if any, of the ERP program on the Ministry’s capacity to implement and eventually scale up early grade reading reform? Are the human, material, and financial management capacities at the MENFP sufficiently improved because of the ERP intervention that the Ministry could successfully expand the ERP program nationwide?

METHODOLOGY

EVALUATION DESIGN

IMPACT EVALUATION – QUASI-EXPERIMENTAL DESIGN

To address the above evaluation questions, Haiti ESS will follow a pre-post IE design with statistical matching. The evaluation seeks to collect and analyze baseline, midline, and endline data on girl and boy students in Grades 1 and 2 from participating schools and compare their respective grade-level performance to a control or comparison group.

Since a randomized control trial, the gold standard for definitively assigning observed outcomes to a program, was not feasible (given that the ERP schools had already been assigned by the implementing partners), a quasi-experimental design (QED) was required to best evaluate the ERP impact. Coarsened

Exact Matching (CEM) was recommended and ultimately chosen as the best QED for the Haitian context. CEM is used to identify a control group that is almost identical to the treatment groups.

The evaluation design included two treatment groups (as opposed to designing two separate evaluations) to allow for isolation of the effects of enhanced treatment (“enhanced” group) over the “core” group. This approach increases the efficiency of the evaluation by providing supplemental information about the different characteristics of the two interventions. With two treatment arms, however, it becomes difficult to identify a common control group that is almost identical in all aspects to both treatment arms, particularly when the selection criteria for different treatment groups are different. To overcome these difficulties, the ET used basic school-level characteristics (such as teacher qualifications and experience, access to infrastructure, school management indicators, and student-teacher ratio) that were easily available across the control and both treatment groups and matched two different treatment arms separately with control schools. The ET then combined (merged) these matched groups over common control schools so that each control school would correspond to both types of treatment schools. The team discarded the control schools that matched with one treatment group, but not the other. A full explanation of how CEM is used and its benefits and limitations can be found in Annex C.

TREATMENT GROUPS

The study includes samples from two treatment groups, which are identified based on the intervention type: one “core” group and one “enhanced” group. The main differences between the two groups can be found in Table 1. The **core treatment group** consists of 189 schools benefitting from interventions by Save the Children (39 schools in Ouest department) and AIR (150 schools in Nord, Nord-Est, Ouest, and Artibonite departments). These schools receive the co-developed MENFP and USAID curriculum as well as teacher training in implementing the curriculum. The ESS team randomly sampled 65 schools from these 189 schools in the core treatment group to form the core treatment sample.

Table 2: List of ERP implementers and evaluation sample schools

Implementer	Total number of schools with interventions across Years 1 and 2	Number of schools in evaluation sample: CORE	Number of schools in evaluation sample: ENHANCED
AIR	150 (33%)	49 (38%)	--
Save the Children	39 (9%)	16 (12%)	--
UND	160 (35%)	--	37 (28%)
UNICEF	106 (23%)	--	28 (22%)

Note: Figures in bracket indicate proportional distribution in the treatment groups and treatment sample. The sum of treatment schools does not include TBD schools during time of sampling.

The enhanced treatment group consists of 269 schools benefitting from interventions implemented by UND (160 schools in Sud and Grand’Anse departments) and UNICEF (106 schools in Nord, Sud-Est, and Nord-Est departments). While all schools in the enhanced group also benefit from the core group’s

curriculum and teacher training, several UND schools introduced additional activities, including school libraries and summer camps, and some UNICEF schools have reading clubs, PTA-focused activities, and Pre-K interventions. It is these extracurricular activities that differentiate the core from the enhanced group. A total of 65 schools were randomly selected to form the enhanced treatment sample for the evaluation.

SELECTION CRITERIA FOR TREATMENT SCHOOLS: Prior to ERP intervention implementation, targeted schools were chosen by the four implementing partners based on a set of pre-determined criteria. The ET utilized the same criteria to the maximum extent possible to identify and randomly select the control group of schools. The implementing partners' specific school selection criteria, which varied across partners, can be found in Annex C.

COMPARISON/CONTROL GROUP

To measure the impact of the interventions, the evaluation strategy requires the identification of a similar group of schools and students to enable comparison of the outcomes achieved by the treated group.¹¹ To identify a suitable comparison group from non-treated schools, the ET required access to robust school-level data. However, Haiti lacks an Education Management Information System (EMIS) from which school-level data can be easily obtained. To account for this, the ET used the same lists of schools that each implementing partner (IP) used to identify treatment schools. Lists from all four IPs were pooled to create a sampling frame that contained a total of 2,734 schools across seven departments. From this pooled list of schools, the ET randomly selected a total of 355 schools (160 treatment and 195 control) to conduct a rapid survey of schools (known as a school listing exercise) to perform the matching.

After the partial list was created, the ET conducted the school listing activity. (A detailed description of the listing activity is provided in the Data Collection Activities section.) Enumerators collected data from the 355 randomly selected treatment and control schools mentioned earlier. Data collection focused on questions related to school-selection criteria used by each implementer, as well as other questions on location, typical school demographics, basic enrollment, and existence of Grade 1 and Grade 2 students and teachers. Using CEM (explained above), data collected from non-treatment schools through the listing activity were then compared to the treatment schools and matched along a set of defined variables. At the end of the initial round of the CEM matching process, 65 control schools that matched the 130 treatment schools were randomly identified to form the common control group. The ET then used this total group of 195 schools to conduct the full-fledged data collection.

QUALITATIVE SUPPLEMENTATION

To supplement the quantitative data, SI collected qualitative data about the program's impacts. Whereas the quantitative data will elucidate the program's impact, qualitative data will provide deeper insight into the mechanisms for how and why these impacts took place. At baseline, qualitative data was collected via two methods—focus group discussions (FGDs) and key informant interviews (KIIs). FGDs and KIIs were carried out with six distinct populations: boy students, girl students, mothers of students, fathers of students, teachers at school, and MENFP departmental-level officials. Approximately six to 10

¹¹ A comparison group is a group that is similar in the best possible way to the target group, except that it has not been exposed to the project intervention in any way.

participants were chosen for each FGD via convenience sampling. KIs consisted of one to two respondents. Qualitative events were carried out at both control and treatment schools at baseline.

SAMPLING

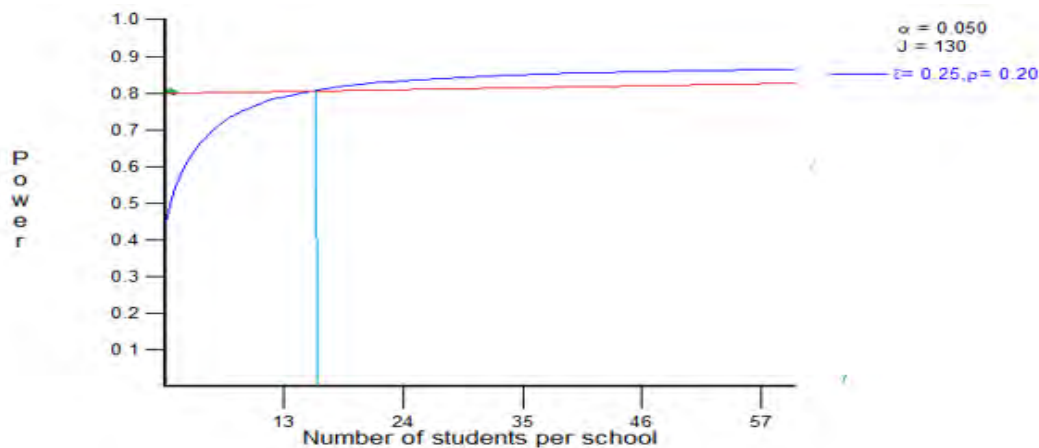
POWER ANALYSIS

Statistical power informs the reader about the likelihood that the study will detect an effect where there is an effect there to be detected. If statistical power is high, the probability of concluding there is no effect when, in fact there is one, reduces. Generally, a statistical power of 80 percent and above is enough to distinguish an effect with confidence. A study that is sufficiently powered also ensures external validity of the evaluation results.

For the ERP evaluation, sample selection was conducted at the school level while outcomes were measured at the student level. Due to a lack of school level information and student performance typically found in a functioning EMIS, the ET used common hypothetical parameter values during the design period. A minimum detectable effect size (MDES) of 0.25 standard deviation (SD) in treatment effect was used between baseline and endline, as demonstrated in other similar EGRA studies conducted in Haiti.¹² The ET also considered a conservative value of intra-class correlation (ICC) of 0.2 to capture grade and higher-level variation. Given that clustering affects the optimal number of students sampled per school, the team assessed the tradeoff between adding more schools or more students per school to reach a fixed sample. Based on these parameters and taking cost, time, and school enrollment numbers into consideration, the team estimated that 65 schools per treatment arm with at least 16 students per school was optimum to maintain 80 percent of analytical power (Figure 1). To mitigate the risk of sample size reductions due to day-of absence of students across schools, the ET increased the total sample size per school by two to 18 students. These 18 were randomly selected at each school—nine from each grade—and equally stratified across gender to the extent possible. In the case that there were multiple classes per grade level, the team sampled from one classroom only per grade.

¹² Our consideration of effect size to be 0.25 SD comes from the ToTAL study in Haiti. Based on the individual effect sizes reported for listening comprehension, letter sound and name identification, familiar word reading, and oral reading fluency under treatment A (Table 2: Summary of EGRA 2012–2013 Scores by Grade and Language, pp 17-18 of ToTAL, 2015), we calculated the average effect size between 0.227 SD (with listening comprehension) and 0.276 SD (without listening comprehension). We used only Treatment A, because Treatment A schools were supported with all the necessary materials, teacher training, and ongoing support.

Figure 2: Sample size calculations and power analysis for the ERP Impact Evaluation



Note: We used Optimal Design (OD) statistical package for sample size calculation. The parameters used for this calculation are $\alpha=0.05$; number of schools per arm=65; MDES (δ)=0.25 and Intra-class Correlation (ρ)=0.2. Based on these parameters, the number of students per school required is estimated to be 16 to maintain at least 80 percent power.

QUALITATIVE SAMPLE

For the qualitative sample, the ET decided to stratify the six qualitative events equally across the seven departments, with one FGD type (boys, girls, mothers, or fathers) per department and two KIIs (teachers, MENFP personnel) per district. This resulted in a total sample size of 56 events, and approximately 200-300 participants (Table 3).

Table 3: Planned qualitative sample size at baseline

Type	Location	Targeted respondents	Number of planned events	Expected sample size per event	Rounds
FGD	School	Grade 1 + 2 boys	7	6-10	BL, EL
FGD	School	Grade 1 + 2 girls	7	6-10	BL, EL
FGD	Community	Mothers of students	7	6-10	BL, EL
FGD	Community	Fathers of students	7	6-10	BL, EL
KII	School	Teachers	14	1-2	BL, EL
KII	Departmental-level MENFP	MENFP personnel	14	1-2	BL, EL
Total			56	200-300	

DATA COLLECTION ACTIVITIES

DESK REVIEW

The ET carried out an initial desk review of available documents as the first step in preparing for the evaluation. The desk review included documents, reports, and tools related to previous early grade reading evaluations conducted in Haiti and of the EGRA tool itself, in general. Documents were selected for their relevance to the evaluation, such as those presenting best practices in early grade reading evaluations, EGRA usage, and supplemental tools for evaluating issues such as teacher training and the school environment. The review identified current data gaps and informed the development of the data collection tools that would be

used to gather the primary data for answering the EQs. In total, the review covered approximately six key documents and more than 10 tools from previous evaluations.

Key findings from this desk review included: (1) To be most effective, the EGRA tool should be adapted to the current context and validated by all relevant partners; (2) Quantitative data is best triangulated through the use of qualitative data to better contextualize and add nuance to the numbers; (3) The school and home environment are critical factors for understanding reading scores and, to that end, evaluations do well to include snapshot surveys of the school and home environment; (4) Teacher instruction and teacher quality play a large role in reading scores—linking teacher observation data to reading results is critical; and (5) Reporting objective disaggregated data from the seven EGRA subtasks is important for understanding a holistic picture of reading. More details on the Desk Review phase can be found in the Annex C.

SCOPING TRIP

To better prepare for and facilitate a smoother, higher quality baseline study, the ET undertook a brief scoping trip in early October 2017. The goals of the scoping trip were to: (1) obtain a better sense of the geographic scope of ERP, geographic project areas, and other donor activities in study locations, (2) coordinate with stakeholders, implementing partners, and the local data collection firm for the school listing activities from which to construct the sampling frame, (3) meet face-to-face with other stakeholders to discuss likely cooperation and coordination and obtain grey literature that might not be available online, and (4) plan for full data collection with the local data collection firm. All information gathered during the scoping trip, including the list of schools, was utilized to prepare and plan the final evaluation design.

EGRA VALIDATION WORKSHOP

Per the findings of the desk review, at baseline the ET adapted the EGRA tool. The original tool, developed by Research Triangle International (RTI) in 2009, was available in Creole and French. To ensure validation by all parties and context appropriateness, the ET carried out a validation workshop in late October 2017. Over the course of three days, the EGRA tool was discussed and adapted with input from key stakeholders including MENFP, USAID, ERP implementers, IFOS (the data collection firm), and Social Impact. The tool was also piloted in two schools with 50 children, the pilot data was analyzed, and the tool was further revised based on the findings. Workshop results led to dropping the EGRA in French for Grade 1 and dropping the non-familiar word sub-task in Creole for both grades. Upon finalization of the tool, the entire research protocol was submitted to a local Institutional Review Board (IRB) in Haiti as well as Social Impact's IRB. Both institutions subsequently approved the research.

SCHOOL LISTING ACTIVITY

As explained earlier in this report, to obtain a sample frame from which to conduct the CEM matching necessary to construct a control group, the ET, in conjunction with IFOS, conducted a school listing activity in late October through early November 2017, just prior to data collection. The ET visited 160 treatment and 195 control schools, selected randomly from lists of schools used by implementing partners. IFOS conducted the activity over a period of three weeks, administering a short 10-minute survey to head teachers at each school. A key finding from the school listing activity was that many schools have very low enrollment numbers, at times barely reaching the 18 students (nine each from Grades 1 and 2) necessary to ensure statistical validity. Any school with fewer than this required number was excluded from the school list. This finding allowed the ET to proactively mitigate delays associated with insufficient sample sizes.

TRAINING, PILOTING, AND PRE-TESTING TOOLS

From November 6-12, 2017, the ET worked with a local data collection firm to facilitate a six-day training for all enumerators, supervisors, and managers in preparation for baseline data collection. The training introduced the program and evaluation goals and touched upon the following specific aspects: the sampling approach, data collection methodologies, data collection ethics, introduction to the various tools (both qualitative and quantitative), and hands-on, field-based practice. A day of field-testing was conducted with students, teachers, head teachers, and community members in 15 local schools (schools not involved in the study). The tools, especially the head teacher survey, classroom observation tool, and qualitative tools, were revised based on findings from the training and pre-test. All enumerators underwent a final test in which their EGRA inter-rater reliability (IRR) scores were compared to ensure consistency across marking of the various tools. The lowest performing enumerators were not selected to continue with the full fieldwork. In total, 122 enumerators, supervisors, and managers were trained over the six-day training. IRR results from the training can be found in Annex J.

DATA COLLECTION

Baseline data collection took place between November 13 and December 8, 2017, in the seven targeted departments. In total, 15 teams of seven enumerators and supervisors each were dispatched across the seven departments, managed by two to three field coordinators. Given that annual testing for all schools began on Monday, December 11, it was imperative that all data collection was completed by Friday, December 8. Most of data collection was completed by Friday, December 1, with a small number of schools revisited the following week and completed by December 8. Enumerator teams faced many challenges during data collection, notably torrential rains and flooding in Sud and Grand'Anse and school closures in the Nord due to a Presidential visit. In total, 17 of 195 schools (9 percent) were replaced due to insufficient sample sizes or head teacher refusal to participate. Administration and testing of all tools was completed on electronic tablets using Open Data Kit (ODK) software (for the head teacher survey and classroom observation tools) and Tangerine software (for EGRA). Qualitative data collection was carried out in tandem with quantitative data by separate teams. All qualitative events were digitally recorded and later transcribed and translated. It should be noted that as far as qualitative data was concerned, the ET achieved data saturation (no new major findings) approximately halfway through data collection. To mitigate time and budgetary risks associated with this finding, the ET may reduce the number of qualitative events in future data collection rounds.

DATA COLLECTION TOOLS

All data collection instruments are explained below. The full tools are included in Annex B.

EARLY GRADE READING ASSESSMENT

The centerpiece of the data collection methodology for this evaluation was the Early Grade Reading Assessment (EGRA) tool, which was used to collect student-level reading scores. After adaptation, the team used the EGRA tool originally prepared by RTI for earlier EGRA evaluations in Haiti, which was available in both French and Creole. The EGRA tool contained the following key sub-tasks: (1) letter-

sound mastery (phonetic awareness), (2) familiar word reading, (3) non-familiar word reading,¹³ (4) reading connected text/oral reading fluency (ORF), (5) reading comprehension,¹⁴ (6) writing of a sentence dictated to the student, and (7) writing of a sentence the student generates in response to an open verbal prompt (listening comprehension). The EGRA was administered by tablet in a face-to-face manner with each child in a school setting and took an average of 24 minutes to complete at baseline. The EGRA helps to answer the key evaluation questions by examining changes in question responses between baseline, midline, and endline, and by comparing control and intervention groups. The EGRA will be repeated with students from the same grades and schools at midline and endline.

The EGRA was administered to students at school, typically in the morning between 8 a.m. and noon. Head teachers were made aware one to two days ahead of a visit by enumerators and prepared spaces where students could be tested. Upon arrival at a school, enumerators obtained permission from the head teacher to test students, and then proceeded to Grade 1 and Grade 2 classrooms to select students for testing. Nine students each in Grades 1 and 2 were randomly selected, stratified equally to the extent possible by gender. These pre-data collection activities were carried out prior to the start of the school day to minimize disruption. After the 18 total students were chosen, enumerators periodically removed them from class to administer the EGRA test. Tests were administered in quiet spaces outside the classroom, face-to-face, and in line of sight of other adults while minimizing noise disruption. EGRA tests took approximately 24 minutes to administer and once a test was complete, the student was walked back to his or her classroom and another student chosen. This process was repeated until the maximum number of 18 students were tested.

HEAD TEACHER/SCHOOL SURVEY

Based on the Snapshot of School Management Effectiveness (SSME) as designed under USAID's EdData II project, the team utilized a head teacher survey (also known as a school survey) as a means of capturing school-level environmental information that could explain key outcomes in schools across communes and departments. The survey focused on the following areas: (1) pedagogical leadership and management; (2) class and classroom management, (3) school management, (4) parent and community involvement in the school, and (5) district- and system-level support and supervision. Each school's head teacher was the survey target or, if the head teacher was unavailable, the survey was conducted with the next most knowledgeable person about the school. The survey was administered by tablet and took approximately 30 minutes to complete at baseline. To minimize disruption, the head teacher was interviewed either before school began or at the end of the school day.

¹³ Non-familiar word reading was removed from the EGRA for both grades after input at the validation workshop. For word knowledge skills, the ET used the lexical route for familiar words only and did not include non-word reading or writing to assess students' skills in decoding (known as the sub-lexical route). The non-familiar word strategy is generally used when familiar words are too familiar for students. In such a case, the identification of non-familiar words through the sub-lexical route presents a more relevant challenge for the students. However, given low scores overall as reported in previous evaluations for Haiti, particularly for the Grade 1 students, a decision was made to take out the sub-task. SI is open to including the sub-task for Grade 1 students in subsequent data collection phases (midline and/or endline) upon USAID request.

¹⁴ While both grades will receive the full Creole EGRA, only Grade 2 will receive a full French version at midline and endline. Grade 1 will receive a short section on French oral reading comprehension as part of the Creole version.

CLASSROOM OBSERVATION TOOL

A classroom observation tool was used to help capture descriptions of a variety of practices in pedagogy and classroom management. The tool used was modeled after the Stallings' Tool¹⁵ adapted for the ERP and Haitian context. Upon arrival at a school, an enumerator consulted with the head teacher and randomly chose two classrooms (one each for Grades 1 and 2) to be observed. Each classroom received approximately 40 minutes of observation, stratified into four blocks of time, where the first block of ten minutes was a waiting period with no observation. For the remaining three blocks of time, an enumerator observed for 10 minutes then answered a set of questions based on observations for the previous 10 minutes only. Observations were recorded by a single enumerator, sitting to the side or in the back of the class to minimize disruption, per classroom. The tool was not prescriptive in that it remained agnostic to “good” and “bad” teaching practices in the Haitian context, but rather captured a teacher’s actions with regard to classroom management, use of educational materials and resources, phonics and reading practices, and gender equity. Data from the classroom observation was ultimately linked to EGRA scores to tease out the effect of certain classroom-level characteristics on reading scores.

STUDENT QUESTIONNAIRE

Prior to administering the EGRA to a student, enumerators carried out a student questionnaire consisting of 40 questions that asked the student about his or her household and school experience. The questionnaire captured household-level topics such as language spoken at home, if a child ate before going to school, school absence and reasons for absence, schoolwork completed at home, existence and use of books at home, if a parent reads to a child at home, and marks in school. Classroom-level questions asked if and to what extent teachers and students read aloud in class, as well as questions regarding teachers correcting homework. Findings from the student questionnaire can be found in this report in the section on descriptive statistics of the sample and as key explanatory variables in the regression models.

KEY INFORMANT INTERVIEW AND FOCUS GROUP DISCUSSION GUIDES

The ET developed FGD and KII guides with specific focus on questions that addressed the five EQs. These guides were sent to the data collection firm for translation and pre-testing. After the pre-test at the enumerator training, the guides were further revised to reduce and streamline questions. FGDs and KII facilitators and note-takers underwent an intensive 5-day training during the full data collection training in early November 2017. The firm accounted for gender-specific considerations, such as ensuring that qualitative events with females (girls, mothers) were carried out by female moderators. In the end, enumerators completed 28 FGDs and 28 KIIs with boy students, girl students, mothers of students, fathers of students, teachers at school, and MENFP departmental level officials, stratified across departments. Each FGD lasted approximately 90-120 minutes, with KIIs lasting 60-90 minutes. FGD and KII participant selection was typically done by purposive sampling, with head teachers selecting which students, parents, and teachers to interview. The data collection firm was given a specific list of departmental-level MENFP personnel to interview.

¹⁵ World Bank. 2015. Conducting classroom observations: analyzing classrooms dynamics and instructional time - using the Stallings 'classroom snapshot' observation system: user guide (English). Washington, D.C.: World Bank Group.

DATA ANALYSIS

BASELINE DATA COLLECTION

At baseline data collection, the ET conducted several types of analysis with quantitative and qualitative data. **Quantitative data analysis** was conducted in Stata software and included: (1) CEM matching and balancing, (2) summary statistics for outcomes and covariates by treatment and control groups, in part to determine the statistical balance of the control and treatment groups at baseline, (3) descriptive statistics of outcome variables and covariates, with the objective of determining if there were any pre-existing gaps in outcomes between treatment and control groups, and (4) regression and decomposition analyses of potential explanatory variables on the outcomes of interest (reading fluency). **Qualitative data analysis** was conducted in Dedoose software. The analysis included: (1) development of a codebook, (2) initial reviews of imported transcribed (and sometimes translated) qualitative data, (3) coding of transcripts based on the pre-defined codebook, and (4) Grouping, tallying, and content and pattern analysis to analyze the qualitative data as it pertained to the EQs.

ANALYZING EGRA SCORES

Evaluation Questions 1 and 2 look at the impact of core and enhanced ERP interventions on student reading outcomes in Grades 1 and 2. Reading fluency and comprehension are higher-order skills in the reading acquisition process, and they build upon several lower-order, foundational skills such as phonological awareness, alphabet knowledge, decoding and vocabulary, all of which can be detected through an oral assessment like EGRA. Lower-order skills also include initial sound recognition and identification of letters and letter sounds. At baseline, the ET administered the tests in Haitian Creole only, with one short subtask on oral French comprehension. The EGRA subtasks include the following:

Table 4. EGRA subtasks

Untimed	Timed
Creole letter dictation	Correct letters per minute (CLPM)
Creole word dictation	Correct letter sounds per minute (CLSPM)
Initial sound recognition	Correct words per minute (CWPM)
Listening comprehension	Oral reading fluency (ORF)
Reading comprehension	--

The EGRA was administered to students at school, typically in the morning between 8 a.m. and noon. Head teachers were made aware one to two days ahead of a visit by enumerators and prepared spaces where students could be tested. Upon arrival at a school, enumerators obtained permission from the head teacher to test students and then proceeded to Grade 1 and Grade 2 classrooms to select students for testing. Nine students each in Grades 1 and 2 were randomly selected, stratified equally to the extent possible by gender. These pre-data collection activities were carried out prior to the start of the school day to minimize disruption. After the 18 total students were chosen, enumerators periodically removed them from class to administer the EGRA test. Tests were administered in quiet spaces outside the classroom, face-to-face, and in line of sight of other adults while minimizing noise

disruption. EGRA tests took approximately 24 minutes to administer and once a test was complete, the student was walked back to his or her classroom and another student chosen. This process was repeated until the maximum number of 18 students were tested.

DATA QUALITY ASSURANCE MEASURES

The ET observed a set of stringent data quality assurance (DQA) measures throughout the data collection process. For instance, to minimize data entry errors and obtain data quickly for verification, the ET utilized electronic data collection tools to collect most of the data. The ET vetted, tested, and re-tested the data collection instruments, and then trained, tested, and supervised enumerators. Additionally, the ET programmed the tablets with internal quality checks— skip patterns and logic checks— and drop-down menus to ensure accuracy of certain key fields. Further, timestamps and GPS point verification were used to ensure that enumerators were administering surveys at appropriate speeds and that surveys took place where they were supposed to. The data was uploaded to servers directly from each tablet daily for real time review. The ET also hired an external consultant who conducted surprise spot visits on enumerator teams to verify protocols were followed. The consultant and the ET also conducted daily data verifications for all four weeks of data collection. For more information on DQA measures employed by the ET, please see Annex I.

WEIGHTING

To ensure that results are regionally representative, best practice in statistical analysis dictates that the ET apply weights to the results. In the absence of a nationally (or at least regionally) representative sample frame, the ET did not apply weights. The team randomly selected treatment schools from the lists of schools available from IPs. The students from Grades 1 and 2 in each school were also randomly selected to maintain external validity.

LIMITATIONS

Table 5 below details the main limitations associated with the study design, the risk associated with each limitation and the mitigation measures used at baseline to minimize negative impacts of the limitation.

Table 5: Study limitations

Theme	Risk	Mitigation Measures
Child protection	Handling a vulnerable section of the population; children in the study may be as young as five years old. The risk of child protection issues could be high.	All team members are NIH-certified. Child protection and data collection ethics were also reviewed at the training and strictly maintained. The ET obtained ethical approval of the study from a local Haitian IRB and Social Impact's internal IRB.

Theme	Risk	Mitigation Measures
Low sample size/attrition	There is a risk of not enough children in Grades 1 or 2 in some schools to arrive at a sufficient statistical sample size	The school listing activity carried out prior to data collection identified the sample sizes in each school. The ET planned data collection accordingly, ensuring to the extent possible that there were enough children in Grades 1 and 2 in each school sampled. Nevertheless, a small number of schools were replaced due to insufficient sample sizes.
Consent	Three school directors refused to participate in the study.	During the school-listing phase, the survey team and ET worked together to reach out to each school to obtain their consent. The MENFP also provided a letter of authorization to facilitate participation in data collection. All unwilling schools were excluded from the sample frame. When schools refused during data collection, they were replaced.
Unbalance of baseline conditions	The schools of the four implementing partners have varying versions of curriculum and other related materials, potentially biasing their baseline data.	SI rests on the assumption that at the time of midline data collection, all schools will have received the new, Grade 1 and Grade 2 curriculum and will have been using it since after baseline data collection occurred. Under this assumption, the difference-in-difference statistical approach can be used to control for differing starting conditions, such as variety in curriculum.
Contamination/biases	Isolating the effects of different interventions by different partners and intervention starting times could lead to contamination and biases in program effect. One of the biggest challenges is that different partners are implementing different interventions under the umbrella of the Enhanced program. The number of schools under these interventions was not always large enough to identify effects with external validity.	The ET proposed to compare the average impact of Core and Enhanced models and estimate marginal effects separately for each separate subcomponent of the Enhanced model.
Bias in classroom teaching	Classroom observation is inherently biased as it is expected that most teachers, upon having an observer in the back of the classroom, will teach differently than they typically would.	Triangulation of classroom observation with student questionnaires and FGDs helped to clarify and isolate bias at baseline and will be further employed in future data collection rounds.
Weather and other natural events	Data collection took place in the heavy rainy season.	Heavy rains delayed data collection for several days. However, the one-week buffer time allotted for data collection ensured that it was completed by December 8.

FINDINGS

Findings from baseline data collection are presented in this section. We first present the realized sample and baseline balance statistics on covariates and outcomes, which are expected to either mediate

program impact or predict selection into core and enhanced treatment groups. An imbalance on these dimensions may introduce bias during endline analysis. Where imbalance is observed, we discuss strategies for minimizing potential bias during the endline. The next section examines the study sample and provides descriptions of core, enhanced, and comparison schools (and the students, teachers, and materials/resources within them).

The final findings section presents baseline data for each Evaluation Question specific to this study. This includes the baseline status of primary outcomes of interest (reading, teacher quality, and MENFP capacity) at all applicable levels (student, school, department). The purpose of this analysis is to assess baseline conditions in areas where IPs will work. This analysis will also set the stage for policy- and program- relevant conclusions and recommendations for further influencing reading performance among learners in the sample frame and beyond. Throughout this findings section, key findings and conclusions are **BOLDED**.

SAMPLE REALIZED

The ET surveyed 65 core, 65 enhanced, and 65 comparison schools resulting in a total of 195 schools, as planned. As detailed in Table 6, 38 percent of surveyed schools are public, and 35 percent are private. The remaining surveyed schools are religious, community, and “others” (such as NGOs and semi-private). Schools are also disaggregated by department: Ouest had the majority of schools with 60 (31 percent of the sample), followed by Nord-Est with 36 schools (18 percent), and Sud with 30 (15 percent). Artibonite and Sud-Est had the fewest schools at 11 each, or 6 percent of the sample each. Overall, sampled schools were 58 percent rural and 42 percent urban, although the rural-urban distribution varied widely across departments. Nord-Est and Sud-Est were the most rural (83 and 82 percent, respectively), while Artibonite was the most urban (91 percent) (Table 6).

Table 6: School sample realized by school type, rural/urban, and department

School Type	Artibonite	Grand' Anse	Nord	Nord-est	Ouest	Sud	Sud-est	Total	Pct.
Public	2	13	8	23	10	10	9	75	38%
Private	5	5	11	8	35	3	1	68	35%
Religious	3	6	1	3	6	15	1	35	18%
Community	1	2	0	2	5	1	0	11	6%
Other	0	0	1	0	4	1	0	6	3%
Total	11	26	21	36	60	30	11	195	100%
Pct. Total	6%	13%	11%	18%	31%	15%	6%	--	--
Pct. Urban	91%	46%	35%	17%	55%	37%	18%	--	42%
Pct. Rural	9%	54%	65%	83%	45%	63%	82%	--	58%

Note: “Other” denotes semi-private and NGO run schools.

Though fewer Grade 1 and 2 boys and girls were tested than originally planned, this was largely due to insufficient sample sizes of students at some schools and, to a lesser extent, non-compliance or non-consent. However, due to planned over-sampling, the reduced sample sizes did not affect statistical robustness of results. The reductions in the overall sample were proportional across treatment and comparison schools, with slightly fewer Grade 2 boys in the core model and fewer Grade 2 girls in the enhanced model. Across Grades 1 and 2, baseline testing included 1,776 boys and 1,718 girls. Table 7 presents the realized baseline sample of students by treatment status and grade level.

Table 7: Final EGRA sample realized, by grade and gender

Grade	Control Girls	Control Boys	Control Total	Core Girls	Core Boys	Core Total	Enhanced Girls	Enhanced Boys	Enhanced Total	Grand Total Girls	Grand Total Boys	Grand Total
Grade 1	279	302	581	286	297	583	293	299	592	858	898	1756
Grade 2	281	299	580	279	297	576	300	282	582	860	878	1738
Total	560	601	1161	565	594	1159	593	581	1174	1718	1776	3494

The ET also interviewed 194 school directors across core, enhanced, and comparison schools; one school director in the Nord department refused to participate in the survey. In addition to these interviews, the ET also aimed to conduct two teacher observations per school, successfully completing a final number of 381 of 390 planned observations. In nine schools, logistical and time considerations or the lack of a second eligible classroom to observe (where Grade 1 and 2 classrooms were mixed) resulted in only one valid observation at the school.

Lastly, the ET conducted focus group discussions with parents (both mothers and fathers, but separately) of children that attend sampled schools, as well as the children (boys and girls separately) themselves. In total, 28 FGDs were carried out with mothers, fathers, girls, and boys, stratified equally across departments. Key informant interviews were also held with 14 teachers and 14 MENFP personnel across the seven departments, for a total sample size of 309 participants, 158 (51 percent) of which were female (Table 8).

Table 8: Final sample realized for qualitative data collection

Qualitative Activity	Number of events completed	Females interviewed	Males interviewed	Final sample size
Boys FGD	7	--	66	66
Girls FGD	7	68	--	68
Fathers FGD	7	--	70	70
Mothers FGD	7	77	--	77
Teacher KII	14	11	3	14
MENFP KII	14	2	12	14
Total	56	158	151	309

ASSESSMENT OF EVALUATION DESIGN

Following data collection in November-December 2017, during which 17 of the original 195 selected schools were replaced,¹⁶ the ET conducted a second round of CEM with the fully collected data to check the robustness of the matching.¹⁷ Following this second round of matching, the final set of treatment and control schools was reduced to a total of 182 schools consisting of 61 control schools, 60 core treatment schools and 61 enhanced treatment schools. Balance statistics displayed in this section

¹⁶ Schools were replaced due to non-consent by a school director or insufficient sample sizes in Grades 1 and 2.

¹⁷ The list of covariates used for the second round of matching is provided in Annex C.

only (as well as all data in subsequent reports) correspond to the matched schools. Descriptive statistics and findings displayed in the rest of the report correspond to the full set of 195 unmatched schools.

When looking at covariates, both treatment groups are sufficiently balanced with the control group. To confirm the validity of the matching, the team closely examined the extent of balance of the covariates used in CEM matching. We used Stata command, *iebal*¹⁸. Table 9 shows that both treatment groups are sufficiently balanced. The respective F-statistic shows that the null hypothesis, that covariates are jointly significant, is rejected strongly,¹⁹ although a few variables, such as proportion of public schools and language of instruction, are individually significant at the 5 percent level. To further confirm balance across these three groups, we present school level indicators in the descriptive statistics section below (also found in Annex C) that shows no significant difference across treatment and control groups.

Table 9: Covariate balance statistics after CEM matching

Variable	Mean/[SE] Control	Mean/[SE] Core	Mean/[SE] Enhanced	p-value Core	p-value Enhanced
--	(1)	(2)	(3)	(1)-(2)	(1)-(3)
Public school	0.295	0.283	0.607	0.027**	0.872
What is the normal exit age for students at this school?	14.836	13.83	13.180	0.501	0.232
How many years' experience do you have as the director of this school?	10.721	12.967	10.082	0.065	0.965
Are textbooks usually shared?	0.000	0.000	0.033	N/A	0.313
Does every classroom have a chalkboard?	0.984	1.000	0.951	0.407	0.599
What is the main language of instruction?	0.770	0.900	0.787	0.027**	0.662
Have there been any dropouts since the beginning of the school year?	0.508	0.386	0.623	0.079	0.165
What is the proportion of part-time teachers?	0.459	0.367	0.393	0.473	0.852
Rural urban location (rural=1)	0.541	0.517	0.672	0.713	0.554
Does your school provide any health services to students (e.g., vitamin A supplements)?	0.328	0.300	0.328	0.778	0.234
Proportion of trained male teachers more than 25 percent?	0.902	0.933	0.852	0.544	0.971
Proportion of trained female teachers more than 25 percent?	0.885	0.800	0.770	0.103	0.206
What is student-teacher ratio?	20.747	23.785	26.411	0.397	0.649
N (Number of schools after balancing)	61	60	61	--	--

¹⁸ 'iebal' is a Stata command that generates balance tables (difference-in-means tables). The command tests for statistically significant difference between the categories defined in the covariates. The command can either test one control group against all other groups or test all groups against each other. The command also allows for fixed effects, covariates, and different types of variance estimators.

¹⁹ F-test was used to test the hypothesis that covariates in treatment and control samples are jointly significant. Based on the high level of significance, we can reject the hypothesis and safely say that covariates across treatment and control samples are jointly insignificant and the samples are balanced.

Variable	Mean/[SE] Control	Mean/[SE] Core	Mean/[SE] Enhanced	p-value Core	p-value Enhanced
F-test of joint significance (p-value)	--	--	--	0.000***	0.000***
F-test, number of observations	--	--	--	121	122

*Note: Values displayed for t-tests are p-values for two separate balancing calculations. The values displayed for F-tests are p-values. Standard errors are clustered at variable department level. The covariate variable cem_strata10 is included in all estimation regressions. Missing values in balance variables are treated as zero. *** and ** indicate significance at the 1 and 5 percent critical level.*

BALANCE IN LEARNING OUTCOMES: TREATMENTS VS CONTROL

To test balance in learning outcomes, we consider correct letter per minute (CLPM), correct letter sound per minute (CLSPM), correct word per minute (CWPM), and oral reading fluency (ORF) among timed subtasks. From non-timed subtasks, we consider listening and reading comprehension, Creole letter and word dictation, and initial sound recognition. Among those subtasks, reading fluency and comprehension are higher-order skills in the reading acquisition process, and they build upon lower-order, other foundational skills included in EGRA, such as phonological awareness, alphabet knowledge, decoding, and vocabulary.²⁰

Students' performance by grade shows mixed results in balance statistics. There is no significant difference between Grade 1 students from core treatment and comparison groups in six of nine subtasks. Both higher order subtasks, oral reading fluency, and reading comprehension, are similar across groups for Grade 1. Among the lower order subtasks, the only significant differences we observed are in CLPM, Creole letter dictation, and Creole listening comprehension. Between the enhanced and comparison groups, we found significant differences in four of nine subtasks including CLPM, CLSPM, CWPM, and Creole letter dictation, where the enhanced group scored higher than the comparison group. Similarly, in the case of enhanced versus comparison, oral reading fluency and reading comprehension are similar across groups for Grade 1.

For Grade 2 students, however, balance statistics show that students from the core treatment group significantly outperform their counterparts from the comparison group in almost all subtasks, including higher order ones. The difference between enhanced and comparison groups are, however, negligible (Table 10). Grade 2 students from the enhanced group performed significantly better than the comparison group only in CLPM, and vice versa in Creole listening comprehension. The difference in the two higher order subtasks, oral reading fluency and reading comprehension, was not significant.

²⁰ For an explanation of how these tests were conducted and how they measure the different stages of learning to read, see Annex C.

Table 10: EGRA scores (for all), with p-values

Grade	EGRA subtasks	Core Treatment (T1)	Enhanced Treatment (T2)	Control Treatment (C)	p-value [T1-C]	p-value [T2-C]
Grade 1	CLPM	10.4	8.2	5.8	0.00	0.00
Grade 1	CLSPM	6.4	5.9	4.4	0.35	0.00
Grade 1	CWPM	2.8	2.6	1.7	0.76	0.01
Grade 1	ORF	2.9	3.5	3.8	0.34	0.72
Grade 1	Creole letter dictation score (out of 5)	3.1	2.8	2.5	0.00	0.01
Grade 1	Creole word dictation score (out of 3)	0.1	0.1	0.0	0.15	0.04
Grade 1	Initial sound recognition score (out of 10)	1.4	1.3	1.1	0.57	0.33
Grade 1	Listening comprehension score (out of 5)	3.3	3.1	3.1	0.02	0.72
Grade 1	Reading comprehension score (out of 5)	0.1	0.1	0.1	0.34	0.58
Grade 2	CLPM	25.8	20.4	18.2	0.00	0.04
Grade 2	CLSPM	14.1	10.1	10.2	0.00	0.90
Grade 2	CWPM	13.3	9.1	7.6	0.00	0.05
Grade 2	ORF	17.0	12.0	11.6	0.00	0.76
Grade 2	Creole letter dictation score (out of 5)	4.2	3.9	3.8	0.00	0.19
Grade 2	Creole word dictation score (out of 3)	0.7	0.4	0.4	0.00	0.66
Grade 2	Initial sound recognition score (out of 10)	3.0	1.8	1.6	0.00	0.22
Grade 2	Listening comprehension score (out of 5)	3.72	3.73	3.58	0.056	0.04
Grade 2	Reading comprehension score (out of 5)	0.8	0.5	0.5	0.00	0.84

Note: A p-value in red font indicates a statistically significant difference at the $\alpha > 0.05$ level.

While different initial values may pose a threat of different trends in learning outcomes, generally the threat can be removed by using a difference-in-difference (DID) approach at mid and endline. Irrespective of the case, it is worth exploring the reasons behind such baseline differences, particularly when the schools are sufficiently balanced.

The EGRA literature suggests that in many developing countries the high proportion of zero scores in subtasks is responsible for significant variation in average overall scores. However, the ET also notes that quartile distribution of scores vary as well. To probe further into the differences across groups in this study, the ET adopted the following approach. First, the team identified the proportional distribution of all students (Grades 1 and 2 together) across each decile of learning outcome scores, including the proportion of zero scores as a separate category. Next, the ET separately analyzed the distribution of zero and non-zero scores for each learning outcome across the three groups. Figure 2 shows that the differences are mainly found at the proportion of zero scores. It shows that a visibly lower proportion of students scored zero in core treatment schools as compared to the students from enhanced and control schools. This incidence is more pronounced in CLPM, where findings showed that approximately 15 percent more students in the enhanced schools scored zero as compared to the

students from core schools. However, differences in other categories (such as 20-to-30 percent and 30-to-40 percent in CLPM, or 10 to 20 percent in CLSPM) are also observed in the figure. Because of this observation, in Table II, and box and whisker plot (Figure 3), the analysis compares the proportion of non-zero scores and then finally compares quartile distribution of non-zero scores across groups.

Figure 3: Decile distribution of timed learning test scores (percent of students in each category)

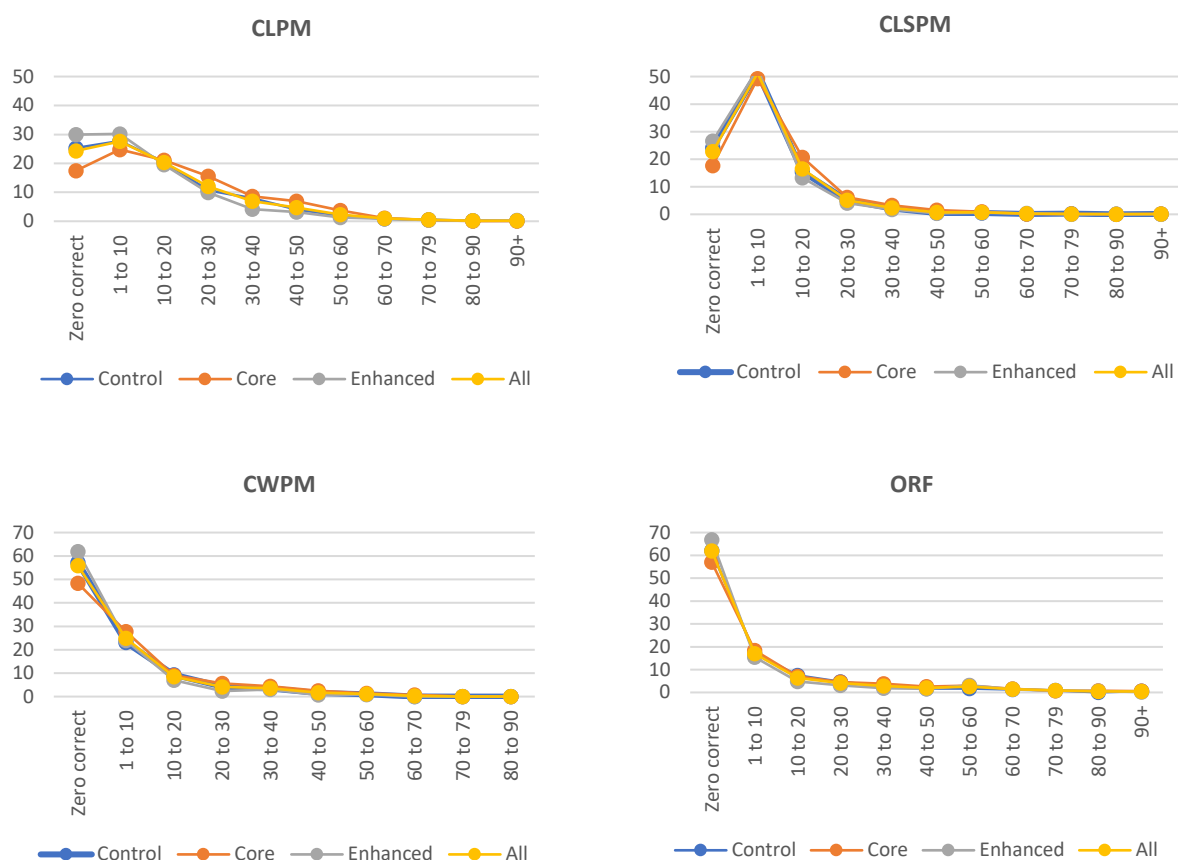


Table II compares treatment and comparison schools with respect to proportion of students who scored zeros in each subtask. The distribution clearly shows that significantly lower proportions of Grade 2 students in core treatment schools received zero scores in all subtasks compared to those in enhanced and comparison schools. The ET also observes that the subtasks in Table 10 with significant differences in average scores among Grade 1 students are the same subtasks in Table II where the proportion of zeros are significantly different across treatment and comparison schools. This high correlation between the imbalance in grade-specific subtask scores by treatment groups (in Table 10) and corresponding proportion of zero scores (in Table II) leads the ET to believe that difference in proportion of zero scores plays a key role in the observed imbalance in grade-specific subtask scores, even after the school-level covariates are sufficiently balanced among matched schools.

Table 11: Proportion of zeros in subtasks across treatment groups

Grade	Timed and Untimed Subtasks	Core Treatment (T1)	Enhanced Treatment (T2)	Control Treatment (C)	p-value [T1]	p-value [T2]
Grade 1	CLPM	29%	47%	41%	0.000	0.032
Grade 1	CLSPM	29%	41%	36%	0.029	0.060
Grade 1	CWPM	67%	76%	73%	0.054	0.226
Grade 1	ORF	74%	80%	73%	0.870	0.017
Grade 1	Creole letter dictation score (out of 5)	6%	15%	13%	0.000	0.353
Grade 1	Creole word dictation score (out of 3)	92%	97%	93%	0.261	0.018
Grade 1	Initial sound recognition score (out of 10)	75%	76%	75%	0.916	0.599
Grade 1	Listening comprehension score (out of 5)	4%	5%	7%	0.013	0.205
Grade 1	Reading comprehension score (out of 5)	91%	92%	93%	0.188	0.576
Grade 2	CLPM	5%	12%	10%	0.00	0.21
Grade 2	CLSPM	5%	11%	12%	0.00	0.64
Grade 2	CWPM	29%	47%	41%	0.00	0.047
Grade 2	ORF	39%	53%	49%	0.00	0.26
Grade 2	Creole letter dictation score (out of 5)	1%	1%	3%	0.01	0.22
Grade 2	Creole word dictation score (out of 3)	61%	77%	75%	0.00	0.39
Grade 2	Initial sound recognition score (out of 10)	54%	72%	70%	0.00	0.38
Grade 2	Listening comprehension score (out of 5)	0%	1%	2%	0.01	0.17
Grade 2	Reading comprehension score (out of 5)	58%	74%	72%	0.00	0.48

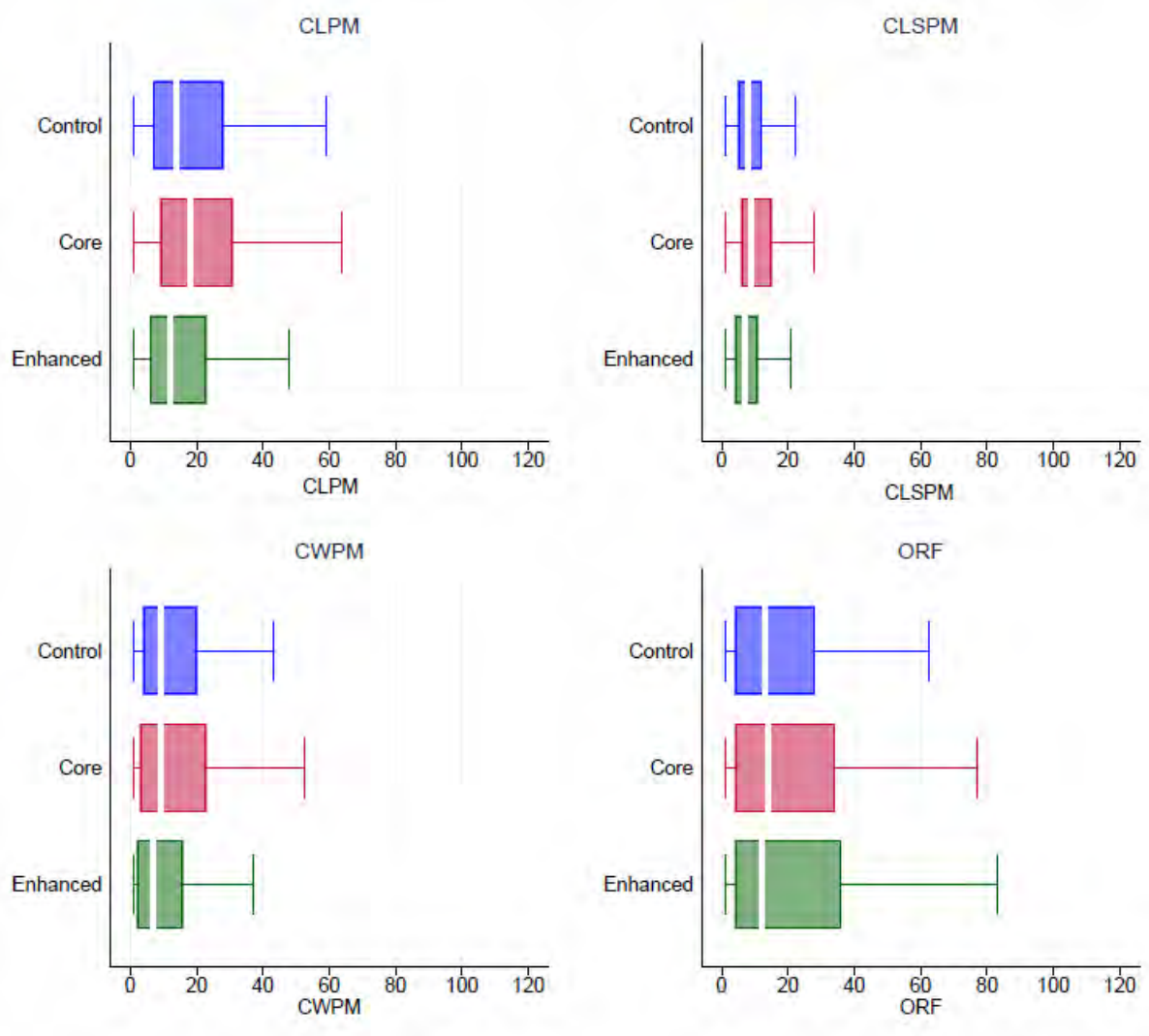
To examine further differences in non-zero distribution, the ET used box and whisker plots and analyzed the quartile distribution of non-zero scores across treatment groups.

Figure 3 shows that the distribution of core and comparison groups is similar for CLPM, CLSPM, and CWPM in the bottom 25 percent and at the median (the white line in Figure 3). The spread of the top 25 percent in the comparison group is, however, wider than the core group in the CWPM and ORF subtasks. The corresponding median value is also higher for the comparison group. The enhanced group has a narrower spread in all three subtasks except for ORF. In ORF, the spread of the top 25 percent of the comparison group is much wider, with a higher maximum score than the core group. For Grade 2, the comparison and enhanced groups have similar distributions, while the core group has a higher median as well as higher maximum values in all four subtasks.

To summarize, Grade 2 students in core treatment schools outperformed the other two groups of Grade 2 students at each quartile. However, the main differences come from lowest proportion of zero scores and the top 25 percent of scores. The only case where the

top 25 percent of students (measured by the length of the right whisker) from enhanced schools outperformed their counterparts from core schools was in ORF for Grade 1.

Figure 4: Quartile distribution of non-zero scores in timed outcomes (for all students)



REVISED POWER ANALYSIS

Because of the absence of actual grade level values, the overall mean minimum detectable effect sizes (MDES) for all subtests for Grades 1 and 2 was considered 0.25 SD in the design report for sample size calculation based on global trends, with overall ICC assumed to be 0.20. Using actual baseline statistics (EGRA subtask scores by grade) and intra-cluster correlations (ICC), the ET finds that the respective MDES for several subtasks are higher than the value the ET considered. This is primarily because the mean scores across learning outcomes are very low, with high levels of variance. ICCs across grades in the actual sample are also found to be higher in some cases than the assumed overall value of 0.20. Due to these differences between the assumption the team made during design phase and the findings from actual data, the ET re-calculated the MDES for each subtask so that the evaluation is sufficiently powered to maintain external validity. Adopting standard assumptions of 0.8 for statistical power (β) and

0.05 as the standard significance level (α), the analysis presents updated MDES for each of the EGRA subtests in Table 12 below:

Table 12: Updated MDES, by EGRA sub-tasks and Grade

Grade	EGRA subtasks	Mean values	Standard deviation	Intra cluster correlation (ICC)	Design effect	Detectable difference	Expected value in baseline	MDES
Grade 1	CLPM	8.25	11.76	0.24	2.89	3.44	11.69	0.29
Grade 1	CLSPM	5.61	8.44	0.19	2.5	2.29	7.9	0.27
Grade 1	CWPM	2.39	6.48	0.16	2.24	1.67	4.06	0.26
Grade 1	ORF	3.32	10.99	0.15	2.22	2.81	6.13	0.26
Grade 1	Initial sound recognition score (out of 10)	1.28	2.76	0.16	2.28	0.72	2	0.26
Grade 1	Reading comprehension score (out of 5)	0.12	0.46	0.13	2.04	0.11	0.23	0.24
Grade 1	Listening comprehension score (out of 5)	3.19	1.32	0.06	1.49	0.28	3.47	0.21
Grade 1	Creole letter dictation score (out of 5)	2.78	1.58	0.30	3.39	0.50	3.28	0.32
Grade 1	Creole word dictation score (out of 3)	0.08	0.38	0.11	1.86	0.09	0.17	0.24
Grade 2	CLPM	21.7	17.75	0.27	3.14	5.4	27.1	0.30
Grade 2	CLSPM	11.6	11.46	0.17	2.32	3	14.6	0.26
Grade 2	CWPM	9.99	14.54	0.27	3.19	4.47	14.46	0.31
Grade 2	ORF	13.5	21.18	0.25	2.97	6.27	19.77	0.30
Grade 2	Initial sound recognition score (out of 10)	2.23	3.53	0.13	2.03	0.87	3.1	0.25
Grade 2	Reading comprehension score (out of 5)	0.6	1.06	0.22	2.76	0.30	0.9	0.28
Grade 2	Listening comprehension score (out of 5)	3.69	1.13	0.06	1.48	0.24	3.93	0.21
Grade 2	Creole letter dictation score (out of 5)	3.96	1.21	0.21	2.66	0.34	4.3	0.28
Grade 2	Creole word dictation score (out of 3)	0.48	0.85	0.14	2.09	0.21	0.69	0.25

Note: The estimation of MDES is based on significance level, α (.05), power, β (0.80), and class size, n (9) and number of schools per arm, k (60). The analysis used mixed model with fixed and random variation by school and grade, and `clustersamps` in Stata to find out the value of MDES for each indicator. MDES is estimated by dividing detectable difference by standard deviation.

DESCRIPTIVE STATISTICS OF SAMPLED SCHOOLS AND STUDENTS

SCHOOL LEVEL

Public and private school types were the most prominently represented in the sample, as noted in the realized sample section above. Although most schools in the country are private schools (80-90 percent), the ERP schools and schools chosen to represent the comparison group were not necessarily representative of that reality. ERP schools were chosen by IPs based on a set of pre-selected criteria (found in Annex C), which do not necessarily take into account the schools' funding sources. When constructing a valid comparison group, the ET tried, to the extent possible, to match characteristics of treatment (ERP) schools with control schools, including funding type (as detailed above). Nearly all sampled schools were mixed (co-ed), particularly in core and comparison samples, with 98 percent and

100 percent of schools, respectively, reporting co-ed student populations at the baseline. The proportion is 93 percent for the enhanced schools (Table 13).

Enhanced schools have larger class sizes relative to core schools. The average number of students in an average Grade 2 school classroom at baseline was 38 (for both core and comparison schools) and 54 in an enhanced school classroom. Furthermore, there is on average one teacher for 24 students in core schools and one teacher for 26 students in enhanced schools. Average entry and exit age of students to and from primary school was consistent across treatment and comparison samples; students generally enter Grade 1 when they are four or five years of age and exit the school when they are between ages 13 or 14 (Table 13).

The ET finds that a greater proportion of core schools had all teaching positions filled by the start of the school year relative to enhanced schools (85 percent and 61 percent, respectively). Three out of every four comparison schools reported that their teaching positions were filled (Table 13).

Table 13: Basic school-level information across schools

Indicators	Core	Enhanced	Control
Number of boys and girls in Grade 1	37.9	60.1	36.7
Number of boys and girls in Grade 2	37.8	53.7	33.8
Student to teacher ratio	23.8	26.4	20.7
Normal entry age for students at this school	4.9	5.21	4.62
Normal exit age for students at this school	13.8	13.2	14.8
Average number of hours taught per day	5.3	5.2	5.2
% of schools reporting dropout since the beginning of the school year.	39%	62%	51%
Number of girls dropped out since the beginning of the school year	6.4	9.5	5.7
Number of boys dropped out since the beginning of the school year	6.7	6.6	7.2
% Mixed (co-ed) school	98%	93%	100%
% of schools with all teaching positions filled by school year start date	85%	61%	75%

Note: Please see Annex D to check level of significance

Table 14 presents school-level management and programs/services. School management committees and parent-teacher associations in Haiti include teachers and parents. The committees and associations attempt to ensure public funds are utilized transparently and effectively. Most core schools (84 percent) reported having an active parent-teacher association, while 77 percent of enhanced schools reported having one. Approximately 83 percent of core schools' management assist students with their school fees, and in enhanced schools, 61 percent of management provide support. This fact is critical considering that the cost of tuition and school materials is one of the main barriers to school enrollment in the country.

Oversight from MENFP also helps to ensure appropriate use of funds and resources in schools. Most core schools (88 percent) had been monitored by MENFP in the last year at baseline, while just over two thirds of enhanced schools reported receiving a monitoring visit (Table 14).

Most core schools do not offer student programs or services such as a feeding program, reading clubs, or extracurricular activities. This is also true of enhanced schools, with a few exceptions. Just under 70 percent of enhanced schools offer a feeding program and over half distribute student kits, which include academic materials and resources. Regarding school infrastructure, almost all surveyed schools have a chalkboard. Most schools, regardless of ERP model or treatment status, however, do not have libraries or computers/computer rooms for students to use. More core schools have libraries relative to

enhanced schools (47 percent and 25 percent, respectively). Likewise, while just under a quarter of core schools reported having computers/computer rooms, only 7 percent of enhanced schools have them (Table 14).

Exactly one third of both core and enhanced schools reported benefitting from updated or renovated infrastructure. Additionally, just over one third of schools reported benefitting from a water, sanitation and hygiene (WASH) intervention (42 percent of core schools and 33 percent of enhanced schools). Approximately two thirds of core and enhanced schools had separate bathrooms for boys and girls. More core schools had hand-washing stations close to a girl's latrine relative to enhanced schools (64 percent of core schools and 52 percent of enhanced schools) (Table 14).

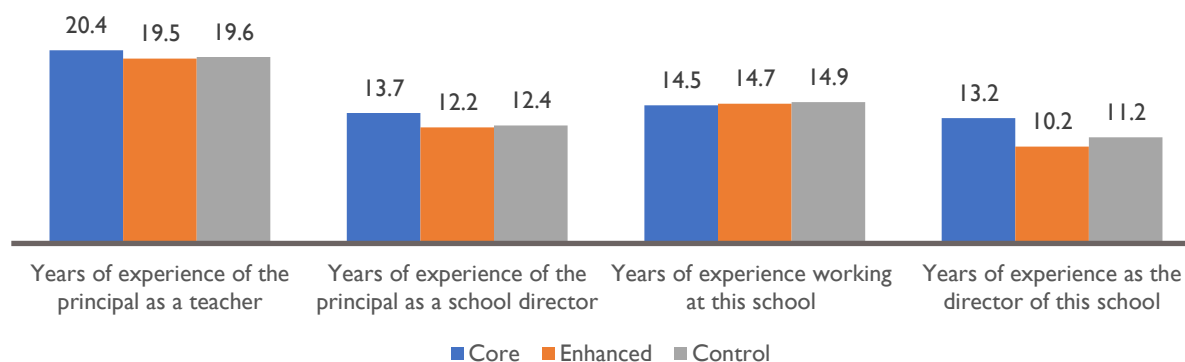
Table 14: School level resources and activities, by treatment and comparison

Indicators	Core	Enhanced	Control
% Schools that have an active parent-teacher association	84%	77%	72%
% School managements that support students with their school fees	83%	61%	76%
% Schools that have after-school extra-curricular activities	17%	20%	18%
% Schools that have a school feeding program	37%	69%	33%
% Schools that have a reading club	27%	15%	15%
% Schools that distribute student kits	33%	57%	26%
% Schools that provide any health services to students	30%	33%	33%
% Schools visited by a member of the MENFP in the past year to monitor facility	88%	70%	71%
% Schools that have computers or a computer room that students can use	22%	7%	13%
% Schools that have a library	47%	25%	31%
% Schools that have a chalkboard in every classroom	100%	95%	98%
% Schools that have had infrastructure renovations	33%	33%	33%
% Schools that have water, sanitation, and hygiene (WASH) intervention	42%	33%	39%
% Schools that have separate toilets for girls and boys	66%	70%	60%
% Schools that have a hand washing station near the girls' latrines	64%	52%	56%

Note: Please see Annex D to check level of significance.

School director characteristics were similar across the sample. Directors, in general, had just under two decades of teaching experience and had served in a director post for just over a decade. Interestingly, directors worked several years on average in their current school before taking on the position as director. Figure 4 shows director experience disaggregated by core, enhanced, and comparison schools.

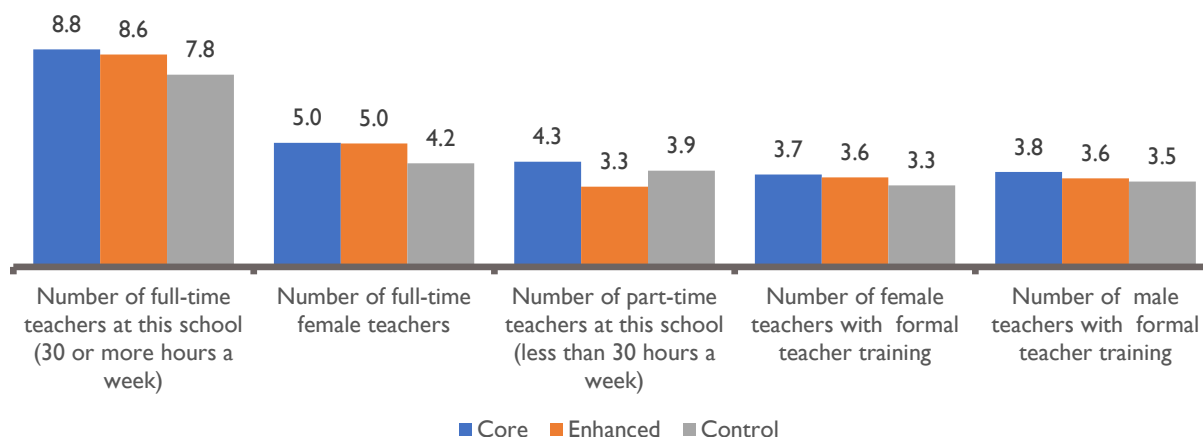
Figure 5: School Director experience



Sampled schools in both treatment and comparison groups had more full-time than part-time teachers. As shown in Figure 5, core and enhanced schools reported nine full-time teachers per school on

average. Just over half of these full-time teachers are female. In Haiti, teacher qualifications vary widely. On average, a similar number of male and female teachers across the schools have formal teacher training; however, teachers with formal teacher training represent just over a quarter of teachers in sampled schools. The data collection tools, in addition to documenting formal teacher training, also documented highest level of academic education completed. On average, sampled schools had more male teachers completing education above secondary school (such as a university degree or vocational training) than female teachers.

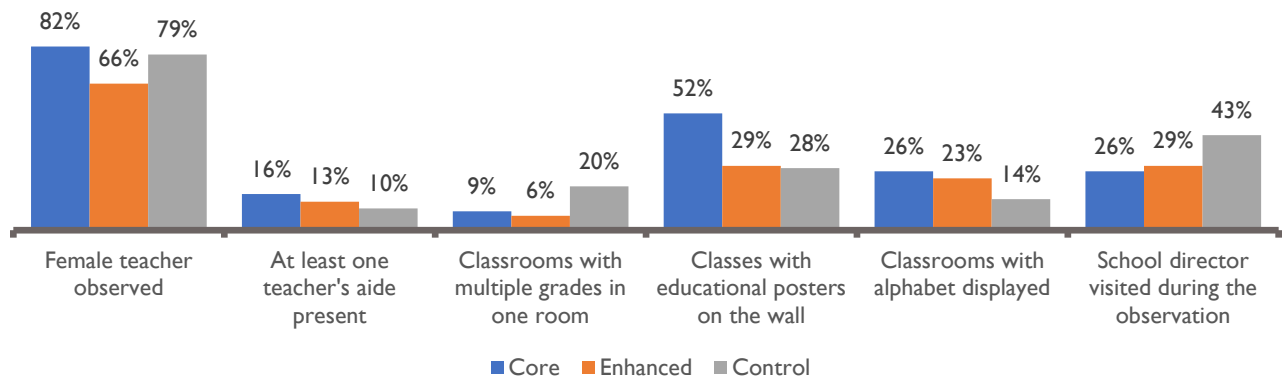
Figure 6: Teacher characteristics, per school



CLASSROOM OBSERVATIONS

The ET observed a total of 381 classrooms, or approximately two classrooms per each treatment and control school (one Grade 1 and one Grade 2). Just over half of the observations were conducted in Grade 1 classrooms. A greater percentage of core classrooms had a female teacher relative to enhanced classrooms (82 and 65 percent, respectively). Only a small number of classrooms had a teaching aide (16 percent in core and 12 percent in enhanced schools). A greater percentage of core classrooms had educational posters on the wall than did enhanced classrooms, and approximately a quarter of both core and enhanced classrooms had an alphabet displayed. Twenty-six percent of core classrooms that were observed received a visit from the school director while the observation was taking place, and 29 percent of enhanced classrooms received such a visit (Figure 6). A visit from a school director might make a teacher teach differently than normal.

Figure 7: Descriptive statistics for classroom-level variables

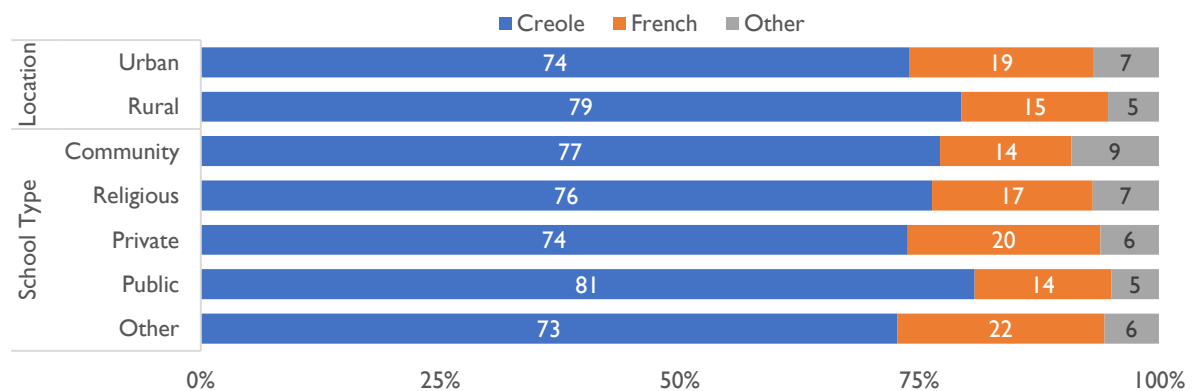


STUDENT LEVEL

In this section, the evaluation team reports descriptive statistics from the Student Questionnaire. In the body of this report, the team only discusses a few of these interesting findings. Additional analyses can be found in Annex D.

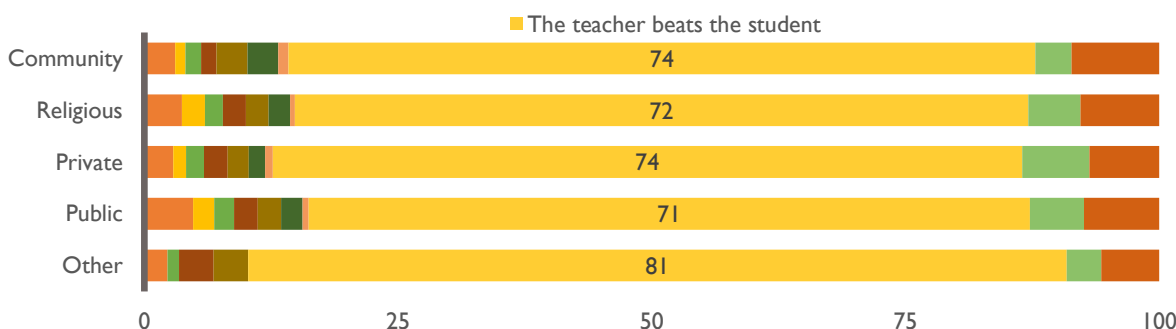
Most students speak Haitian Creole at home (over 70 percent in all school types). More private school students speak French at home relative to public school students (20 percent and 14 percent, respectively). In terms of rural-urban locational distribution, 26 percent of urban students (5 percent more than rural students) speak either French or other languages such as English or Spanish. (Figure 7).

Figure 8: Language spoken by students at home, by school type and location



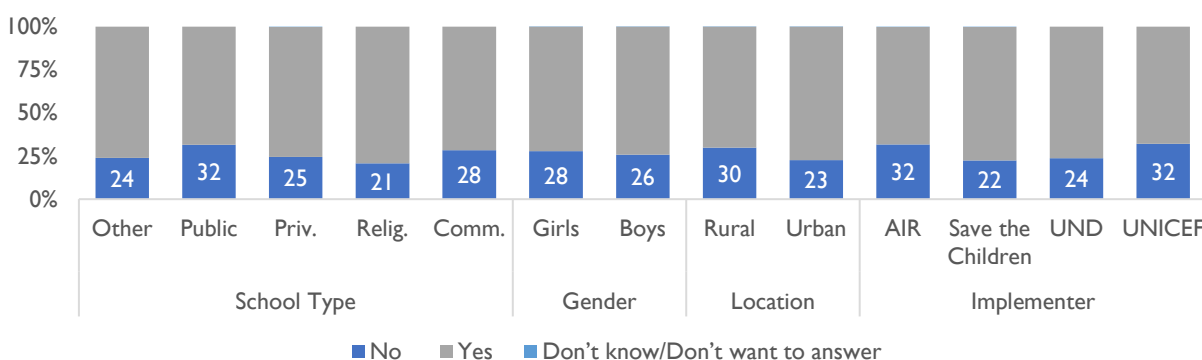
Students report pervasive corporal punishment across school types. When students are unable to provide an answer to the teacher (or gave an incorrect answer), a majority of students report that teachers use corporal punishment (more than 70 percent across school types) (Figure 8). Enumerators conducting classroom observations observed corporal punishment in only 11–13 percent of classrooms, as a comparison, although this figure is likely biased by having an observer in the classroom.

Figure 9: Extent of punitive measures (as reported by students), by school type (percentage)



Roughly a quarter of students in surveyed schools go to school without breakfast—with some variation across IPs, school type, and location (urban/rural). Slightly more urban school students reported eating breakfast before school relative to rural school students (77 percent and 70 percent, respectively). Approximately two thirds of students report eating while at school. Although this result did not vary by gender, there are differences across IP, school type, and location. For example, 71 percent of students in private schools reported eating while at school compared to just 57 percent in community schools (Figure 9).

Figure 10: Percent of students who eat in the morning before school



Absenteeism is somewhat high across all school types. When asked, “Have you ever been absent from school,” 42 percent of all students said yes. Among IPs, students in STC schools reported that they had been absent 46 percent of the time compared to 38 percent in UNICEF schools, 40 percent in UND schools, and 43 percent in AIR schools. Little variation was noted across school type and location, however. For those who reported being absent, 40 percent reported they were absent “once or twice a week,” and 33 percent said, “once or twice a month.” Although students at STC schools were most likely to have been absent, those in AIR schools were absent at more frequent intervals (50 percent reported being absent “once or twice a week.”) The main reason for absenteeism across all school types, genders, and IPs was “because I was sick” (50 percent), followed by “other” (12 percent), and “bad weather” (5 percent).

EQ1 AND EQ2: READING SCORES

Evaluation Questions 1 and 2 look at EGRA scores between treatment status (control and treatment) and ERP model (core versus enhanced). Sub-questions go into the difference in EGRA scores between gender (girl and boy), implementer (Save the Children, AIR, UND, and UNICEF), and type of school (public, private, community, religious, other). The questions also touch upon environmental factors at the household-, school-, and classroom- level that are associated with reading performance.

EGRA SCORES IN CONTEXT

ERP EGRA scores at baseline were similar to previous EGRA studies conducted in Haiti, particularly core school scores. Table 15 presents three EGRA subtask scores for several EGRA studies conducted in Haiti. CWPM represents an initial reading skill and the other two tests— oral reading fluency and reading comprehension— represent higher-level skills. Only Grade 2 student scores are included here. (Comments about how schools were selected for each study are found in the footnotes.) Each study was conducted in a different year and included schools selected based on different criteria, but a simple comparison of these selected scores for Grade 2 students can help provide context for the ERP scores. Specifically, Grade 2 students in Haiti score poorly on the reading comprehension subtask. Little variation exists across different Haiti EGRA studies regarding this subtask score, with students unable to answer one question correctly regarding a piece of text. Haitian Grade 2 students in all EGRA studies noted here have higher ORF scores relative to CWPM scores. The ERP enhanced treatment schools are similar to schools tested in the Concern project, representing the lowest subtask scores across the studies. Other studies have Grade 2 reading scores similar to the ERP core treatment schools.

Table 15: EGRA subtask scores from other Haiti EGRA studies

Country: Program	CWPM	ORF	Reading Comprehension
Haiti: ERP core	13.28	16.87	0.8
Haiti: ERP enhanced	7.79	12.02	0.5
Haiti: Concern ²¹	4.69	5	0.24
Haiti: ToTAL ²²	16.43	24.56	0.92
Haiti: Ann Ale/CLAMS ²³	13.61	--	0.86
Haiti: PHARE ²⁴	--	23	--

²¹ Means from baseline treatment group in 2012. The Baseline Report, Section 1.5 explains school selection criteria. *Early Grade Reading Assessment (EGRA)*. (2013). Concern Worldwide.

²² Means from endline control group in 2014. The Final Report, Section 2.2 explains school selection criteria. *Tout Timoun Ap Li – ToTAL (All Children Reading), Final Report, Revised*. (2012-2014). Education Data for Decision Making EdData II, USAID.

²³ Means are from 2016 data collection. Sampling of schools is explained in the recently released report. *Early Grade Reading and Writing Assessment in An n Ale Schools With the Coach-Led Learning Assessment Monitoring System (CLAMS)*, Final Report- French Version, May 2016, USAID

²⁴ Means are from 2009 data collection. Sampling of schools is explained in Programme Haitien d'Appui à la Réforme de l'Éducation (PHARE) & American Institutes for Research (AIR). (2009). *Draft: Student diagnostic assessment baseline report*. Prepared for USAID/Haiti and the Government of Haiti. Washington, DC: AIR.

EGRA SCORES, BY TREATMENT STATUS AND TREATMENT MODEL

Overall EGRA results by treatment status, model, and grade are presented earlier in the Balance Statistics section of this report (Table 10). To avoid duplication, those results are not repeated here. Discussion of the differences in reading scores between treatment status and ERP model will be brought in at midline and endline after the ERP has had time to show potential impacts. In this section, baseline EGRA results are presented by gender, implementer, school type, and other situational and environmental factors. We introduce regression models that examine factors that significantly affect students' oral reading fluency, followed by a decomposition of the regression results showing the overall share of those factors on oral reading fluency. Lastly, we synthesize qualitative data that help explain several factors found to influence reading scores at baseline. At endline, EGRA results will be compared among core, enhanced, and comparison schools to determine the impact (if any) of the core and enhanced ERP interventions. Evaluation Questions 1, 2, and all sub-questions will be answered completely at that time.

In the Balance Statistics section earlier in this report and in subsequent sections, EGRA scores are presented as averages, and disaggregated by treatment status, model, gender, implementer, and type of school. Figure 2 shows the proportional distribution of EGRA scores. However, to paint a clearer picture of fluency, Table 16 presents the actual proportion of the sample that scored in each ORF bracket (zero CWPM, 1 to 10 CWPM, 11-20 CWPM, etc.). Given that the Haitian government defines a child having oral reading fluency if he or she can read 40 CWPM or greater, the table shows the ultimate distribution, across grades and treatment models, of students who can, effectively, read. Table 16 shows that for Grade 1, 2.31 percent of all sampled children scored 40 CWPM or higher and were considered to have oral reading fluency. For Grade 2, this proportion increased to 13.1 percent. Table 17 and Table 18 show the same breakdown for letters (CLPM) and letter sounds (CLSPM).

Table 16: Distribution of ORF scores across grades and by CWPM categories

Grade 1 CWPM	Grade 1 Control	Grade 1 Core	Grade 1 Enhanced	Grade 1 Total	Grade 2 (%) Control	Grade 2 (%) Core	Grade 2 (%) Enhanced	Grade 2 (%) Total
None correct	75.05	74.59	79.96	76.55	49.64	39.22	53.62	47.51
1 to 10	14.86	18.05	12.27	15.04	18.8	18.96	18.92	18.89
10 to 20	4.77	3.68	2.53	3.65	10.04	10.59	7.05	9.23
20 to 30	2.02	0.92	1.08	1.34	7.3	8.18	5.01	6.83
30 to 40	1.47	1.29	0.54	1.1	3.83	6.32	3.15	4.43
40 to 50	0.92	0.37	0.36	0.55	2.55	4.65	2.78	3.32
50 to 60	0.18	0.37	1.44	0.67	3.1	5.58	4.82	4.49
60 to 70	0.37	0.37	0.54	0.43	2.19	2.79	2.41	2.46
70 to 79	0	0.18	0.72	0.3	1.64	1.3	1.11	1.35
80 to 90	0	0.18	0.36	0.18	0.18	1.3	0.93	0.8
90+	0.37	0	0.18	0.18	0.73	1.12	0.19	0.68
Total	100	100	100	100	100	100	100	100

Table 17: Distribution of CLPM scores across grades

Grade 1 (%) CLPM	Grade 1 (%) Control	Grade 1 (%) Core	Grade 1 (%) Enhance	Grade 1 (%) Total	Grade 2 (%) Control	Grade 2 (%) Core	Grade 2 (%) Enhance	Grade 2 (%) Total
None correct	41.56	29.78	47.22	39.59	9.78	5.19	12.52	9.18
1 to 10	30.13	31.8	34.11	32.02	25.18	17.81	26.7	23.26
10 to 20	15.25	22.06	12.39	16.53	25.54	19.67	26.7	23.99
20 to 30	6.53	10.11	4.49	7.02	14.67	20.78	15.29	16.89
30 to 40	4.72	3.49	0.9	3.03	11.23	14.29	7.37	10.95
40 to 50	0.91	1.65	0.36	0.97	7.07	12.24	5.89	8.38
50 to 60	0.36	0.92	0	0.42	3.62	6.49	2.58	4.22
60 to 70	0.36	0	0	0.12	1.27	2.41	1.66	1.77
70 to 79	0.18	0	0.18	0.12	0.91	0.93	0.74	0.86
80 to 90	0	0.18	0	0.06	0.36	0	0.37	0.24
90+	0	0	0.36	0.12	0.36	0.19	0.18	0.24
Total	100	100	100	100	100	100	100	100

Table 18: Distribution of CLSPM scores across grades

Grade 1 (%) CLSPM	Grade 1 (%) Control	Grade 1 (%) Core	Grade 1 (%) Enhance	Grade 1 (%) Total	Grade 2 (%) Control	Grade 2 (%) Core	Grade 2 (%) Enhance	Grade 2 (%) Total
None correct	36.5	29.96	41.55	36.04	12.34	5.57	11.97	9.98
1 to 10	47.81	52.02	48.56	49.45	55.17	45.83	56.72	52.6
10 to 20	10.77	14.89	6.29	10.62	19.96	27.09	20.26	22.41
20 to 30	2.74	1.65	1.62	2	7.62	10.76	6.45	8.27
30 to 40	0.91	0.55	1.8	1.09	3.27	5.94	1.66	3.61
40 to 50	0.18	0.18	0	0.12	0.73	2.97	0.74	1.47
50 to 60	0.55	0.37	0	0.3	0.36	1.48	1.29	1.04
60 to 70	0	0.18	0	0.06	0.18	0.19	0.74	0.37
70 to 79	0.36	0	0	0.12	0.18	0.19	0	0.12
80 to 90	0	0	0.18	0.06	0	0	0.18	0.06
90+	0.18	0.18	0	0.12	0.18	0	0	0.06
Total	100	100	100	100	100	100	100	100

EGRA SCORES, BY GENDER

Across all “timed” EGRA tests, Grade 1 and 2 girls scored higher relative to boys. For example, Grade 2 boys were able to identify 20 correct letters per minute compared to the 23.5 identified by girls (Figure 10). Notably, this figure includes students who scored zero as well. As noted above, many students in the study at baseline scored zero. In Figure 11, zero scores for timed tests are displayed by grade and gender for the entire sample. More boys in Grades 1 and 2 scored zero on the timed tests relative to girls. Over 70 percent of students (boys and girls) in Grade 1 did not read any

words correctly from the text in the timed CWPM and ORF test. In Grade 2, the number of girls and boys scoring zero decreased, but over 40 percent still could identify a word correctly on either test. The box and whisker plots displayed in

Figure 12 show distribution and level of non-zero scores for timed tests for both grades taken together. Girls have higher median scores relative to boys, and girls' scores vary more relative to boys' in all timed tests (girls achieved higher scores for all tests). The better performance of girls relative to boys in timed subtasks is, therefore, due not only to fewer girls scoring zero, but also to the fact that the top 25 percent of girls achieved much higher scores than boys.

Figure 11: Average scores (in words) in timed subtasks, by grade and gender

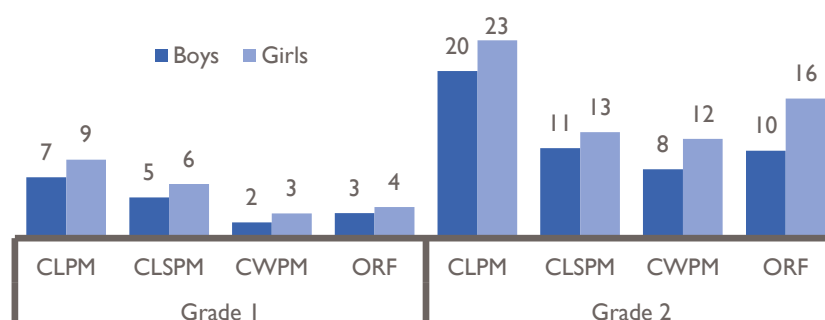


Figure 12: Proportion of zero scores in timed subtasks, by grade and gender

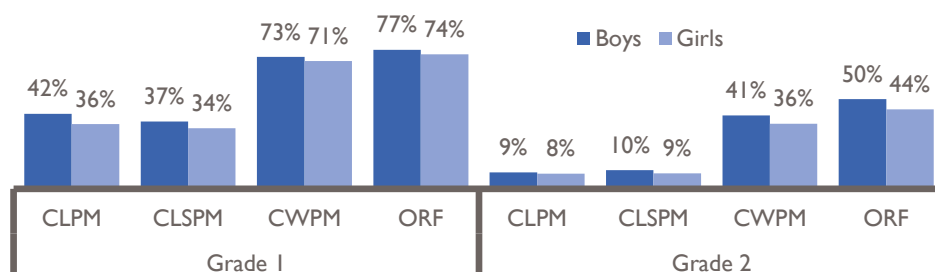
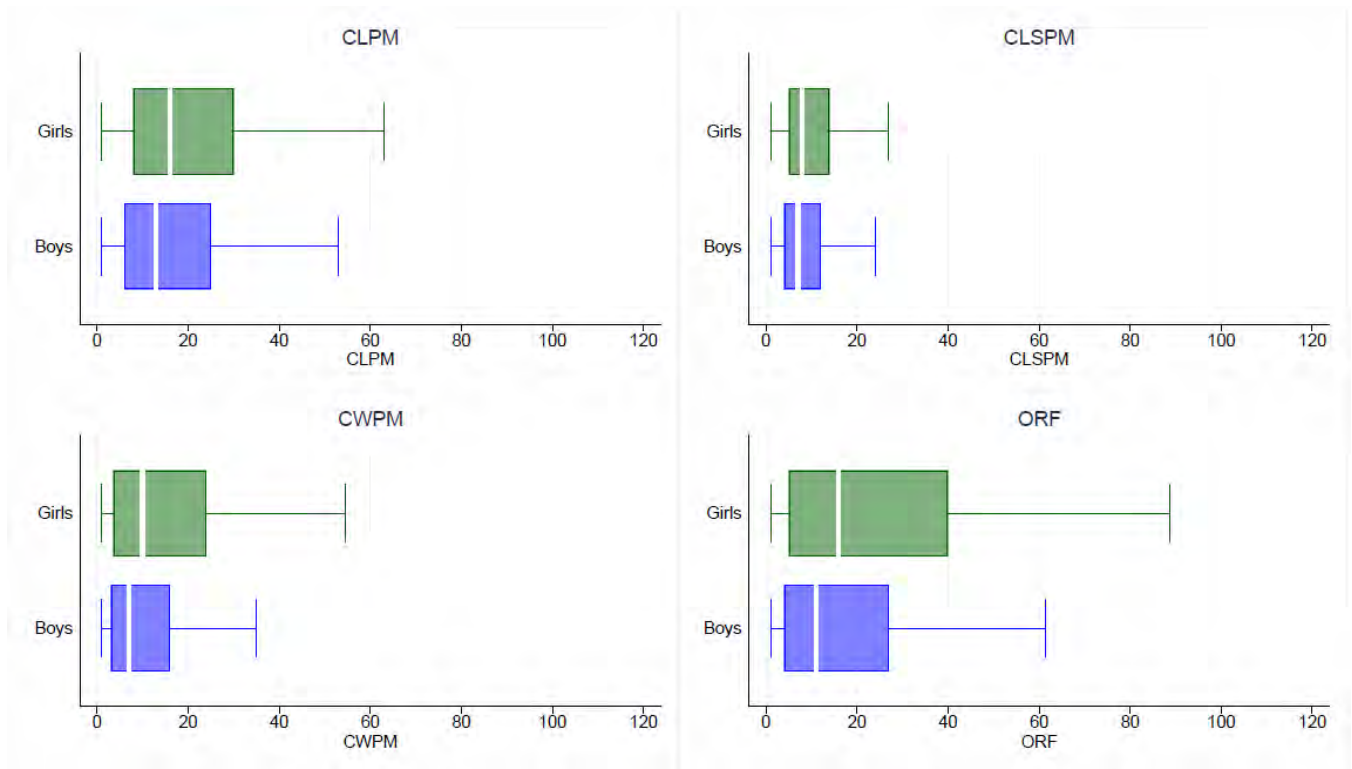


Figure 13: Timed learning outcomes non-zero scores, by gender



Boys' and girls' scores were more similar at baseline in non-timed tests than they were in timed tests. Figure 13 shows non-timed test scores by grade and gender. Boys and girls tested highest in listening comprehension and letter dictation and lowest in word dictation and reading comprehension. Over 90 percent of boys and girls in Grade 1 scored zero on reading comprehension and word dictation (Figure 13). This percentage improved to approximately two thirds of boys and girls in Grade 2. The proportion of zero scores (Figure 14) and quartile distribution of non-zero scores in non-timed subtasks (Figure 15) reveal that girls do relatively better in some subtasks, such as Creole word dictation or reading comprehension, because fewer girls received zero scores than boys. The distribution of boys and girls with non-zero scores in each quartile is almost identical for initial sound recognition, listening and reading comprehension, and Creole letter dictation.

Figure 14: Average scores in non-timed subtasks, by grade and gender

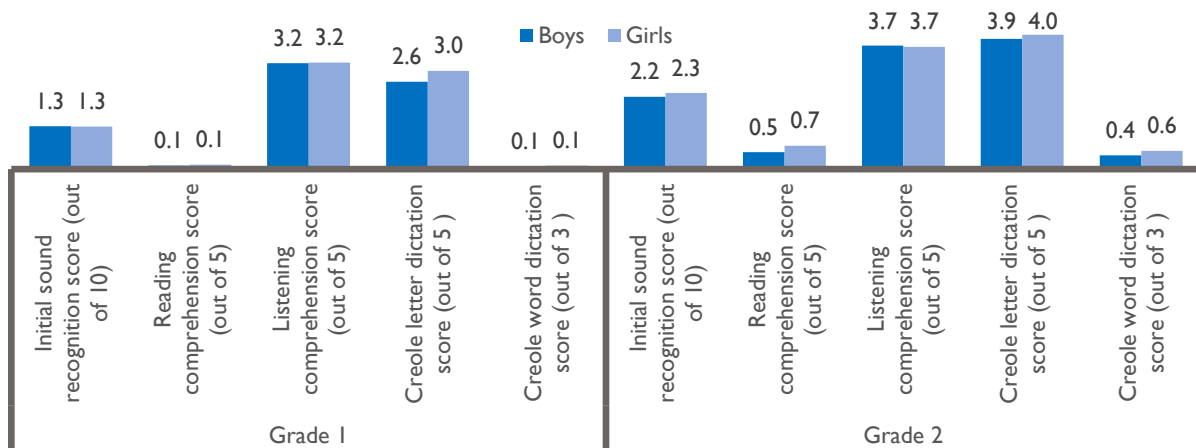


Figure 15: Proportion of zeros in non-timed subtasks, by grade and gender

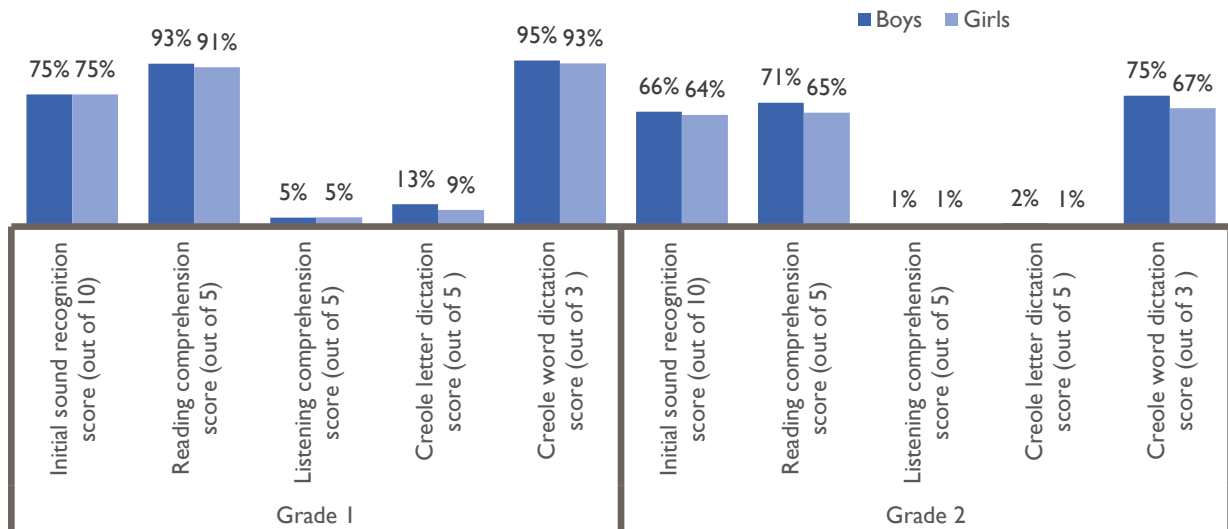
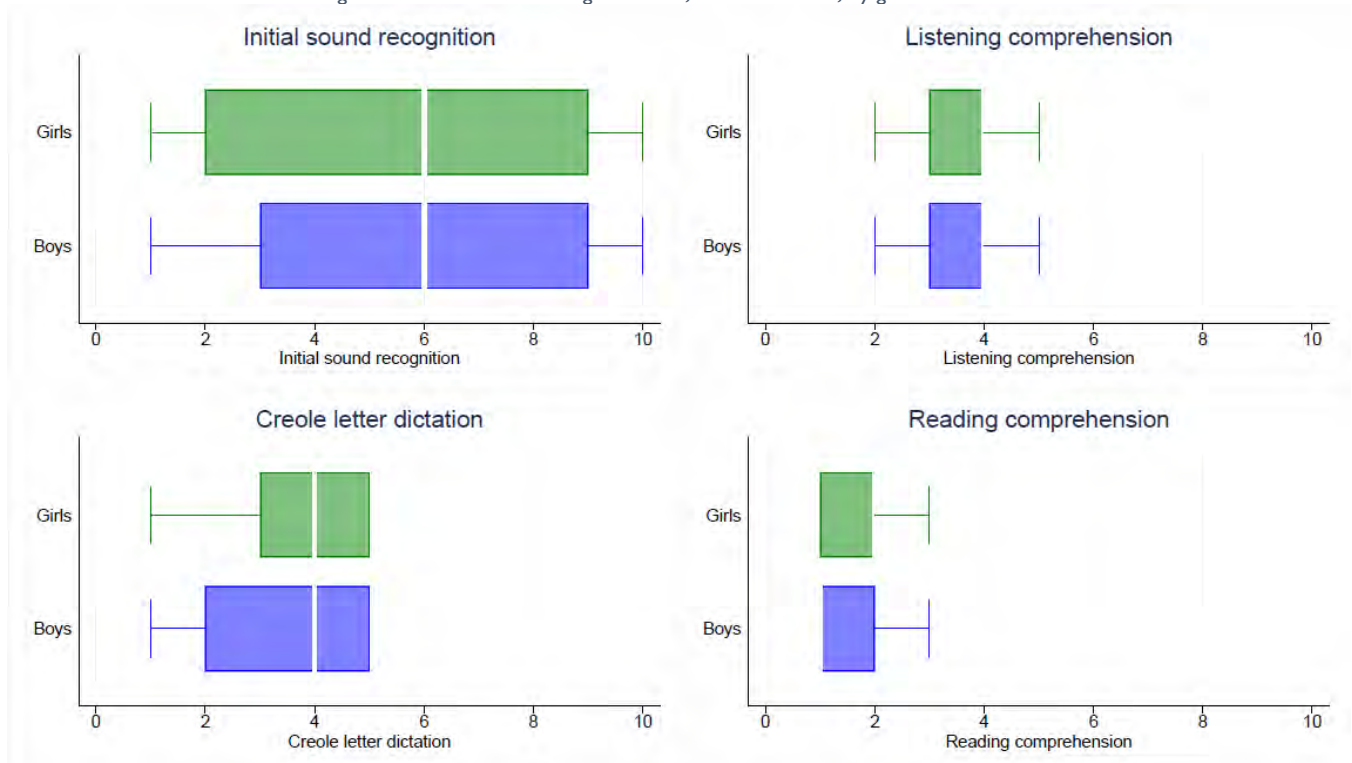


Figure 16: Non-timed learning outcomes, non-zero scores, by gender



EGRA SCORES, BY IMPLEMENTER

At baseline, there was variation in EGRA scores across implementers.²⁵ Figure 16 shows timed EGRA results by grade and IP demonstrating that scores were not similar across implementers at baseline. For example, Save the Children school students scored higher on the EGRA test, in general, across timed subtasks, particularly in Grade 2 (Figure 16). For non-timed test scores, the scores for implementers were much closer (Figure 17).

²⁵ It should be mentioned that the starting conditions of the schools under the implementing partners varied substantially in some areas. For example, at baseline, UNICEF schools had the older version of the curriculum, AIR and UND schools had not received any curriculum, and Save the Children schools were located exclusively in the urban Port-au-Prince regions. While these characteristics may play some role in their reading scores at baseline, the initial conditions will be accounted for using the difference-in-difference statistical approach at midline and endline to further help isolate the effect of the program.

Figure 17: Timed learning outcomes, by grade and implementer

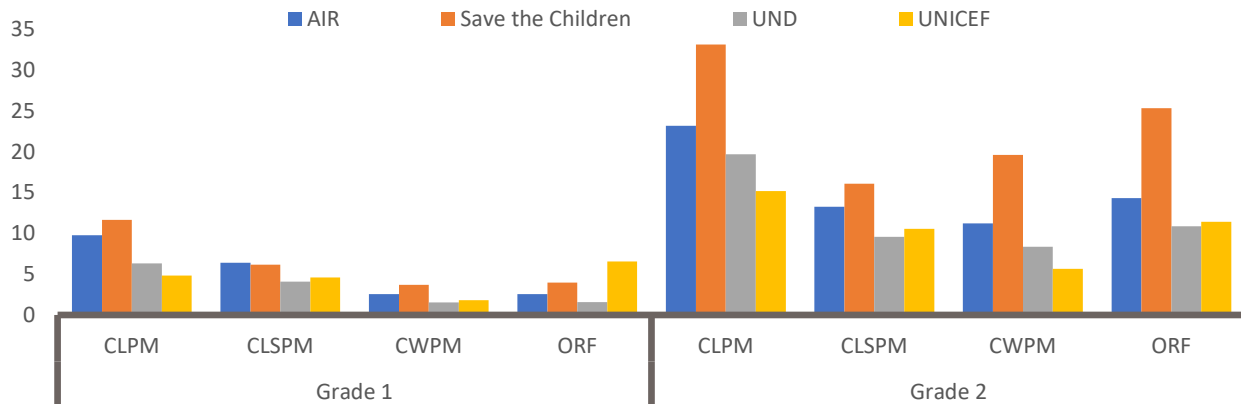
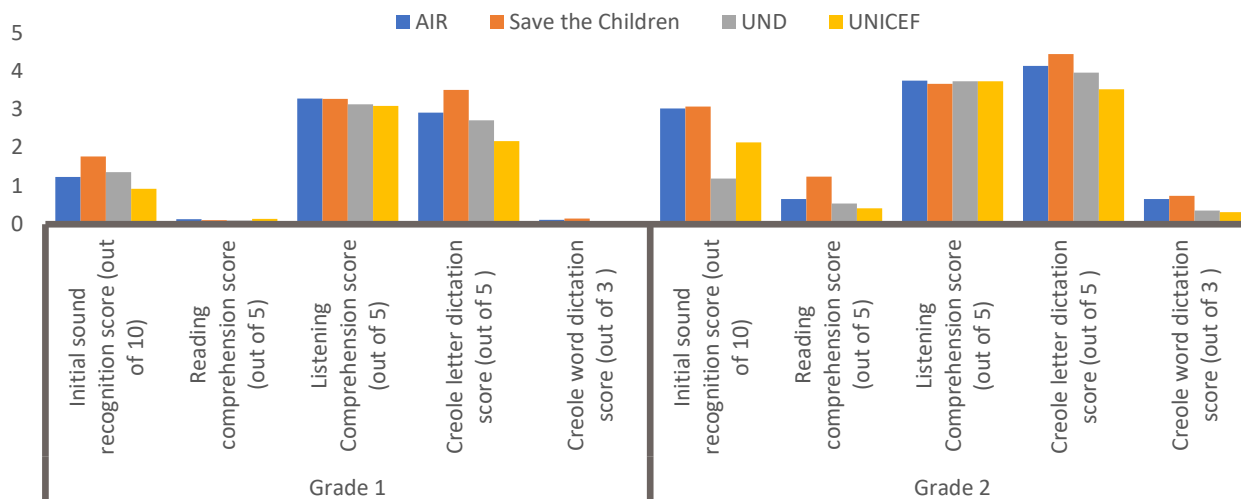


Figure 18: Non-timed learning outcomes, by grade and implementer



EGRA SCORES, BY SCHOOL TYPE

Variation in timed and non-timed EGRA scores also occurred across school types. The ET considered EGRA scores at the baseline disaggregated by school type, namely public, private, religious, community and “other” schools. “Other” schools refer to those listed as “semi-private” or “NGO.” Figure 18 presents timed EGRA scores by school type and grade. In Grade 1, all school types performed similarly on timed tests, particularly on the ORF test. However, by Grade 2, private, religious, and “other” schools performed much better than public and community schools. The differences were much less pronounced on non-timed tests, although private schools score slightly to across the five sub-tasks than other types of schools (Figure 19).

Figure 19: Timed learning outcomes, by grade and type of school

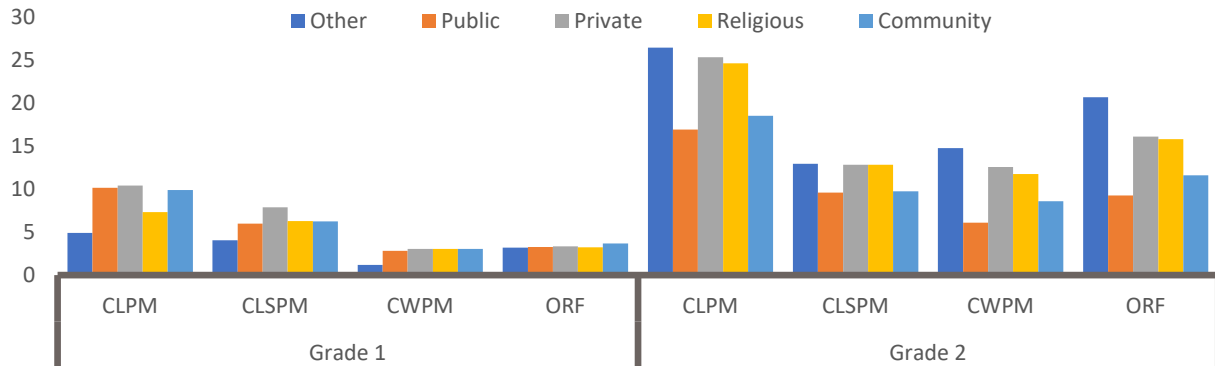
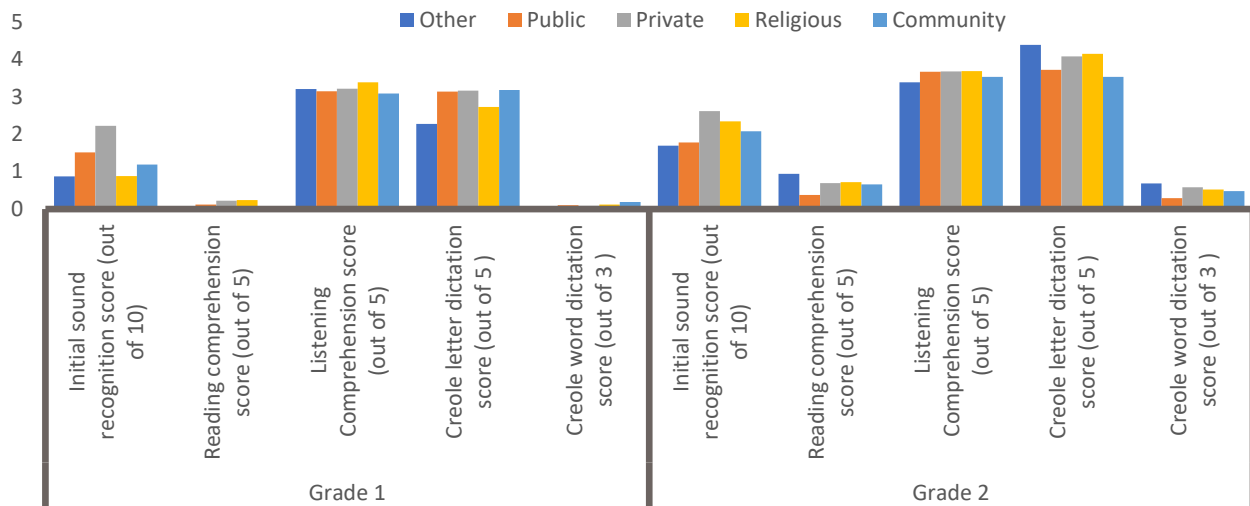


Figure 20: Non-timed learning outcomes, by grade and type of school



ENVIRONMENTAL AND SITUATIONAL DRIVERS OF READING FLUENCY

Thus far, this report has described the surveyed schools and students, noting variation where it occurred. The report has also explored the variation (or lack thereof) in EGRA scores across treatment groups, gender, school type, IP, and other factors. Next, the report examines the association of school, students, teachers, and classroom observation variables with oral reading fluency using regression analysis. The inclusion of variables from multiple sources (school, classroom observation, student) at baseline ensures a holistic approach to regression analysis as it helps account for as many potential explanatory variables of reading performance possible (Table 19). The above distributional analysis of EGRA test scores shows many zero scores, suggesting significant floor effects. Therefore, the ET used a censored regression model (Tobit) with lower limit censoring at zero.²⁶

²⁶ The Tobit model, also called a censored regression model, is designed to estimate linear relationships between variables when there is either left- or right-censoring in the dependent variable due to fixed minimum (floor) or maximum (ceiling) achievable values beyond which the values cannot go. For example, in a scale of 0 to 10, no score can be below zero or more than 10.

REGRESSION: EFFECTS OF SCHOOL, CLASSROOM, AND STUDENT LEVEL VARIABLES ON READING

The ET used four separate models to test out different combinations of explanatory variables. First, the team used a full-fledged model (model one) to account for the separate marginal effects of student, classroom, teacher, and school-level variables. Model three used indices of teachers' teaching methods from classroom observations, such as use of phonics approaches and gender equitable treatment of students. The remaining two models (two and four) repeat the same structure as Models one and three but used department-level fixed effects to control for heterogeneity from unobservable factors (using Artibonite as the reference point).²⁷ In all four models, the team controlled for IPs (using control schools as the reference point), because the school selection criteria for treatment differed by implementer. The team also controlled for school type (for example, public or community) using public schools as the reference point.²⁸

Table 19: Regression models showing the effects of various factors on ORF

Dependent variable: Correct Answers per minute for Oral Reading Fluency	Model (1)	Model (2)	Model (3)	Model (4)
A. From Student Survey	Model (1)	Model (2)	Model (3)	Model (4)
Teacher corrects the homework you completed the previous week?	1.862	2.241	2.030	2.379
Do you have time to do your homework or study the lessons after school?	0.317	0.381	0.103	0.245
Do you eat in the morning before school?	6.160***	4.607***	6.146***	4.624***
Did you (or will you) eat while in school today?	-0.836	0.538	-0.628	0.855
The last time you were given a good grade for a test or an exam at school, your parents were aware of it.	6.223***	4.915***	6.106***	4.928***
Did you attend preschool or kindergarten?	3.070	2.666	2.931	2.482
Do you have time to read in class or in the school's library?	2.986	3.941**	3.058	4.044**
Apart from your schoolbooks, do you have other books to read at home?	3.001**	2.499*	3.050**	2.463*
Student has Creole exercise book at the time of survey	3.044**	2.053	2.979**	1.951
Household asset index	-0.801**	-1.158***	-0.735*	-1.061***
Gender of student (1=Male)	-6.108***	-6.715***	-6.149***	-6.744***
Grade level	26.23***	26.81***	26.31***	26.73***
B. From Classroom Observation	Model (1)	Model (2)	Model (3)	Model (4)
Students spent time repeating teacher's words aloud.	-2.728*	-2.124	--	--
Teacher asks learners to recognize letters and say letter names and/or sounds.	-3.292	-3.183	--	--
Teacher provides instructions on how to blend sounds and/or decode syllables and words.	1.845	4.350**	--	--
Use of phonics approaches (index)	--	--	-0.752	0.110
Teacher assigns reading for learners to do on their own during school time.	2.041	0.347	1.337	-0.327

²⁷ The ET used Artibonite as a reference point because unlike other departments where around 50 percent or more of schools are rural, around 80 percent of sampled schools in Artibonite are urban schools, which is disproportionately higher than any other department.

²⁸ Based on the information from Fig. 18 that average score in reading (ORF) is lowest in public schools, the team used public school as a reference to measure relative performance of other types of schools from the regression model.

Dependent variable: Correct Answers per minute for Oral Reading Fluency	Model (1)	Model (2)	Model (3)	Model (4)
Teacher called on or actively tried to involve a student who was not participating.	0.429	-2.376*	--	--
Boys and girls are mixed together in the seating arrangement.	-5.496**	-5.637**	--	--
Girls are disciplined physically in class.	5.802**	3.393	--	--
Boys are disciplined physically in class.	-3.578	-0.605		
Gender equitable teaching (index)	--	--	0.317	-0.375
Students use the lesson books.	0.00968	1.074	-0.491	0.0875
C. From Administrative data collected from the Principals' Survey	Model (1)	Model (2)	Model (3)	Model (4)
Locational distribution (1=rural)	-6.637***	-8.745***	-6.420***	-8.501***
Average student to teacher ratio in the school	-0.0461	0.0196	-0.0497	0.0136
Average number of hours taught per day	0.331	1.438	0.952	2.166*
Principal's years of experience	0.267***	0.152*	0.304***	0.176*
Principal's years of experience as Director	-0.0622	-0.0843	-0.110	-0.135
Number of female teachers with formal teacher training?	1.198***	1.298***	1.160***	1.289***
Number of male teachers with formal teacher training?	-0.181	-0.198	-0.227	-0.249*
Does the school have an active parent-teacher association?	-4.050**	-0.858	-3.499*	-0.618
School management supports students with their school fees	0.122***	0.0572*	0.114***	0.0412
School has after-school extra-curricular activities	2.450	4.609**	2.142	4.104**
Does the school have a school feeding program?	2.136	-1.003	2.551	-0.419
Does this school have a reading club?	0.227	-3.012	0.872	-2.845
Does this school distribute student kits?	-2.227	-3.335**	-2.069	-3.838**
Does the school have a library?	5.840***	5.296***	6.213***	5.664***
Does this school benefit from infrastructure renovations?	-2.000	-1.943	-2.272	-1.947
D. Control variables	Model (1)	Model (2)	Model (3)	Model (4)
School Type (Public schools as reference)	--	--	--	--
School Type: Other	9.255**	7.139*	9.587**	6.241
School type: Private	10.79***	9.083***	10.20***	8.539***
School type: Religious	15.81***	13.55***	14.35***	11.33***
School type: Community	7.002**	7.067**	6.504*	6.292*
E. Implementers (control schools as reference)	Model (1)	Model (2)	Model (3)	Model (4)
AIR	1.966	2.605	1.234	1.465
Save the Children	13.03***	9.905***	11.28***	8.776***
UND	-6.355***	2.219	-6.353***	3.302
UNICEF	6.512**	1.641	6.316**	0.833
F. Department Fixed Effect (Artibonite as reference)	Model (1)	Model (2)	Model (3)	Model (4)
Grand'Anse	--	-13.22**	--	-14.60**
Nord	--	1.532	--	-0.0487
Nord-Est	--	9.782**	--	9.751**
Ouest	--	12.28***	--	10.70**
Sud	--	13.29**	--	10.70*

Dependent variable: Correct Answers per minute for Oral Reading Fluency	Model (1)	Model (2)	Model (3)	Model (4)
Sud-Est	--	31.33***	--	31.26***
Constant	-76.60***	-87.19***	-80.45***	-90.24***
Observations	3,102	3,102	3,102	3,102

Note: Asterisks represent the level of significance: (***) is strong ($p < 0.01$), (**) is moderate ($p < 0.05$), and (*) is weakly significant ($p < 0.1$).

Evidence from the regression analysis suggests that eating a meal before school, parental involvement in education, time spent reading, access to a library and to books at home are all positively associated with ORF scores. Among student-level indicators, students appear to have significantly better reading fluency if they are fed before coming to school or if the parent/parents or legal guardian are aware when the student received a good grade. These indicators are strongly significant across all models irrespective of specification. The above findings also suggest that there is a strong positive association between a student's oral reading fluency and possession of his or her Creole exercise books at the time of the survey, and if a student has other books to read at home apart from his or her schoolbooks. The team also found associations with positive ORF scores if a student has time to read in class or in the school library.

Across the board, girls score significantly better than boys. The two student-level variables that had strong significant effects on students' performance across all models are gender and grade level. The ET found that girls achieved significantly higher scores relative to boys. A clear progression in reading fluency was also observed as a student advances from Grade 1 to Grade 2.

"Girls are more engaged in [classroom] activities than the boys. Boys prefer to play – girls are more studious and alert. Sometimes while working with them, the boys will play under the bench, and you always have to check on them. Girls are more interesting, while you need to always discipline the boys," **Teacher from Nord-Est** (on students' participation in class)

Public schools are strongly associated with lower ORF scores, as are rural schools. School-level indicators suggest that there is a wide variation in reading scores across type of schools. With reference to public schools, students from private, religious, semi-private, or community schools performed better in terms of oral reading fluency. Religious schools were associated with the highest reading scores. Location mattered as well. After controlling for school type (private, public, community, or religious,) students in urban schools, on average, had significantly better oral reading fluency than their rural counterparts.

On average, more hours of instruction per day and lower teacher-student ratios are positively associated with ORF scores. Students do significantly better in terms of reading fluency when the average number of hours of instruction (per day at school) is higher. Also, higher student-to-teacher ratios lead to lower student reading scores. Although not significant in all four models, the quality of teaching (and therefore student learning, which is affected by certain measures of teacher quality) is adversely affected when the class size is larger.

School-level factors such as the existence of a library, school management assisting with students' school fees, school director level of experience, and the number of formally

trained female teachers at a school are all significantly associated with positive ORF scores.

Using school-level administrative information, the team found that the school director's years of experience is positively associated with overall reading scores. A higher proportion of formally trained female teachers in the school also has a positive influence on student reading performance. From a school management perspective, the ET did not find any conclusive evidence of significant (and positive) effects of parent-teacher associations (PTAs), school feeding programs, reading clubs, or distribution of student kits on reading outcomes, although the team did find evidence that extracurricular activities matter. Additionally, students perform significantly better in reading when school management assists students with their school fees. Overall, the team found strong evidence that students perform better when there is library in the school.

At the classroom-level, the use of phonics approaches by a teacher is not always associated with improved reading scores, and not all phonics approaches produce positive results.

Phonics approaches include asking students to repeat after the teacher, providing instructions on how to blend sounds and/or decode syllables and words, or asking learners to recognize letters and say letter names and/or sounds. The overall use of phonics approaches has an inconclusive effect on learning since this effect is not statistically significant. Individually, the various components of phonics approaches have different effects. For example, teachers providing instructions on how to blend sounds and/or decode syllables and words were associated with improved ORF scores. In contrast, in at least one regression model, teachers asking students to spend more time repeating the teacher's words aloud were associated with lower ORF scores. This result may be because students are more focused on repeating a word without understanding its meaning.

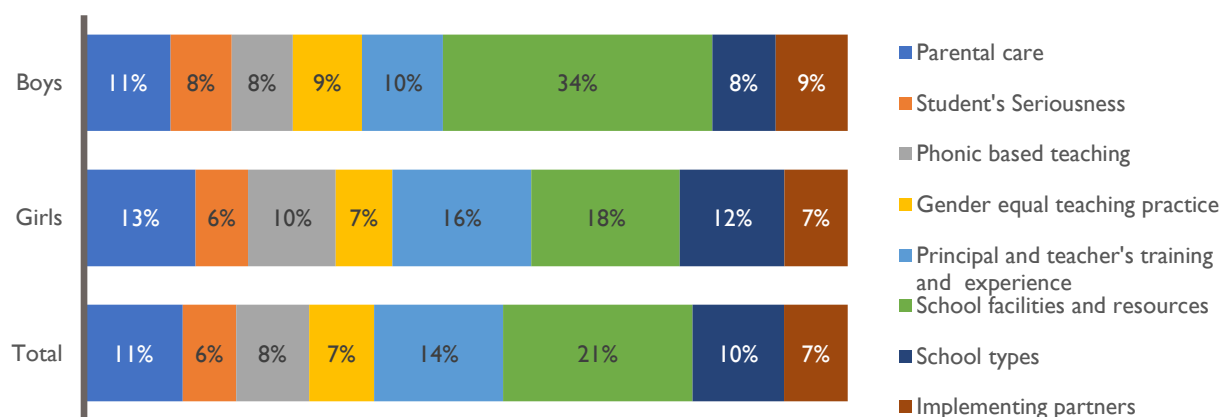
Mixing boys and girls together in seating arrangements is associated with significant adverse effects on reading. This result could be because teachers spend more time managing unruly behavior when girls and boys are seated together. On the other hand, it could be that mixing of students together is a proxy for a lack of sufficient materials (books, pencils, etc.) and resources in class. Corporal punishment in class appeared to have no consistent significant effects on reading.

DECOMPOSITION MODEL: RELATIVE SHARE OF DIFFERENT DRIVERS ON READING SCORE

The above regression analysis showed the relative importance of each driver in explaining variation in oral reading fluency. The discussion examines how each driver contributes in a group (with other similar covariates) to the total explained variation in ORF scores and whether these drivers have similar contributions across gender. To find out the relative contributions, or the share of how each driver contributes to overall ORF score, the ET grouped similar factors (drivers) into eight groups. The ET used a Shapley decomposition method²⁹ to isolate relative shares for each group of drivers. Based on this explanatory variable grouping, the decomposition analysis was done for all students together to find the overall contribution of various drivers to ORF and then separately for boys and girls to identify if there is any differential effect on gender (Figure 20). More information on the Shapley decomposition and eight factor groups can be found in Annex C.

²⁹ Shapley decomposition is a method that provides an additive decomposition of the statistic, allowing the reader to see the relative contribution of each regressor or contribution of a group of regressors of same kind in the total change that the regressors can explain.

Figure 21: Decomposition of drivers that explain ORF performance, overall and by gender



Overall, school-level drivers (for example, extracurricular activities, longer hours of operation, and student-teacher ratio), as well as teacher and director experience levels, explain the most variation in boys' and girls' ORF scores. Figure 20 shows that school-level activities and drivers (including provision of student kits, existence of extracurricular activities, existence of a feeding program, number of hours a school is open, and student to teacher ratio) contribute the highest share to ORF scores (21 percent overall). Strikingly though, there is wide variation between boys and girls with these school-level drivers explaining 18 percent of girls' ORF compared to 34 percent of that of boys.

Parental-level drivers play a somewhat small role in explaining ORF score variation, contributing to an average of 11 percent. Parental-level drivers (including making sure children eat before going to school or if a parent is aware of their child's grades) explain 11 percent of overall ORF, slightly more for girls (13 percent) than for boys (11 percent). A student's reading habits (including whether they have time to do homework and reading, have access to a library, have access to books at home) explains 8 percent of boys' ORF scores as compared to 6 percent for girls. Classroom teaching practices (including use of phonics approaches or gender equity practices) contribute to 15 percent of ORF scores overall (8 and 7 percent, respectively, for phonics-based approaches and gender-equitable practices) and explains equal shares of ORF for girls and boys. Similar to the findings from the regression analysis, gender-inclusive classroom practices are slightly less effective for girls than boys.

For girls, variation in reading scores is explained the most by teacher and principal experience levels and school-level factors, together accounting for 34 percent of ORF scores. For boys, school-level factors alone (in the form of school-sponsored activities, longer hours, and better student-teacher ratios) contribute to more than one-third of their ORF scores. Director or teacher experience and training explains 6 percent more of girls' ORF scores than boys (16 percent compared to 10 percent). Combined with previous regression analysis results, this could indicate a positive relationship between more and better-trained female teachers in a school and students' reading fluency, particularly for girls. This may be explained by the notion that more trained female teachers create an environment for girls to perform better.

INSIGHTS FROM PARENTS AND STUDENTS

The ET utilized FGDs and KIs to explain, supplement, and triangulate the quantitative findings from EGRA. Students (boys and girls separately) and parents (mothers and fathers separately) were asked about reading outcomes in the context of household, community, and school-based environmental factors as a means of explaining educational performance.

Parents had generally positive attitudes about gender equity in education. According to the qualitative research findings, in all 14 FGDs conducted with parents (divided by gender, i.e., seven with mothers and seven with fathers), mothers and fathers believed that both parents should play a role in deciding who will attend school. In all the parent focus groups, respondents indicated that it is critical to send both boys and girls to school so that they can be contributing members of society, have a profession, and earn a living to take care of their family. Fathers in Artibonite, Grande'Anse, Nord, Sud-Est, and West indicated that girls should attend school just like boys. One mother from the Sud-Est focus group remarked: *“Girls must go to school because they can become accountants, nurses, so that they can serve the population.”*

“School is life, hope for tomorrow... even if you are rich [but] cannot read, [then] you are sick.” -Mother of a female student, Nord-Est (on the importance of reading)

Parents appear slightly more likely to choose a girl for schooling over a boy, reflecting a recent change in behavior. Qualitative data suggest that when families need to make difficult decisions regarding which child to send to school, there is a slight tendency to believe sending the girl is a better investment. In focus groups for mothers held in Sud and Nord, all mothers responded that if they had to choose, they would send a girl to school, saying girls are a better investment because girls were more interested in and performed better in school. Other parents believed boys should be sent more often because men need to have a livelihood so that they can care for their families later in life. Many parents said the once common inclination to send boys to go to school if a choice had to be made between girl and boy family members is changing and that people are seeing the importance of educating children of both sexes.

Financial issues are the largest barriers for households when it comes to providing consistent education for children. Despite financial challenges, parents and family members will go to great lengths to support sending their children going to school. Mothers and fathers in all FGDs (parents in the Sud, Sud-Est and Ouest) said families take out loans, sell animals or vegetables, or take on an extra job to earn money for school fees. At the same time, many parents expressed disappointment, anger, and/or frustration with the failure of the government to assist families in educating their children. Parents explained that while school is technically “free,” schools and school districts are underfunded by the government and can’t function without the financial assistance of families who pay fees, including for school materials and uniforms, to keep their children in school. The regression results show that financial support provided by the school management to students is significantly associated with positive reading performance. Nearly all adult interviewees (parents, teachers, and MENFP staff) pointed out the significance of the financial barriers to education— all stemming from poverty and the inability of a family to marshal the necessary financial resources to send or keep children in school. A few parents said another barrier to school attendance was the distance

from home to a school. Children of young, single mothers were identified as the most marginalized in terms of educational opportunities.

Parent-teacher associations exist, though it is unclear how active they are. Most parents seem unaware of PTA activities. Quantitative data shows that 78 percent of schools have an active parent-teacher association (PTA). However, nearly all men and women in the parents FGDs, across all seven departments, said they knew of no parent-teacher program at their children's school. None of the 24 parents who participated in the Nord-Est FGD, the 19 parents interviewed in Artibonite, nor the 25 parents in the Grand/Anse FGD, knew of a parent-teacher program in their children's schools. However, most parents in the FGD conducted in the Sud said there was an active PTA at their school. Parents in Sud-Est and Ouest said that though there are no parent-teacher programs in their children's schools, or any other organized ways to influence the school's curriculum, they did work with teachers and principals to be more involved in their children's education. The regression analysis also shows no significant marginal effect of PTA on reading skills.

Most parents help their children with homework, although adult illiteracy limits parents' ability to assist with reading work specifically. Most parents believe that both parents should play an active role in their children's education, including helping the children with homework, purchasing school materials, and providing emotional support. Both fathers and mothers in all 14 FGDs said that they helped their children with homework as they were able. Literacy only in Creole was noted as a barrier for some parents in helping their children with their homework. Participants noted that mothers should or do assist their children with homework more often than fathers who might not have time. In the regression analysis, the effect of a parent helping with homework is positive but not significant in improving reading fluency.

Children report doing schoolwork often at home but are rarely involved in after-school activities related to reading. Girls have disproportionately more household responsibilities than boys. Few children indicated that they were involved in organized reading-related activities after school, such as a club or camp (except for two girls in the Ouest FGD). Rather, both boys and girls indicated that after school they did their homework and read Creole books or the Bible. In addition, children described having household responsibilities, including "working before school" and cleaning, cooking, and washing after they returned from school. These responses were largely attributed to the girls interviewed.

Generally, children enjoy going to school. All children interviewed said they enjoyed coming to school and liked to learn. They were enthusiastic about their teachers and classmates and about learning in general. Students believed that going to school and learning would lead to a profession and job that will allow them to help their families later in life.

Most girls (in five of seven FGDs) said that they enjoyed most things about school except the noise and disruptions caused by other students that prevented them from learning. There were no comments from parents or students related to safety while commuting to school or the state of school facilities (for example, classrooms, latrines).

Children speak openly about corporal punishment in class, though the frequency of corporal punishment use in the classroom is unclear. In three separate FGDs conducted in Ouest and Nord, boys and girls noted that part of why they did their lessons and paid attention in class

was so the teacher would not hit them. One girl from Grand'Anse reported that students would be punished physically if they did not do their lessons, as described below.

“[If a sibling does my work for me]...the teacher would recognize their handwriting and give me a whipping” –female student, Grand’Anse (on corporal punishment)

Early-age pregnancy and other issues act as barriers for many girls to attend school. At least one regional MENFP inspector said that she specifically works with schools on programs to avoid early-age pregnancy. Mothers from an FGD in Artibonite also suggested that some girls don't go to school because they practice prostitution to bring in money for their family.

EQ3: PRE-K INTERVENTION

Evaluation Question 3 is not addressed in this report as the status of the intervention was not known at the time of baseline data collection. Given that the Pre-K intervention will begin in Fall 2018, the ET will only collect data at endline in Spring 2020 on the first cohort of students exposed to the Pre-K intervention as they finish Grade 1. Their scores will be compared to Grade 1 students' endline scores.

EQ4: TEACHING QUALITY

Evaluation Question 4 looks at different aspects of teaching quality among Grade 1 and 2 teachers in treatment and control schools. Specifically, it seeks to answer the question of whether teachers in ERP schools teach differently than those in non-ERP schools and in what ways their teaching practices differ. Supplemented with qualitative data and regression analyses, the answers to EQ4 paint a picture of the potential impact of ERP interventions on teacher quality. To assess teacher quality, the evaluation team employed an adapted version of the Stallings' Classroom Observation Tool, described earlier in the report in the Data Collection Tools section.

Given that the ERP intervention had not yet begun at baseline, observation questions were general in nature and focused on good classroom management techniques, gender equity, the use and existence of certain educational materials, and phonics approaches. At midline and endline, the classroom observation tool may focus on more ERP curriculum-specific indicators. This baseline evaluation only reports fully totaled data, aggregated across all treatment and control schools. The results below highlight the most interesting findings.

REGRESSION ANALYSIS

Earlier in this report in the section on EQ1 and EQ2 findings, regression analysis demonstrated the role of classroom-level variables, such as teaching quality, on the key outcome of reading scores. The regression models included 10 key classroom-level variables (Table 19). The analysis showed an association between positive reading scores and the use of general phonics approaches by a teacher but with different positive and negative effects for different methods. Interestingly, the analysis also found that when boys and girls are mixed together in seating arrangements, there are significant adverse effects on reading. Corporal punishment in class appeared to have no consistent significant effects on reading. These findings suggest that how a teacher teaches and runs his or her classroom will continue to play an important role in the evaluation of early grade reading.

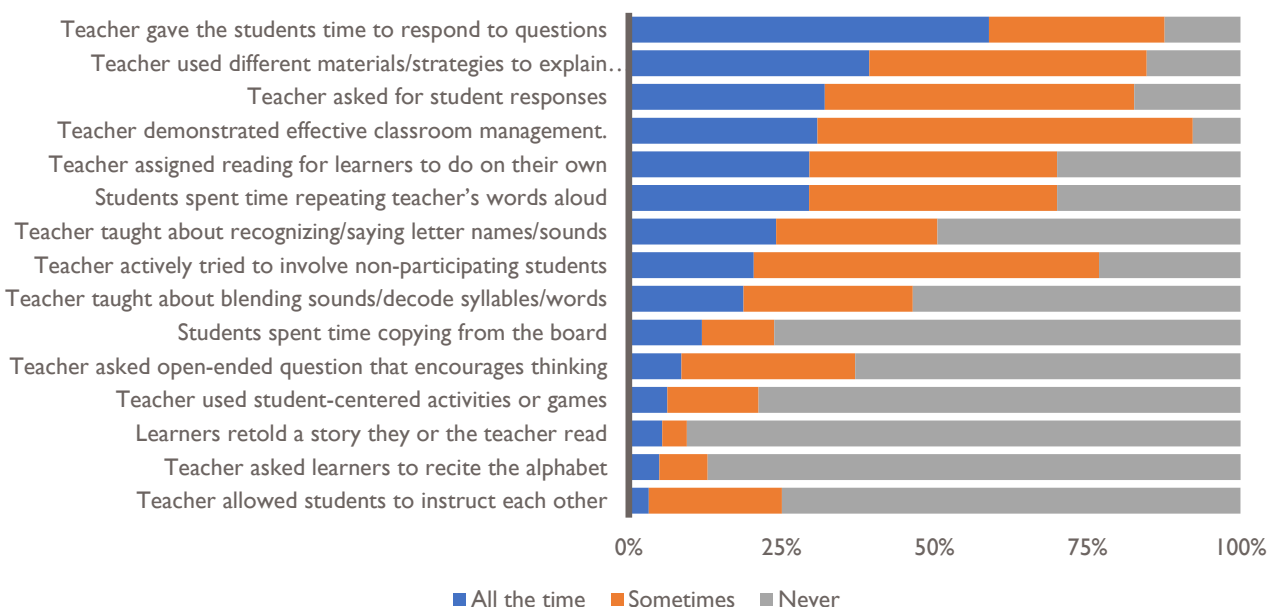
CLASSROOM MANAGEMENT PRACTICES

Teachers overall generally demonstrated effective classroom management (for example, efficiently managing materials, transitions between activities, and start and finish times), with over 90 percent of teachers doing so all the time or sometimes. Similarly, 85 percent of teachers were observed sometimes or all the time using different resources (books, tools,) and strategies (audio, visual) to explain concepts. Eighty-three percent of teachers were observed to solicit questions and/or answers from students. Just over three in four teachers were observed calling on or actively trying to involve a student who was not participating. Seventy percent of teachers assigned reading for students to do on their own during class time. In looking at classroom management techniques employed less often, the evaluation revealed that 37 percent of teachers utilized open-ended questions that encouraged thinking. Only one in four teachers were observed allowing students to instruct each other (come to the board to demonstrate something or explain a concept to a classmate). Students spent time copying from the board in 24 percent of classrooms. Further, 21 percent of teachers used student-centered games or activities in their lessons sometimes or all the time (Figure 21).

PHONICS AND READING PRACTICES

About half of teachers were observed using some type of phonics pedagogy in their classrooms. Given the focus of this evaluation on early grade reading, teaching quality around reading and phonics-based indicators was an important piece of the classroom observation activity. To that end, the tool incorporated several observation questions aimed at gauging a teacher's use of typical early grade reading methods. Observers marked behaviors "not applicable" if the educational activity being observed had no opportunity for incorporating reading practices (for example, if a teacher was teaching math). The "not applicable" data was removed from this analysis. In 70 percent of all applicable lessons, students were observed sometimes or all the time repeating teachers' words aloud, and in exactly 50 percent of classrooms, teachers asked learners to identify letters and say letter names and/or letter sounds. In 47 percent of classrooms, teachers were observed sometimes or all the time to have provided instructions on how to blend sounds and/or decode syllables and words. Thirteen percent of teachers asked students to recite the alphabet and in 9 percent of classrooms, children were observed retelling a story that they or the teacher had read (Figure 21).

Figure 22: Classroom-level variables on classroom management and phonics approaches



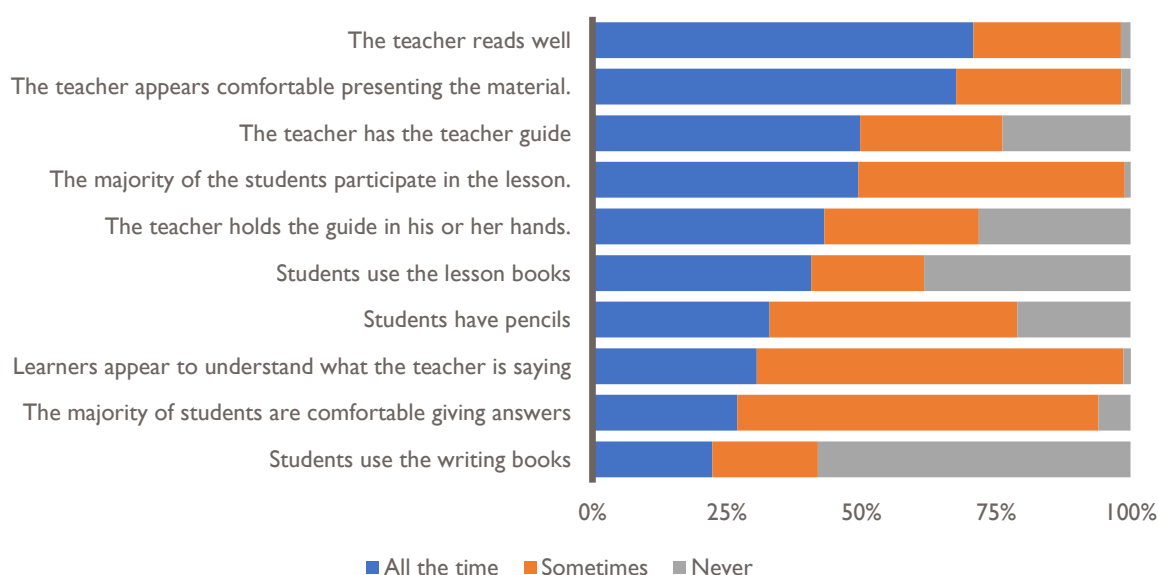
GENDER EQUITY

While not necessarily a focus of the ERP interventions, ensuring that girls and boys have equal opportunities to participate within the classroom is critical for equitable learning; it is also a tenet of good teaching practice. Teachers should create a friendly and encouraging environment where both girls and boys are comfortable participating. To that end, the evaluation included gender equity indicators in its classroom observation tool, focusing on the availability of materials and resources for both sexes, differences in verbal treatment and teacher encouragement of students, and the reality of corporal punishment. With respect to resource equitability, girls had equal access to desks and learning materials “all the time” in 65 percent of observed classrooms, “sometimes” in 28 percent, and no equal access in 16 percent of classrooms. Mixed seating arrangements, another potential predictor of gender equity, was observed “all the time” in 63 percent of classes, in nearly one in four classrooms, girls and boys were seated separately. On the topic of verbal treatment and encouragement, for all classrooms where teachers asked questions to students, three in four teachers asked questions equally of girls and boys. Twelve percent asked boys more questions and 11 percent asked girls more questions. Girls and boys were given equal time to respond to questions 57 percent of the time. Teachers provided encouraging feedback to students in 60 percent of classrooms, nearly equally to girls and boys. Teachers were observed to use a harsh tone in speaking to students in 15 percent of classrooms. Among those classrooms where a teacher used a harsh tone, it was used mostly equally with respect to boys and to girls (63 percent of the time) compared to classrooms where teachers used harsh tones when speaking to boys more often (26 percent) than when speaking to girls (11 percent.) Corporal punishment was observed in 12 percent of classrooms; it was used slightly more with respect to boys (13 percent) than to girls (11 percent).

USE OF LEARNING MATERIALS

At baseline, the ET did not include curriculum-specific observation questions because the ERP intervention had not yet been implemented. Instead, the team focused on the general use and availability of standard educational materials in the classroom for both students and teachers. Observation data showed that at the teacher level, slightly more three in four teachers (76 percent) had a teacher guide, and 72 percent of teachers held the guide in his or her hands sometimes or all the time while teaching. Students used lesson books in 62 percent of classrooms; they used writing books in 42 percent of classrooms. The availability of pencils was variable: 33 percent of students had their own pencils all the time, 46 percent had a pencil sometimes, and 21 percent did not have pencils (Figure 22).

Figure 23: Classroom-level variables on classroom materials and other topics



About half of MENFP personnel interviewed believed that most schools in their departments hire non-qualified teachers, and that many teachers are hired as political favors. Seven MENFP representatives reported that schools maintain a great deal of autonomy, but some schools occasionally accept recommendations from the Ministry for vacant positions. All MENFP interviewees mentioned that the Ministry implores public schools to recruit qualified teachers with teacher training certifications and university degrees. However, officials representing Grand'Anse, Nord, and Sud departments said in interviews that most schools tend to recruit or appoint individuals with no teaching qualifications. While many teachers have specific degrees in education and teaching experience, MENFP respondents pointed to other teachers whose specializations are in non-related fields, such as agronomy, business administration, or computer science. MENFP respondents proposed that these teachers are hired in return for political favors.

For teachers, a major barrier to effective teaching was a lack of professional development/training support. Most teachers said they received their most recent training from NGOs. A teacher in Nord department said, “I admit not being well trained.” He explained that his only training in the past three years was a 13-day training session on reading and writing offered by

NGOs, including CARE and Save The Children. Seven of the 14 teachers interviewed said more training in reading and writing and general teaching practices was needed.

There is a general lack of educational resources for teaching and students. Six teachers from the Sud, Nord, and Artibonite reported that many of their students come to class without books of their own, including students who attend the entire scholastic year without having a book of their own. Four teachers noted that the students who do not have materials are the ones who show little interest in reading and learning in class. One teacher from the Nord said although she requests books and other classrooms supplies, such as writing utensils, chalk, and notebooks, from MENFP, she has yet to receive materials in response. All teachers interviewed noted that the school and MENFP provide insufficient classroom materials to teach reading. One private school teacher from the Sud noted that her school does not receive classroom resources like the public schools do.

Teachers report that girls participate in and are more enthusiastic about reading than boys. This report is triangulated with other qualitative and quantitative data. Seven teachers from Sud and Artibonite, Grand'Anse, and Nordeast mentioned that girls seem to be more enthusiastic about reading and learning in general than are boys. For example, a teacher from the Sud indicated that “the boys are troublemakers” and don’t put as much effort into reading or volunteer as much often as girls do. When looking at EGRA scores, boys scored overall worse than girls (a significant finding at the $p < .01$ level).

Teachers with more experience are more able to discuss specific reading teaching methodologies and are more likely to use them. Most teachers only mentioned that they use general methods taught in teacher colleges. Depending on their teaching experience, some teachers mentioned specific methods, such as letter sounds, letter recognition, and repetition. Seven teachers described their literacy and reading teaching methods as students read out loud with the teacher correcting the student when they mispronounce a word. Teachers with more teaching experience said that they use sound and letter repetition. One teacher from the Nord-Est with 14 years of teaching experience said she has students repeat each portion of a text going “*step by step, word by word and sometimes letter by letter.*” Two other teachers, one from Artibonite with 17 years of experience and one from the Sud with 14 years of experience, said they used similar methods to teach reading.

“If we are teaching about the letter M, we just write it down in small and capital letters on the board. We pronounce the sound of the letter to them, afterwards we can ask them to make some words related in order to be sure they truly understand...” -Teacher from Artibonite (on reading pedagogical methods)

No teachers spoke of using corporal punishment in their classrooms to elicit or control behavior, yet students in FGDs reported the opposite. During FGDs, students spoke openly about the threat of corporal punishment being one of the methods teachers use to get students to complete assignments. However, during classroom observations, the use of corporal punishment as observed in only 12 percent of classes. The practice was not mentioned at all in teacher interviews. Given the sensitive nature of the topic and the potential bias created by having an observer present in a classroom, it is difficult to estimate the degree to which corporal punishment is used in the classroom.

Other school and household-level issues pose problems for teachers as well. All teachers interviewed noted some students travel long distances to school and late arrivals are a problem.

Interviewed teachers also unanimously agreed that large class sizes impede effective teaching and learning. A teacher in Grand'Anse cited large classes sizes as an impediment and reported that she has at least 70 children in her classroom. Two teachers from the Nordeast raised the issue of children needing to eat, noting that children who haven't eaten all day participate less in class.

EQ5: MENFP CAPACITY

EQ5 looks at MENFP capacity to implement and scale up ERP. Data for this question was collected entirely through qualitative key informant interviews with MENFP personnel. The ET interviewed two MENFP employees, typically a zonal inspector and a departmental coordinator or director, from each department. Interviewers asked respondents specifically about MENFP's financial, human, and material capacities and constraints and how those factors aided or prevented improvement in ERP implementation and scale up. Given that the ERP model focuses on curriculum implementation and teacher training, qualitative interviews focused on knowledge of reading pedagogy, capacity to financially facilitate access to curriculum, and ability to train and support teachers. The ERP program had not started at baseline, so most respondents were unaware of the program. At future data collection periods, qualitative data collection with MENFP personnel will be expanded to include questions on changes to their capacity, why the changes came about, and who and/or what was responsible for any changes.

A majority of MENFP officials interviewed were unable to name specific reading approaches or methodologies. According to MENFP officials' interviews, the MENFP employs few, if any, experts in literacy and reading pedagogy and methods. Only four out of the 14 officials interviewed were able to identify specific approaches to teaching reading or name a specific reading/pedagogical method used, promoted, or endorsed by the MENFP. Respondents mentioned using picture books and repeating letter sounds and letter names to teach reading, but they were unable to elaborate on the specific literacy teaching learning methodologies used or promoted by MENFP and recognized internationally. The MENFP interviewees said instead that MENFP promotes the use of general teaching methods taught in teacher colleges, adding that most schools have adopted such methods. While some interviewees mentioned older early grade reading programs in Haiti (ToTAL, Ann Ale), they did not mention the key component parts of these programs such as scripted lessons, teacher guides, or books.

MENFP staff can identify the challenges that affect effective teaching of reading, but they lack available avenues and resources to address these challenges. Three inspectors noted that MENFP encourages all schools to utilize Haitian Creole as the language of instruction in early grades, but certain schools, particularly the religious ones, tend to ignore this policy and teach mostly in French. Although the officials believe that this policy of noncompliance contributes to the low literacy levels among early grade students, no MENFP officials mentioned programs or mechanisms to directly address this policy.

No concrete plans appear to exist to improve teacher vetting and recruitment processes. MENFP interviewees noted that most schools appoint unqualified individuals with no education degrees. Out of the 14 MENFP officials interviewed, only one commented, briefly, on MENFP plans to ensure professionals work in their respective fields. One MENFP official from the Nord indicated that the MENFP is developing plans to ensure that all professionals are working in their respective fields and possess substantial qualification, though no further details was provided regarding these efforts.

“There are some trainings provided through NGOs. For example, we have the ‘Ann Ale’ program – this is the way it’s done. We do not directly conduct training ourselves as we have no means for that.”-MENFP official from Ouest Department (on professional development training for teachers)

The government is unable to consistently provide training and continuous professional development for teachers. Four MENFP officials from the Ouest and Nord-Est noted that the

“Some school principals do not collaborate with the Ministry and even though you go to the school to get the information, it is not available or just does not exist. You may have an appointment with them and you don’t find them and for no reason, they are not present ...” -MENFP official from the Sud department (on barriers regarding monitoring visits at schools)

MENFP provides training programs for teachers, especially when schools are performing below MENFP expectations. But due to the small budget allocated to their department, these programs are seldom actualized, and very few teachers receive training on a regular or even one-time basis. This information matched that from teacher KII findings where at least half of interviewed teachers reported they had gone years without exposure to any regular training opportunities offered by MENFP. Findings from parents FGDs and MENFP KIIs noted that teachers lack proper qualifications and training.

MENFP relies heavily on NGOs to provide materials and funds to schools. NGOs appear to provide significant educational support to teachers and schools. Despite the lack of government and school administration support, some teachers mentioned that they receive help from NGOs. Four teachers from the Nord and Artibonite departments specifically mentioned the USAID programs ToTAL and Ann Ale as providing them with teaching resources, such as curriculum books. MENFP KIIs validated these findings: All MENFP respondents confirmed that NGOs collaborate with MENFP in most departments to provide necessary resources to schools. Eleven of 14 MENFP officials stated that MENFP works closely with NGOs to provide teacher training to schools. Ten MENFP officials noted that schools rely mostly on billing parents for fees and NGO assistance to operate. Nearly all MENFP officials interviewed mentioned that to implement any programs, or provide school materials, they mostly work with NGOs.

Ministry officials face occasional barriers to making monitoring visits at schools. According to MENFP officials, it is mostly the schools that create barriers to conducting successful monitoring visits. Four officials from the Nord and Sud reported that some school principals often refuse to cooperate with the Ministry by restricting access to documents or canceling appointments. Four additional officials from Grand’Anse and Artibonite said that schools often misplace pertinent documents such as attendance sheets, tests results, enrollment records, and teachers’ education diplomas. Two officials shared their frustrations of being assigned too many schools for inspections and not being able to carry out all inspections assigned. Quantitative data from the head teacher survey, however, shows that 75 percent of schools were visited by an MENFP official in the previous 12 months.

CONCLUSIONS

GENERAL

LACK OF EMIS SYSTEM

The lack of an EMIS system in Haiti hampers educational research and program implementation. The availability of school-level data is essential for conducting rigorous research and for better understanding gaps in educational access throughout the country. Without regular, authoritative census data on schools nationwide, implementers, evaluators, and public-sector actors at the MENFP will be limited in their ability to conduct high quality programming and respond to the needs of communities.

EQ1 AND EQ2: READING SCORES

Disproportionate numbers of zero scores have led to a situation where EGRA scores at baseline are imbalanced across treatment status, ERP model, gender, implementing partner, and school type. The evaluation team found that students in Grade 2 in the core treatment group outperform their counterparts in the enhanced and control groups significantly in almost all subtasks (timed and untimed). Variation in EGRA scores at baseline was also noted across genders, school types, and implementers, with students in Save the Children schools scoring much better than students in other implementer's schools. Girls score significantly better than boys across all tests (and in both grades), and public and rural schools had lower oral reading fluency scores relative to public, community, religious, and urban schools. Where variation was noted between treatment and comparison groups, the difference-in-difference approach will ensure non-biased analysis at the endline, although some initial differences may still exist.

Surveyed students have Creole language skills but do not yet have reading skills. Students in both Grade 1 and 2 scored highest in Creole letter dictation and listening comprehension, but lowest in Creole word dictation and reading comprehension. This was to be expected in Grade 1 and somewhat in Grade 2 as students begin their formal education, but higher grades in primary school should begin to show higher reading scores. This is in line with other EGRA studies in Haiti that found students were not able to read at sufficient speed (with comprehension to understand text) until Grade 3 or later.³⁰

The home environment plays a critical role in a student's reading scores. At the student-level, students who eat before school, have parents involved in their education, spend time reading, have access to a library and have books at home have higher oral reading fluency scores than students without these factors. Qualitative data showed that parents are involved and invested in their children's education (for both girls and boys) but are mostly unaware of PTA activities even though school directors say nearly three in four schools have active PTAs. Parents across the board noted that school fees are a major barrier to school attendance, although they are willing to go to great lengths to pay for their children's education. While at baseline girls reported taking on a disproportionate amount of

³⁰ EDDATA II: Education Data for Decision Making. The Early Grade Reading Assessment (EGRA) in Haiti. June 2014. Prepared by RTI International. Contract Number EHC-E-00-04-00004-00. www.eddataglobal.org.

housework outside of school, a shift in community attitudes seems to suggest an evolving opinion that education for girls is as important as education for boys.

School-level factors explain the largest amount of variation in oral reading fluency for girls and boys. Regression and decomposition analysis found that a host of school-level variables have significant positive associations with reading scores. School management assisting with students' school fees, school director experience level, and number of formally trained teachers at a school are all associated with positive oral reading fluency scores. Notably, while a minority of schools report having extra-curricular programs and services, having these programs is associated with higher oral reading fluency. Furthermore, students are associated with better reading fluency when the average number of hours of instruction (per day at school) is higher. As to be expected, as the student-to-teacher ratio increases, student reading scores are lower. Girls appear to benefit most from more and better-trained female teachers at a school.

In the classroom, use of certain phonics approaches works better than others, and seating arrangement matters. While about half of teachers used a phonics approach while teaching reading, those teachers who practice blending sounds and decoding syllables saw better results in reading scores than those who simply asked students to repeat words aloud. Treatment of girls and boys in the classroom was gender equitable; however, regression analysis showed that seating boys and girls together was associated with significant adverse effects on reading. The finding could be a proxy for the need to share educational resources among children.

EQ4: TEACHING QUALITY

Classroom management is generally good, though availability, access to, and use of educational materials in the classroom are highly variable. Grade 1 and 2 teachers practice good classroom management and gender equitable treatment of students. Although students openly claim that teachers practice corporal punishment, classroom observations revealed corporal punishment to be very low. Classroom observations demonstrate both inconsistent use of books and a lack of availability of basic items such as pencils. Qualitative data from MENFP, teachers, and parents suggest a persistent scarcity of books and materials.

Use of pedagogical reading approaches in class is not consistent. Only about half of teachers used a recognized phonics approach during teacher observation. Teachers with more years of experience were more likely to be able to specify good phonics practices in teacher KIs. Reading aloud to students was the most practiced phonics approach, although it does not necessarily correlate to improved reading scores.

Low levels of teacher formalization probably contribute to low reading scores. Only about a quarter of teachers in sampled schools were formalized. Teacher trainings are nonexistent and/or very irregular; many teachers have no formal training. Most teachers reported they may not receive training for years at a time. The MENFP admits that it lacks the resources to hold regular trainings. This situation could pose a major barrier to improvements in reading, given the regression analysis finding that formalized teachers with more experience are associated with higher oral reading fluency scores, especially for girls.

EQ5: MENFP CAPACITY

Departmental-level MENFP lacks the financial resources to adequately support schools. Quantitative and qualitative data supplement each other to paint a picture in which schools do not

receive books and other educational materials, teachers are sometimes delayed payment, and there is no MENFP-supported regular teacher training. Families must take on the burden of supporting education for their children, and when they cannot; children go to school without necessary supplies or are held back. At baseline, MENFP offices relied heavily on NGOs to support and sustain schools with materials, funds, and teacher training.

Departmental-level MENFP staff are unaware of policies or plans through which obvious deficiencies in the educational system can be improved. Officials were able to cite problems, such as non-compliance by teachers of the official language policy, high levels of unqualified teachers, and irregular teacher training, but they could not specify mechanisms or actions through which these issues could be resolved.

MENFP officials carry out monitoring visits often but do not appear to be deeply knowledgeable about phonics/reading pedagogy. Nearly three in four schools sampled had received an MENFP visit in the past year. However, most MENFP officials interviewed could not discuss reading pedagogy in depth. This could have implications for policy development and quality oversight of reading programs in the future.

GENDER

As a critical part of all evaluations, the ET looked at issues surrounding gender when evaluating the ERP at baseline. In general, the team found positive findings around gender equity in the household and classroom when it came to education. For instance, girls are sent to school at an equal or greater rate than boys and treated equally well (or punished equally) in the classroom. In fact, qualitative FGDs with parents revealed that parents choose girls to go to school more than they have in the past, signaling a shift in attitudes towards girls' education. Some discrepancies remain, however. For instance, girls have disproportionately more work to do at home than boys.

Perhaps the most important finding around gender was that girls have significantly higher reading scores across all measures when compared to boys, though this finding is not entirely surprising. Studies have shown that girls do better than boys in most subjects, probably due to the fact that girls study more, have more positive perceptions of school, and behave better,³¹ findings supported by teacher KIIs and girls' FGDs held at baseline. An interesting finding of the regression analysis showed that when a school has more and better-trained female teachers, girls in particular have better reading scores. This could suggest that female teachers enable and encourage girls to attend and participate in school more than the girls typically would.

DISSEMINATION PLAN AND UTILIZATION

Transparency is a key principle of USAID evaluations, and ADS 201³² requires operating units to plan for evaluation dissemination to promote transparency and learning. As noted above, the primary stakeholders for this evaluation include the MENFP, USAID/Haiti, the implementing partners of the ERP (Save the Children, AIR, UND, and UNICEF), and other donors and entities active in Haiti's education sector (including the World Bank, JICA, IDB, World Vision), as well as a variety of NGOs and religious

³¹ Voyer, Daniel, and Susan D. Voyer. "Gender differences in scholastic achievement: A meta-analysis." *Psychological bulletin* 140.4 (2014): 1174.

³² <https://www.usaid.gov/ads/policy/200/201>.

affiliations devoted to improving learning in Haitian primary schools. Additionally, the Scope of Work for this evaluation notes that evaluation findings will:

- Inform USAID/Haiti's future programming in early grade reading in Haitian schools;
- Inform Government of Haiti policy-level decisions around implementation of reading curriculum and teacher training on a national scale; and,
- Guide future programming for other donors' activities as they relate to early grade reading in Haiti.f

With this wide and varied group of stakeholders, and with potential uses of evaluation findings already identified, it is helpful to develop a dissemination plan to ensure utilization. Although dissemination activities were not written into the evaluation Scope of Work, the evaluation team includes recommendations here based on its interaction with the primary evaluation stakeholders.

After conducting a stakeholder analysis to ensure that the list of above stakeholders is comprehensive and accurate, USAID should identify and/or further clarify the goal in disseminating evaluation results to each stakeholder. For example, with MENFP, the goal may be policy change. With ERP implementers, the goal may be to clarify interventions and approaches based on a detailed baseline context. With other donors and entities active in Haiti's education sector, the goal may be two-fold: to share technical/education-related knowledge and to prevent repetition of ineffective approaches.

After a goal is identified and/or clarified for each stakeholder, the Mission should identify appropriate communication tools to reach their audience. These may include tools such as reports, briefs, blog posts, press releases, graphics, and e-mails to listservs. They may also include more involved efforts including presentations, meetings, facilitated workshops or discussions, videos, and journal articles.

Lastly, the Mission should consider the timelines required for dissemination of evaluation findings. It may be more critical to provide ERP implementers with baseline findings relative to other donors, for example. A helpful template to organize the dissemination plan can be found In Annex H.

Developing and implementing a dissemination plan will help ensure that these baseline evaluation findings are utilized not only for the improvement of the ERP interventions, but they also may be useful in increasing the efficiency of education work being planned or conducted in Haiti today.

ANNEX A: FULL LISTING OF REFERENCES IN REPORT

Abadzi, H. (2006). *Efficient learning for the poor: Insights from the frontier of cognitive neuroscience*. Washington, DC: The World Bank.

Ann Ale EGRA Baseline Evaluation. 2015. Various tools including EGRA, Student Stimuli, and Classroom Observation. Prepared by FHI360.

Concern Worldwide. Early Grade Reading Assessment (EGRA) Saut d'Eau – Haiti. Final Report. May 2013. CIA World Factbook, updated 1 January 2018. Obtained on 2 March 2018. EDDATA II: Education Data for Decision Making. The Early Grade Reading Assessment (EGRA) in Haiti. June 2014. Prepared by RTI International. Contract Number EHC-E-00-04-00004-00. www.eddataglobal.org.

EDDATA II. The Snapshot of School Management Effectiveness: Report on Pilot Applications. EdData II Technical and Managerial Assistance, Task Number 1. Contract Number EHC-E-00-04-00004-00. Strategic Objective 3. Prepared by RTI International. October 2008

Early Grade Reading and Writing Assessment in Ann Ale Schools With the Coach-Led Learning Assessment Monitoring System (CLAMS), Final Report- French Version, May 2016, USAID

“Early Grade Reading Barometer”. n.d. Accessed March 14, 2018.
<http://earlygradereadingbarometer.org/iraq/comparisons/sdg>

King, G., Blackwell, M., Iacus, S., & Porro, G. “Coarsened exact matching in Stata.” (2010); Rosenbaum, PR, and DB Rubin. “The central role of the propensity score in observational studies for causal effects.” *Biometrika* 70, no. 1 (1983): 41-55.

Stefano M. Iacus, Gary King and Giuseppe Porro (2011): *Causal Inference without Balance Checking: Coarsened Exact Matching*, Published by Oxford University Press on behalf of the Society for Political Methodology

Programme Haitien d'Appui à la Réforme de l'Education (PHARE) & American Institutes for Research (AIR). (2009). *Draft: Student diagnostic assessment baseline report*. Prepared for USAID/Haiti and the Government of Haiti. Washington, DC: AIR.

Results for Education and Child Health (REACH). 2015. Various tools, including enumerator manual, trainer manual, school observation tool, school director KII guide, and teacher KII guide. Prepared by FHI360.

RTI International. 2015. Early Grade Reading Assessment (EGRA) Toolkit, Second Edition. Washington, DC: United States Agency for International Development.

Shorrocks, A. F. (2013): “Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value.” *Journal of Economic Inequality*, Vol. 11, Issue 1, pp. 99-126.

Suzata, Eriko. 2011. Education in Haiti: An Overview of Trends, Issues, and Plans. World Innovative Summit for Education.

Tout Moun Ap Li (ToTAL) Final Report, revised, January 2015, RTI International.

USAID, ADS Chapter 201, Program Cycle Operational Policy, Functional Series 201, Partial Revision
Date: 02/12/2018, Responsible Office: PPL , File Name: 201_021218,
<https://www.usaid.gov/ads/policy/200/201>.

USAID. Haiti ESS ERP Impact Evaluation Scope of Work. October 2017.

Voyer, Daniel, and Susan D. Voyer. "Gender differences in scholastic achievement: A meta-analysis."
Psychological bulletin 140.4 (2014): 1174.

Wagstaff, A.; H. T. H. Nguyen; H. D.; and S. Bales (2014): Encouraging Health Insurance for the Informal Sector: A Cluster Randomized Trial, World Bank Working Paper Series, No. 6910.

World Bank. 2015. Conducting classroom observations: analyzing classrooms dynamics and instructional time - using the Stallings 'classroom snapshot' observation system: user guide (English). Washington, D.C.: World Bank Group.

ANNEX B: DATA COLLECTION TOOLS

Impact Evaluation of USAID/Haiti's Early Reading Program

Guide to Teachers' interview

Consent script:

"Good morning/afternoon. Thank you for finding time to meet us. We are here on behalf of Evaluation Support Services, Social Impact, a research company working on a research program into education in Haiti. This is part of the Early Grade Reading Project implemented by USAID. We are trying to understand the situation related to early grade reading in the schools in this community. We are looking for typical teachers engaged in early grade reading as an instructor to tell us their opinions and common conditions in this area related to early grade reading. There are no right or wrong answers. We want you to have the opportunity to express your views freely. We would like to ask your permission to participate in this activity. If you agree, we would hold this open discussion with you and record the discussion on this device [show recorder]. Although the topics we discuss should not be sensitive, we will not ask any of your names, and we will not share this recording with anyone who is not involved in this study. In this way we will protect the confidentiality of your responses. You are free to refuse to answer a particular question if they don't want to. There are no risks to participating in this discussion other than the 1 hour of your time that you will give up. There is no direct benefit to you if you participate, other than knowing that you are helping us to assess challenges related to education that may later help Somalia.

We would like to express our appreciation for your time through these modest refreshments.

- **Do you have any questions about this activity?**
- **Do you wish to participate? (record yes or no)**

PRE-CODING:**Date:** _____ **Community Name:** _____**District:** _____ **Village:** _____**Nearby evaluation school code:** _____**Nearby evaluation school name:** _____**FGD PARTICIPANT INFORMATION** *(DO NOT RECORD NAMES):*

Person #	Age	Gender (m/f)	Personal education level	Grade(s) currently teaching	Grades taught in past (if different)	Years teaching experience overall
#1						
#2						

TEACHER DEVELOPMENT, SUPPORT AND MOTIVATION:

1. What are the education and training requirements individuals must complete prior to becoming a teacher?
 - a. Do they vary amongst locations?
 - b. Do they have to continue training once they begin teaching?
2. What are the education and training requirements individuals must complete prior to becoming a teacher to train early grade students in reading?
 - a. Do they vary amongst locations?
 - b. Do they have to continue training once they begin teaching?
3. Do you think teachers receive adequate training generally?
 - a. Why or why not?
4. Do you think teachers receive adequate training to train early grade students how to read?
 - a. Why or why not?
5. Are there trainings that are on-going in the school where you work?
 - a. What type of training?
6. Are there trainings that are on-going in the school where you work with a particular focus on early grade students to read?
 - a. What type of training?
7. What additional training do you think teachers either need or would benefit from to improve early grade student's reading ability?
8. What types of materials and resources do teachers use in the classroom for early grade reading?
 - a. Which do you find to be most effective and why?

9. What additional materials and resources do you think would be beneficial to improve early grade reading?
10. What are some of the challenges schools face in acquiring learning materials/resources for improving early grade reading?
11. What kind of support do teachers usually receive in the classroom to train early grade students in reading?
 - a. How often do they receive this support?
 - b. What other things could be done to support teachers in the classroom?
12. Does the Ministry of Education play a role in supporting teachers through providing teaching materials or resources for teaching early grade students read fluently?
 - a. If so, please describe.
13. What interventions or projects have been instituted at your school to promote early grade reading? (PROBE for any that may have focused on advancing girls' education)?
 - a. If so, how well do you think these address any student enrollment, attendance, retention, and/or teaching challenges that affect your school?

PEDAGOGY, STUDENT ENGAGEMENT, AND PARTICIPATION IN EARLY GRADE READING

14. What are the biggest challenges teachers face in the classroom to improve early grade reading?
15. What do you think would be most helpful to address those challenges?
16. What are common methods that teachers use in the classroom to teach early grade reading?
17. Are students highly engaged and interested in reading?
 - a. Why do you think this is?
 - b. Do you see gender differences among students who are most interested and engaged in? If so, why?
18. Do students seem less engaged and interested in reading?
 - a. What do they struggle with more?
 - b. Do you see gender differences among students who struggle with? If so, why?
19. In a typical classroom, what differences do you see among students in terms of participation? In your opinion, what explains some of these differences?
20. Do boys and girls behave differently in reading class?
 - a. If so, in what ways?

Interview Guide

Haiti Education Service or Ministry of Education Officials

Consent script:

"Good morning/afternoon. Thank you for finding time to meet us. We are here on behalf of Evaluation Support Services, Social Impact, a research company working on a research program into education in Haiti. This is part of the Early Grade Reading Project implemented by USAID. We are trying to understand the situation related to early grade reading in the schools in this community. We are looking for typical MOE official employed in education sector to tell us their opinions and common conditions in this area related to early grade reading. There are no right or wrong answers. We want you to have the opportunity to express your views freely. We would like to ask your permission to participate in this activity. If you agree, we would hold this open discussion with you and record the discussion on this device [show recorder]. Although the topics we discuss should not be sensitive, we will not ask any of your names, and we will not share this recording with anyone who is not involved in this study. In this way we will protect the confidentiality of your responses. You are free to refuse to answer a particular question if they don't want to. There are no risks to participating in this discussion other than the 1 hour of your time that you will give up. There is no direct benefit to you if you participate, other than knowing that you are helping us to assess challenges related to education that may later help Somalia.

We would like to express our appreciation for your time through these modest refreshments.

- **Do you have any questions about this activity?**
- **Do you wish to participate? (record yes or no)**

Date: _____ **Time of start of interview:** _____

Office of MoE: _____

Position or Title: _____

Zone of responsibility (if applicable): _____

District of responsibility (if applicable): _____

GES/MoE Overview

1. What are your responsibilities within the Ministry of Education?
2. What do you see as the biggest challenges to Early Grade Reading in the area that you are working in?
3. Are there any major policies or initiatives the MoE is working on right now to address these challenges, or other issues?
 - a. What is the timeframe for their rollout?
 - b. Are other outside organizations helping to address these challenges?
4. Does the GES/MoE have any specific policy, guidelines, or practices related to gender or girls' education? Explain.
 - a. Does this include any structures to support girls education such as a girls education unit, girls education officers etc?
5. How is funding determined for each school, zone and district?
 - a. How long has this been in place?
6. Generally, what has been the level of involvement of NGOs in MoE planning and funding?
 - a. How have NGOs and the MoE worked together?
7. At what level do teachers participate in curriculum development?
 - a. (Probe: Are there national standards? Do they help build those standards and guidelines?)

TEACHER DEVELOPMENT

8. What capabilities / qualifications do teachers need to get a job at a school?
 - a. Are these always enforced? Why or why not?
9. How are teachers recruited to work at schools?
 - a. What difficulties does the ministry have in recruiting teachers?
10. What types of teaching methods does the MoE promote in relation to Early Grade reading?
11. In your experience, what teaching methods do you see being used in schools for Early Grade Reading?
 - a. If they vary from those recommended by the MoE, why do you think this is?
12. Does the MoE currently provide further trainings or coaching for teachers who are already teaching Early Grade Reading in schools?
 - a. If not, why not?
 - b. If so, please describe some of the trainings available to teachers.

- c. If so, who plays this role, and how often?
 - d. If so, are these trainings available to all teachers? Why or why not?
 - e. If the MoE is not currently supporting teacher trainings, why is this? (PROBE: resources, lack of necessity?)
13. Is the MoE providing any other kinds of support to teachers (e.g. teaching materials or resources)?
- a. Why/why not? If so, please describe.
14. Are there NGOs or other agencies that have supported you in your teacher development efforts?
- a. Please describe briefly and comment on their value added, if any.
15. Does the Ministry of Education currently support educational radio or television programming? Please describe.
- a. Is this at the national, regional or community levels?
 - b. How much of an influence do you think such programs have (if yes) or could have (if no) on increasing enrollment, attendance, retention, and learning?
16. As a MoE official, what challenges or areas for improvement do you see with regard to teacher effectiveness or education quality in this country?

DATA/ STUDENT ENGAGEMENT/ IMPROVED LEARNING OUTCOMES

17. What kind of information does the MoE gather from schools, such as enrollment records, attendance records or reading test results?
- c. How are these collected and how often?
18. What are some of the challenges and difficulties that the MoE faces in collecting these information from schools?
19. What is done with the data once the Ministry receives it?
20. Is there any positive or negative action taken by the MoE towards a school if it has high or low learning outcomes? Please explain.
21. Is there anything else you would like to talk to me about with regards to education?

Time interview ended _____

Focus Group Discussion Guide

GIRL/ BOY STUDENTS

PRE-CODING:

Date: _____

School Community Name: _____

Evaluation school name: _____

Evaluation school code: _____ Type of school _____

Departments: _____ Village/Town/City: _____

All girl or all boy participants? (Circle) Girl Boy

FGD PARTICIPANT INFORMATION (do not record names):

Person #	Age	Grade in school
#1		
#2		
#3		
#4		
#5		
#6		
#7		
#8		
#9		
#10		

STUDENT MOTIVATION, ENGAGEMENT, AND PARTICIPATION

I. How do you feel about going to school?

PROBES:

- a. What do you like most about going to school?
- b. What do you like least about going to school?
- c. How does school make you feel – anxious, excited, curious, scared, safe? Why do you feel this way?

2. During a typical day in your school, what activities do you do?

PROBES:

- a. Walk me through a typical day, from the time you arrive to the time you leave, what do you do?
- b. What parts do you like the most?
- c. What parts do you like the least?

3. How do you feel about your class?

PROBES:

- a. What do you enjoy the most in your class?
- b. What do you least enjoy?

4. How do you feel about reading?

PROBES:

- a. What do you like most about your reading in your class?
- b. What do you dislike about reading in your class?

5. How does school help you to be a better reader?

PROBES:

- a. How do you feel about the way you are being taught how to read?
- b. What do you like most about it?
- c. What do you like least about it?

6. What would help you read better?

PROBES:

- a. Listening to the teacher?
- b. Focusing on the class work?
- c. Paying attention to the study materials?

7. How will school and being a good reader help you in the future?

8. What, if any, differences in girls' and boys' participation in class do you notice?

PROBES:

- a. What are those differences?
- b. Why do you think that is?

QUALITY OF TEACHING

9. How do your teachers help you learn?

PROBES:

- a. What else could they do to help you learn even more?

- b. What do you like about the way your teacher teaches you how to read? What do you dislike?

10. What types of learning materials/aids/resources do you or your teachers use at school?

PROBES:

- a. How helpful do you find these?
- b. What additional materials/aids/resources do you wish you had?

LEARNING OUTSIDE OF SCHOOL

11. What do you like to do with your time outside of school?

PROBES:

- a. attend clubs, watch TV, listen to radio, etc.

12. What reading do you do outside of school?

PROBES:

- a. homework, reading, education programs?

13. How often do like you you get help with school work from your parents or other family members?

PROBES:

- a. How often do your parents help you with your school work, especially reading?
- b. How often do your friends' parents help them?
- c. If not why not?

STUDENT ATTITUDES TOWARD EDUCATION

14. How do you feel about discussing about school each day with other people in your family?

PROBES:

- a. What do you discuss about school with your parents?
- b. What do they say to you about school? (PROBE: encouraging you to do well in school, attend school regularly)

15. Is there anything else you would like to tell me with regards to education?

Focus Group Discussion Guide MOTHERS / FATHERS

Consent script:

"Good morning/afternoon. Thank you for finding time to meet us. We are here on behalf of Evaluation Support Services, Social Impact, a research company working on a research program into education in Haiti. This is part of the Early Grade Reading Project implemented by USAID. We are trying to understand the situation related to early grade reading in the schools in this community. We are looking for typical mothers/father in this community to tell us their opinions and common conditions in this area related to early grade reading. There are no right or wrong answers. We want you to have the opportunity to express your views freely. We would like to ask your permission to participate in this activity. If you agree, we would hold this open discussion with the group and record the discussion on this device [show recorder. Although the topics we discuss should not be sensitive, we will not ask any of your names, and we will not share this recording with anyone who is not involved in this study. In this way we will protect the confidentiality of your responses. We ask that all participants please respect the privacy of each person here by not talking about who said what in this meeting outside of this room. All participants are free to refuse to answer a particular question if they don't want to. There are no risks to participating in this discussion other than the 1.5 hours of your time that you will give up. There is no direct benefit to you if you participate, other than knowing that you are helping us to assess challenges related to education that may later help Somalia.

We would like to express our appreciation for your time through these modest refreshments.

- **Do you have any questions about this activity?**
- **Do you wish to participate? (record yes or no for each person in the room. Dismiss those who say "no")**

PRE-CODING:

Date: _____ **Community**

Name: _____

District: _____ **Village:** _____

Nearby evaluation school code: _____

Nearby evaluation school name: _____

All female or all male participants? (Circle) Female Male

FGD participant information (do not record names):

Person #	FGD participant's age	Any child enrolled in school (Y/N)	Child(ren) Grade Level(s)	Gender of child(ren) enrolled in school (m/f/both)	Name of School child enrolled in
#1					
#2					
#3					
#4					
#5					
#6					
#7					
#8					
#9					

Time of start of FGD: _____

1. Do most children in your community attend school?
 - a. If no, why not?
 - b. If yes, why?
2. What do you think is the purpose of attending school?
 - a. For a girl child?
 - b. For a boy child?
3. If a family does not send a child to school, what are some of the reasons why?
4. In a typical household in this community do both girl children and boy children attend school?
 - a. If only/mainly one or the other, why?
5. When children are absent from school, what are the reasons why (probing any differences for boys versus girls)?
6. Who in the household makes the decision about who attends school?
7. What are some of the barriers for children to enroll in and attend school? Barriers at home?
8. Is there anything that can be done in the household to make it easier for children to attend school?
9. Do you believe parents should play a role in supporting children's education, or is it the responsibility of schools, teachers, or others?
 - a. If no role for parents, why not?
 - b. If yes to role for parents, what role should parents play? (PROBE: what they do to encourage enrollment? Attendance? Improved learning?)
 - c. To what degree do parents in this community play these roles for their children?

Children's Reading ability & parent roles

10. What are your opinions about the quality of reading training that the children in their school receive?
 - a. How about in comparison to other communities?
 - b. Are there any differences between the in ability to read for boys and girls?
11. What difficulties do you think the schools face in providing quality reading training in this area? Can you explain the reason for these difficulties?
 - a. Do you think the school cares about improving children's reading ability?
 - b. Are their enough trained teachers in the school to improve student's reading ability?
12. As far as you are aware, have there been any recent or current attempts to try and improve reading program in the school?
 - a. If yes, what has been done?
 - b. Do you think this attempts will be successful?
 - c. Why / why not? If so, how?
13. Can you describe the ways in which the parents should be involved to improve children's reading abilities? (PROBE: Do you play any role in helping them reading?)
 - a. If more involvement by father or mther, why are there differences?

- b. If more involvement by mother or father, do you think involvement should be more equal? How could this be done?
- 14. Can you describe the ways in which the parents participate in school management and decision-making in school curriculum particularly in reading program? (PROBE: Are there formal parent committees? Do individual community members play a role?)
 - a. Do women and men participate equally in school management and decision-making? If no, explain the differences.
 - b. If more involvement by men or women, why are there differences?
 - c. If more involvement by men or women, do you think involvement should be more equal? How could this be done?

Time of end of FGD: _____

--	Haiti ERP IE Baseline Director/School Survey	--
Field	Question	Answer
Section I	Section I: Pre-Interview Coding	--
A1	Enumerator Name:	--
A2	What is your enumerator number?	--
A3	Date of Interview:	--
A4	Select wave of data collection:	I Baseline
Department	Select department:	☐ Artibonite ☐ Grand'Anse ☐ Nord ☐ Nord-Est ☐ Ouest ☐ Sud ☐ Sud-Est
Commune	Commune:	☐ ANSE HAINAULT ☐ ANSES A PITRES ☐ ARCAHAIE ☐ Bainet ☐ BEAUMONT ☐ BELLE ANSE ☐ BONBON ☐ Borgne ☐ CABARET ☐ CAMP PERRIN ☐ CAP HAITIEN ☐ CAPOTILLE ☐ Carice ☐ Castel Pere ☐ CARREFOUR ☐ CHAMBELLAN ☐ CHANTAL ☐ CHARDONNIERES ☐ COTEAUX ☐ CROIX DES BOUQUETS ☐ DAME MARIE ☐ DELMAS ☐ DONDON ☐ FERRIER ☐ FORT LIBERTE ☐ GRAND GOSIER ☐ JEREMIE ☐ KENSCOFF ☐ LES ANGLAIS

		☐ LES CAYES ☐ LIMONADE ☐ MANICHE ☐ MARFRANC ☐ MILOT ☐ Mombin-Crochu ☐ Mont-Orgnise ☐ MORON ☐ OUANAMINTHE ☐ PERCHES ☐ PESTEL ☐ PETIONVILLE ☐ PORT AU PRINCE ☐ PORT SALUT ☐ QUARTIER MORIN ☐ SAINT LOUIS ☐ SAINT MARC ☐ SAINTE SUZANE ☐ ST JEAN ☐ TABARRE ☐ TERRIER ROUGE ☐ THIOTTE ☐ TIBURON ☐ TORBECK ☐ TROU DU NORD ☐ VALLIERES
School	Select School Name:	Dynamic list.
A7	Commune sectional	--
A8	Address of school:	--
A10	Select School Type:	☐ 1 Baseline ☐ 1 Public ☐ 2 Private ☐ 3 Religious ☐ 4 Community ☐ 5 NGO ☐ 6 Semi-private ☐ 98 Don't know
A11	Collects GPS coordinates	--
A12	Please note special directions to assist location for follow-up visit:	--
Section 1_end	--	--
Section 2	Section 2: B	--
B2	On which date did school start this year?	--
B3	Were all your teaching positions filled by that date?	☐ 0 No

		☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
B4	Had all the assigned teachers arrived and started working by that date?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
B5	Has school been closed this year at all for reasons other than holiday?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
B6	How many days has school been closed this year so far for reasons other than holiday?	--
B7	What is the main language of instruction?	☐ 1 Creole ☐ 2 French ☐ 3 Creole et French ☐ 99 Don't know
B8	What is the normal entry age for students at this school?	--
B11	What is the normal exit age for students at this school?	☐ 96 N/A ☐ 99 Don't know
B12	Is this school mixed, girls only, or boys only?	☐ 1 Mixed (boys and girls) ☐ 2 Girls only ☐ 3 Boys only
B13	What is your highest level of education completed?	☐ 1 Primary ☐ 2 Secondary ☐ 3 More than secondary ☐ 97 Don't want to respond ☐ 99 Don't know
B14	How many years of experience do you have as a teacher?	--
B15	How many years of experience do you have as a school director?	--
B16	How much years experience do you have working at this school (as a teacher or school director)?	--
B17	How many years experience do you have as the director of this school?	--
Section 2_end	--	--
Section 4	Section 4: D	--
D1	How many classrooms does this school have?	--
D2	What is the level of access to electricity at this school?	☐ 1 Yes, reliably ☐ 2 Yes, occasionally

		☐ 3 Hardly ever ☐ 4 Never
D3	What is the level of access to water?	☐ 1 Yes, reliably ☐ 2 Yes, occasionally ☐ 3 Hardly ever ☐ 4 Never
Section 4_end	--	--
Section 5	Section 5: E	--
E1	How many teachers work at this school full-time? (30 or more hours a week)	--
E2	How many of these full-time teachers are women?	--
E3	How many teachers work at this school part-time? (less than 30 hours a week)	--
E4	How many of there part-time teachers are women?	--
Section 5 E	Please indicate the number of FEMALE teachers who have completed each level of education	--
E5_a	Highest education level completed: more than secondary school	--
E5_b	Highest education level completed: secondary school	--
E5_c	Highest education level completed: primary school	--
E5_d	How many female teachers have completed formal teacher training?	--
E5_end	--	--
E6	Please indicate the number of MALE teachers who have completed each level of education	--
E6_a	Highest education level completed: more than secondary school	--
E6_b	Highest education level completed: secondary school	--
E6_c	Highest education level completed: primary school	--
E6_d	How many male teachers have completed formal teacher training?	--
E6_end	--	--
E7	What is the average number of hours taught per day?	--
E10	How many days of teaching did teachers miss last week?	--
E10_a	How many teachers are absent today?	--
E10_b	How often are teachers tardy (late) for class?	☐ 1 Hardly ever tardy (one or two times a month) ☐ 2 One or two days a week ☐ 3 About half the time ☐ 4 Most of the time ☐ 99 Don't know

E11	How many female teachers at this school have received any type of teacher training (excluding ERP training) in the past year?	--
E12	How many male teachers at this school have received any type of teacher training (excluding ERP training) in the past year?	--
E13	What is the name of the organization that gives the additional teacher training?	--
Section 5_end	--	--
Section 6	Section 6: F	--
F2	Are textbooks usually shared or does each child uses his/her own textbook?	☐ 1 Usually shared ☐ 2 Each child uses their own textbook ☐ 3 There are no textbooks ☐ 99 Don't know
F3	Do you use chalkboards in your school?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F4	Does every classroom have a chalkboard?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F8	Has this school started using any new learning materials such as textbooks, supplementary books, workbooks, or supplies within the past three school years?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F9	Which materials are new?	☐ 1 Math books/workbooks/curriculum ☐ 2 Reading/language books/workbooks/curriculum ☐ 3 Sports related supplies ☐ 4 School supplies (such as stationary, pencils) ☐ 99 Don't know
F10	Where did these materials come from?	☐ 1 International NGO ☐ 2 Haitian NGO or community-based organization ☐ 3 Government ☐ 4 United Nations ☐ 5 Private sector ☐ 6 Religious entity/church ☐ 99 Don't know
F14_a	Does the school management support students with their school fees?	☐ 0 No ☐ 1 Yes

		☐ 96 N/A ☐ 99 Don't know
F15	In the past year, have any teachers from this school received incentives (either financial or other benefits) to encourage them to stay at this school?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F16	Where did these incentives come from?	☐
F20	Is this school currently receiving any other assistance from an outside group in the form of money, construction, programs, or training?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F21	Please describe all of the current assistance being received.	--
F22	Which group(s) is providing this assistance?	☐ 1 International NGO ☐ 2 Haitian NGO or community-based organization ☐ 3 Government ☐ 4 United Nations ☐ 5 Private sector ☐ 6 Religious entity/church ☐ 99 Don't know
F24	Has a member of the MENFP come to the school in the past year to monitor facilities, teaching quality, or teacher attendance?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F25	What did they monitor?	☐ 1 Teacher attendance ☐ 2 Facilities ☐ 3 Teaching quality
F26_a	How many times has a member of the MENFP visited in the last year?	--
F26	Has there been an increase in monitoring visits from the year before?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F27	After the MoE officials leave, do you receive any feedback or reports about these monitoring visits?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F28	What did you receive? Please describe	--
F31	What is the annual fee for enrollment in this school?	--
F31_a	Year (in gourdes)	--
F31_b	Trimester (in gourdes)	--
F31_c	Month (in gourdes)	--
F31_end	--	--

F32	Do families pay for any other thing in order to have their child attend this school?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F33	Please describe all the types of costs and how much they must pay.	--
F34	Are there any students at this school who receive tuition support?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F35	Does the school have a library?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F36	Does the school have computers that students can use or a computer room?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F37	Does this school have a reading club?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F38	Does this school have an active parent-teacher association?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F39	How many times do they meet each school year?	☐
F40	Is this school associated with any extracurricular activities in the summer before the school year starts?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F41	Does the school have after-school extracurricular activities?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F42	Please describe the after-school activities	--
F44	Does this school have a psychosocial support program?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F45	Does this school have a school feeding program?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F46	Does this school benefit from infrastructure renovations?	☐ 0 No ☐ 1 Yes

		☐ 96 N/A ☐ 99 Don't know
F47	Does this school benefit from water, sanitation, and hygiene (WASH) interventions?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F48	Does this school distribute student kits?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F52	Does your school have a school garden?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F50	Does your school have either a school-based nurse, health care worker and/or health specialist?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
F55	Does your school provide any health services to students (e.g. vitamin A supplementation, deworming, immunizations, etc.)?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
Section 6_end	--	--
Section 3	Section 3: C	--
CI	Number of ENROLLED girls in each grade THIS YEAR:	--
CI_a	Grade 1:	☐ 96 N/A ☐ 99 Don't know
CI_b	Grade 2:	☐ 96 N/A ☐ 99 Don't know
CI_c	Grade 3:	☐ 96 N/A ☐ 99 Don't know
CI_d	Grade 4:	☐ 96 N/A ☐ 99 Don't know
CI_e	Grade 5:	☐ 96 N/A ☐ 99 Don't know
CI_f	Grade 6:	☐ 96 N/A ☐ 99 Don't know
CI_g	Grade 7:	☐ 96 N/A ☐ 99 Don't know
CI_h	Grade 8:	☐ 96 N/A ☐ 99 Don't know
CI_i	Total girls enrolled in all grades this year	--
CI_end	--	--

C2	Number of ENROLLED boys in each grade THIS YEAR:	--
C2_a	Grade 1:	☐ 96 N/A ☐ 99 Don't know
C2_b	Grade 2:	☐ 96 N/A ☐ 99 Don't know
C2_c	Grade 3:	☐ 96 N/A ☐ 99 Don't know
C2_d	Grade 4:	☐ 96 N/A ☐ 99 Don't know
C2_e	Grade 5:	☐ 96 N/A ☐ 99 Don't know
C2_f	Grade 6:	☐ 96 N/A ☐ 99 Don't know
C2_g	Grade 7:	☐ 96 N/A ☐ 99 Don't know
C2_h	Grade 8:	☐ 96 N/A ☐ 99 Don't know
C2_i	Total boys enrolled in all grades this year	--
C2_end	--	--
C3	Have there been any dropouts since the beginning of the school year?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
C4	How many girls have dropped out since the beginning of the school year?	--
C5	How many boys have dropped out since the beginning of the school year?	--
C14_end	--	--
--	Thank you very much for your help. We are very grateful for your time and we will use what you have told us carefully.	--
Section 8	Section 8: Observations (H)	--
D8	OBSERVE: Does school have a roof?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
D9	OBSERVE: Does school have dirt or cement floors?	☐ 1 Dirt floors only ☐ 2 Cement floors only ☐ 3 Both dirt and cement floors
H1	OBSERVE: How many toilets are there for girls that are available to be used?	☐ 96 N/A ☐ 99 Don't know
H2	OBSERVE: How many of these girls' toilets are fully covered by walls for privacy?	--

H3	OBSERVE: How many toilets are there for boys that are available to be used?	--
H4	OBSERVE: How many of these boys' toilets are fully covered by walls for privacy?	☐ 96 N/A ☐ 99 Don't know
H5	OBSERVE: How many toilets are there that are available for either gender to enter?	☐ 96 N/A ☐ 99 Don't know
H5a	OBSERVE: Are there separate toilets for girls and boys?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
H6	OBSERVE: Is there a handwashing station WITH WATER in it anywhere on the school grounds?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
H6_exp lain	Why were you not able to observe if there was a handwashing station on the school grounds?	--
H7	Is there a handwashing station near the girls' latrines?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
H8	If yes, Is there a handwashing station near the boys' latrines?	☐ 0 No ☐ 1 Yes ☐ 96 N/A ☐ 99 Don't know
H9	Is there soap next to the handwashong containers?	☐ 1 Yes, soap next to all ☐ 2 Yes, soap next to some ☐ 3 No soap at any handwashing station ☐ 96 N/A
A9	Language of Interview:	☐ 1 Creole ☐ 2 French ☐ 3 Creole et French
Section 8_end	--	--

--	Haiti ERP IE Baseline Class Observation Tool	--
Field	Question	Answer
Protocol	Please follow these directions for conducting the classroom observation:	--
Protocol_1	Explain the exercise to the Head Teacher and obtain his/her consent to do classroom observations.	--
Protocol_2	Ask the Head Teacher or other knowledgeable person to write a list of every teacher that meets the following criterion: a. Teaches grade 1 or 2 b. Is a regular teacher at this school	--
Protocol_3	Out of this full list, randomly select two teachers, one per Grade 1 and one per Grade 2. If possible, stratify by gender, 1 male, 1 female.	--
Protocol_4	Obtain consent directly from teacher.	--
Protocol_5	Divide time into 4 blocks of 10 minutes each. First 10 minutes, ignore and do not record anything; and then start observation with the second 10 minutes.	--
Field	--	--
Part A	Head Teacher consent and teacher selection	--
A3	Please enter your enumerator number:	--
A3_a	What is your name?	--
A4	Date of Observation:	--

--	Haiti ERP IE Baseline Class Observation Tool	--
Commune	Select commune:	☐ Anse hainault ☐ Anses a pitres ☐ Arcahaie ☐ Bainet ☐ Beaumont ☐ Belle anse ☐ Bonbon ☐ Borgne ☐ Cabaret ☐ Camp perrin ☐ Cap haitien ☐ Capotille ☐ Carice ☐ Castel pere ☐ Carrefour ☐ Chambellan ☐ Chantal ☐ Chardonnières ☐ Coteaux ☐ Croix des bouquets ☐ Dame marie ☐ Delmas ☐ Dondon ☐ Ferrier ☐ Fort liberte ☐ Grand gosier ☐ Jeremie ☐ Kenscoff ☐ Les anglais ☐ Les cayes ☐ Limonade ☐ Maniche ☐ Marfranc ☐ Milot ☐ Mombin-crochu ☐ Mont-ornise ☐ Moron ☐ Ouanaminthe ☐ Perches ☐ Pestel ☐ Petionville ☐ Port au prince ☐ Port salut

		☐ Quartier morin ☐ Saint louis ☐ Saint marc ☐ Sainte suzane ☐ St jean ☐ Tabarre ☐ Terrier rouge ☐ Thiotte ☐ Tiburon ☐ Torbeck ☐ Trou du nord ☐ Vallieresa
School	Select school	See list of schools.
A2	Enter school name:	--
A7	Select wave:	I Baseline
A8	Collects gps coordinates	--
Consent_head_teacher	"Now i would like your permission to observe up to two classrooms at this school. A person will sit quietly in the back of the class and take notes without interrupting or talking at all. This will take about 40 minutes. The purpose is only to provide information that helps the erp program in general. It will not be used in an individual manner to evaluate a particular teacher's job performance, and results will not affect in any way the school's ability to receive support now or in the future. The results will remain confidential and will not be discussed with other teachers, the ministry of education, or the community. The teacher's name will remain confidential. Is this ok?"	--
A9	Was consent given by the head teacher?	☐ I Yes ☐ 0 No ☐ 96 Don't know ☐ 99 Not applicable
A10	How many eligible teachers (see criteria in notes above) did the head teacher list for random sampling?	--
Consent_teacher I	"i would like your permission to observe your classroom. A person will sit quietly in the back of the class and take notes without interrupting or talking at all. This will take about 40 minutes. The purpose is only to provide information that helps the erp program in general. It will not be used in an individual manner to evaluate your job performance, and results will not affect in any way the school's ability to receive support now or in the future. The results will remain	--

	confidential and will not be discussed with other teachers, the ministry of education, or the community. Is this ok?"	
B1	Was consent given by teacher 1?	☐ 1 Yes ☐ 0 No ☐ 96 Don't know ☐ 99 Not applicable
Consent	Consent given	--
B5	Teacher name:	--
B6	Teacher gender:	☐ 0 Female ☐ 1 Male
B7	Class grade:	☐ 1 Grade 1 ☐ 2 Grade 2
B8	Number of adults helping in the classroom in addition to the teacher:	--
B9	Does this class have multiple classes in one room?	☐ 1 Yes ☐ 0 No ☐ 96 Don't know ☐ 99 Not applicable
B12	Did teacher take attendance at the beginning of class?	☐ 1 Yes ☐ 0 No ☐ 96 Don't know ☐ 99 Not applicable
--	Repeat block	--
Block_teacher 1	Select which block you are observing:	☐ 2 Block 2 ☐ 3 Block 3 ☐ 4 Block 4
B_a	Teaching approach	--
B13	Students spent time copying from the board.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B14	Students spent time repeating teacher's words aloud.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B15	Teacher used student-centered activities or games.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable

B15a	Asks learners to recognize letters and say letter names and/or sounds	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B15b	Learners retell a story they or the teacher read	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B15c	Asks learners to recite the alphabet	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B15d	Assigns reading for learners to do on their own during school time	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B15e	Provides instructions on how to blend sounds and/or decode syllables and words	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B16	Teacher allowed students to instruct each other (e.g. Come to board to demonstrate something, or explain to classmate).	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B17	Teacher asked open-ended question (requires more than simple answer) that encourages thinking.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B18	Teacher used questions that ask for student opinions (e.g. "what do you think?" "do you like...?").	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B19	Teacher called on or actively tried to involve a student who was not participating.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable

B20	The teacher makes eye contact with all of the students while presenting the lesson.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B21	The teacher gives the students time to respond to questions.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B22	The teacher asked for student responses (i.e. Answers, questions).	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B22a	The teacher demonstrates effective classroom management (e.g., efficiently manages materials, transition(s) between activities, class start and finish)	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B22b	The teacher uses different resources (e.g., books, tools, manipulatives) and strategies (e.g., audio, visual) to explain concepts.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B_b	Gender equity	--
B23	Girls have equal access to desks, learning materials, etc. (e.g. Less amount of sharing of books, desks).	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B24	Boys and girls are mixed together in the seating arrangement.	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable
B25	Teacher provided positive, encouraging feedback to a boys (e.g. "good answer", "nice try", "you're close").	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don't know ☐ 99 Not applicable

B26	Teacher provided positive, encouraging feedback to a girls (e.g. “good answer”, “nice try”, “you’re close”).	☐ 0 Never ☐ 1 Sometimes ☐ 2 All the time ☐ 96 Don’t know ☐ 99 Not applicable
B27	Do boy students ask more questions than girl students?	☐ 1 boys ask more questions ☐ 2 girls ask more questions ☐ 3 girls and boys ask questions about equally ☐ 4 girls and boys do not ask questions ☐ 98 don't know
B29	Does the teacher call on boy students more than girl students?	☐ 1 teacher calls on boys more ☐ 2 teacher calls on girls more ☐ 3 teacher calls on girls and boys equally ☐ 4 teacher does not call on girls or boys ☐ 98 don't know
B31	When called on, girls and boys were given approximately the same amount of time to answer a question (before moving on to the next student).	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B33	Does the teacher use a harsh tone with boys more than girls?	☐ 1 teacher uses harsh tone with boys more ☐ 2 teacher uses harsh tone with girls more ☐ 3 teacher uses harsh tone with girls and boys equally ☐ 4 teacher does not use a harsh tone ☐ 98 don't know
B35	Girls are disciplined physically in class.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B36	Boys are disciplined physically in class.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time

		☐ 96 don't know ☐ 99 not applicable
B37	The teacher asks girls and boys questions of the same difficulty.	☐ 1 the teacher asks more difficult questions to boys ☐ 2 the teacher asks more difficult questions to girls ☐ 3 the teacher asks difficult questions equally to boys and girls ☐ 4 the teacher does not ask difficult questions ☐ 98 don't know
B_c	<u>Curriculum-specific</u>	--
B39	The teacher has the teacher guide.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B40	The teacher holds the guide in his or her hands.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B41	Students use the lesson books.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B43	Students use the writing books.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B45	Students have pencils.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable

B49	The teacher appears comfortable presenting the material.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B50	The teacher reads well.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B51	The majority of the students participate in the lesson.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B52	The majority of students are comfortable giving answers to the questions asked in the lesson.	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
B52a	Learners appear to understand what the teacher is saying	☐ 0 never ☐ 1 sometimes ☐ 2 all the time ☐ 96 don't know ☐ 99 not applicable
--	End of block-repeat questions for next block(s)	--
B_d	Ending questions	--
B53	Can all the students see the blackboard?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
B54	How many of the program posters are hanging in the class?	--
B55	Is the alphabet displayed in the classroom?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable

B56	Can all the students see the teacher?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
B57	Is there enough room for the teacher to move around in the classroom?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
B58	Did the school director come and observe the class?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
B59	Does the teacher have an attendance book?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
B60	Does the teacher have a class journal?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
B_e	--	--
B62	Observer: in your opinion, did the teacher try to include girls and boys equally?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
B63	Explain your answer.	--
B64	Observer: in your opinion, were girls trying to participate as much as boys?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
B65	Explain your answer.	--
B66	Observer: what do you think the teacher did well to help the students learn?	--
B67	Observer: what do you think the teacher could improve to help the students learn?	--

B68	Do you have any comments to enter?	--
Teacher _2	Is there another grade 1 or 2 teacher and classroom to observe?	☐ 1 yes ☐ 0 no ☐ 96 don't know ☐ 99 not applicable
--	Survey repeats	--
End Survey	--	--

--	Haiti ERP IE Student Questionnaire	--
Field	Question	Answers
1	How old are you?	--
2	What language do you speak most often at home?	☐ Creole 1 ☐ French 2 ☐ Other 3
3	When it's time to go to sleep, what does your mother or father or the person responsible for you tells you?	☐ "It's time to sleep"—or other Creole phrases 1 ☐ "It's time for bed"—Or other French Phrases 2 ☐ Other 3 ☐ Don't know/don't want to answer 888
4	In what grade were you in last year? (Do not ask the child if he/she is repeating the grade to verify what he/she says)	☐ No in school last year 999 ☐ Pre-school 0 ☐ 1st Grade 1 ☐ 2nd Grade 2 ☐ 3rd Grade 3 ☐ 4th Grade 4 ☐ Don't know/don't want to answer 888
5	How do you get to school?	☐ You walk by yourself 1 ☐ You walk with your brother/sister 1 ☐ You walk with your friends from class 1 ☐ You walk with the adult from home 1 ☐ You ride your bike to school 1 ☐ You ride your bike with your brother/sister 1 ☐ You ride your bike with your friends from class 1 ☐ You ride your bike with the adult from home 1 ☐ An adult from home drives me in a car/pickup truck/motorcycle 1 ☐ I take the bus/public transportation by myself 1

		☐ Other 1 ☐ Don't 'know /Don't want to answer 888
6	Can I see your creole exercise workbook book?	☐ Creole exercise book is unavailable 0 ☐ Creole exercise book is available 1 ☐ More than one exercise book 2 ☐ Don't know/don't want to answer 888
7	How many pages have correction marks? (Do not count the teachers comments)	☐ None 0 ☐ Some 1 ☐ Most of the pages 2 ☐ All of the pages 3
8	How many teacher comments were in the excise workbook?	☐ None 0 ☐ Some 1 ☐ Most of the pages 2 ☐ All of the pages 3
9	What action does the teacher take when you know the lessons well, perform well on a test/quiz or do your homework well?	☐ Nothing 0 ☐ She congratulates me 1 ☐ She gives me a prize (star/pencil) 2 ☐ She lets me skip a classroom or homework assignment 3 ☐ Other 4
19	What actions does the teacher take when the student cannot respond to a question or gives an incorrect answer?	☐ The teacher repeats/explains the questions 1 ☐ The teacher encourages the students to response 2 ☐ The teacher asks another student 3 ☐ The teacher asks the questions again 4 ☐ The teacher repeats the questions without being frustrated with the student 5 ☐ The teacher becomes frustrated with the student 6 ☐ The teachers kick the student out of the classroom 7

		☐ The teacher beats the student 8 ☐ Other 9
20	Last week, how many times were you assigned homework?	☐ Never was assigned 0 ☐ One time 1 ☐ Two times 2 ☐ Three times 3 ☐ Four times 4 ☐ Everyday 5 ☐ Don't know/don't want to respond 888
21	Does the teacher correct the homework you completed at home last week?	☐ No 0 ☐ Yes 1 ☐ Don't know/Don't want to answer 888
22	Do you have the time to study or complete your homework assignment after school? If not, what other work do you do?	☐ No 0 ☐ Yes 1 ☐ Other 1 ☐ Don't know/don't want to respond 888
23	Who assists you with homework assignments when you are home? (Check all that applies)	☐ No one 1 ☐ Brother/Sister 1 ☐ Mother/Father 1 ☐ Grandmother/grandfather 1 ☐ Friends 1 ☐ Other 1 ☐ Don't know/don't want to respond 888
24	Do you eat in the morning before coming to school?	☐ No 0 ☐ Yes 1 ☐ Don't know/don't want to respond 888
25	Will you eat (have you eaten) food in school today?	☐ No 0 ☐ Yes 1 ☐ Don't know/don't want to respond 888
26	Were you absent from school yesterday? (If yes) Why were you absent from school?	☐ No, I was not absent from school last week 0

		☐ Yes, I was sent home because my tuition wasn't paid 1 ☐ Yes, because I was sick 2 ☐ Yes, because I woke up late 3 ☐ Yes, because I had nothing to eat 4 ☐ Yes, because I went to a funeral 5 ☐ Yes, because I had to go work at the market/the day to prepare for the market 6 ☐ Yes, because I had to take care of my brothers and sisters 7 ☐ Yes, because I had to take care of a family member that was sick at home 8 ☐ Yes, because I had other work to do at home 9 ☐ Yes because I had no transportation means/ my transportation mean 10 ☐ Yes because I had no uniform to wear 11 ☐ Yes because the students and the teachers mistreats me at school 12 ☐ Yes because school is not safe 13 ☐ Yes because school is too challenging 14 ☐ Yes because school is not interesting 15 ☐ Yes because there was bad weather 16 ☐ Other 17 ☐ Don't know/don't want to answer 888
27	When the teacher is absent, is there a substitute teacher that comes and teach in your calls?	☐ No 0 ☐ Yes 1 ☐ Don't know/don't want to answer 888
28	Were you late to school last week? If yes, why were you late to school	☐ No I was not late last week 0 ☐ Yes, because I was sick 1 ☐ Yes, because I woke up late 2

		☐ Yes, because I had to take care of my brothers and sisters 3 ☐ Yes, because I Had to take care of someone in my family that was sick 4 ☐ Yes, because I had other work to do at home 5 ☐ Yes, because I did not have transportation means/my transportation means was late 6 ☐ Yes, because I could not find my uniform, or my uniform was not ready 7 ☐ Yes, because the students and the professors mistreat me at school 8 ☐ Yes, because the weather was bad 9 ☐ Other 10 ☐ Don't know/ don't want to respond 888
29	The last time you received a good grade for a quiz or a test at school, were your parents or your guardian aware?	☐ No 0 ☐ Yes 1 ☐ Don't know/don't want to answer 888
30	If yes, what did they do?	☐ They knew, but they did nothing 1 ☐ They congratulated me/ they encouraged me 2 ☐ They hugged me / gave me kisses 3 ☐ They rewarded me 4 ☐ Other 5 ☐ Don/t know/Don't want to respond 888
31	Did you do preschool or kindergarten?	☐ No 0 ☐ Yes 1 ☐ Don't know/don't want to answer 888
32	Who is responsible for you at home? Is this person literate	☐ No 0 ☐ Yes 1

		☐ Don't know/don't want to answer 888
33	Do you have time to read in class or in the library at school?	☐ No 0 ☐ Yes 1 ☐ Don't know/don't want to answer 888
34	Apart from your school books, do you have books to read at home?	☐ No 0 ☐ Yes 1 ☐ Don't know/don't want to answer 888
37	Does someone at home reads you stories? If yes, how often does this happen?	☐ Never 0 ☐ Sometimes 1 ☐ Once a week 2 ☐ 2-3 times a week 3 ☐ Every day 4 ☐ Don't know/Don't want to respond 888
38	(If yes, for the previous answer) In what language do they normally read to you in?	☐ Creole 1 ☐ French 2 ☐ Other (indicate) 3 ☐ Don't know/ don't want to respond 888
--	In the house you live in, is there ...?	--
39	Radio	☐ I yes ☐ 0 no
40	Television	☐ I yes ☐ 0 no
41	Bicycle	☐ I yes ☐ 0 no
42	Car	☐ I yes ☐ 0 no
43	Watch	☐ I yes ☐ 0 no
44	Motorcycle	☐ I yes ☐ 0 no

45	Transport Cart	☐ I yes ☐ 0 no
46	Car/van	☐ I yes ☐ 0 no
47	Bat	☐ I yes ☐ 0 no
48	Electricity	☐ I yes ☐ 0 no
49	Computer	☐ I yes ☐ 0 no
50	Indoor Kitchen	☐ I yes ☐ 0 no
--	What do you use to go to the bathroom at home?	--
51	Latrine	☐ I yes ☐ 0 no
52	A toilet that can flush	☐ I yes ☐ 0 no
53	A Bucket	☐ I yes ☐ 0 no
54	Outdoor bathroom	☐ I yes ☐ 0 no
55	Other	☐ I yes ☐ 0 no
56	Don't know/don't want to answer	☐ I yes ☐ 0 no
--	What do you use at home to cook food?	--
57	Fire wood to prepare food	☐ I yes ☐ 0 no
58	Outdoor grill (with wood or charcoal)	☐ I yes ☐ 0 no
59	Electric stove or gas stove	☐ I yes

		☐ 0 no
60	Other	☐ 1 yes ☐ 0 no
61	Don't know/Don't want to answer	☐ 1 yes ☐ 0 no
--	What is the source of the water you drink at home?	--
62	River	☐ 1 yes ☐ 0 no
63	Reservoir	☐ 1 yes ☐ 0 no
64	Faucet outside the house	☐ 1 yes ☐ 0 no
65	Water truck	☐ 1 yes ☐ 0 no
66	Rain	☐ 1 yes ☐ 0 no
67	Rain water	☐ 1 yes ☐ 0 no
68	Merchants of water bags	☐ 1 yes ☐ 0 no
69	Bottled water	☐ 1 yes ☐ 0 no
70	Company that sells water	☐ 1 yes ☐ 0 no
71	Other	☐ 1 yes ☐ 0 no
77	Times up	--

EARLY GRADE READING ASSESSMENT TOOL

Early Reading Program (ERP) Baseline Impact Evaluation

2017

(Tool was administered in Creole only)

Introduction and Verbal Consent

CONSENTEMENT VERBAL POUR LES ELEVES :

Bonjou. Koman ou ye? Mwen rele _____. Mwen ta renmen pale ou de mwen. [di kèk bagay sou laj pitit ou, spò w renmen fè, etc.]

1. Eske ou kapab pale m de ou oubyen de fanmi w? [tann yon repons. Si elèv la retisan, mande kèsyon 2, men si yo alèz, ou met komanse ak konsantman vèbal la]

2. Kisa ou renmen fè lè w pa lekòl ?

Kite m diw poukisa mwen la jodi a. M ap travay pou Ministè Ledikasyon Nasyonal epi n ap eseye konprann kouman timoun aprann li. Mwen ta renmenw ede m nan sa.
Ou te chwazi konsa pou patisipe nan travay n ap fè a.
Avèk ti aparèy sa a, mwen pral wè konbyen tan ou pran pou li.
Se pa yon egzamen, li pa gen anyen pou li wè ak nòt lekòl ou.
Si ou pa vle patisipe, fè m konnen.
Mwen gen pou poze w kèk lòt kesyon sou fanmi w, tankou ki lang yo pale lakay ou ak kèk lòt bagay sou lakay ou tou.
Mwen pap pran non w pou pèsan pa konnen se ou menm ki ban m repons sa yo.
Yon fwa nou kòmanse, si ou pa vle reponn yon kesyon, ou gen dwa pa reponn li.
Eske w gen kesyon ?
Ou dakò pou kòmanse ? Ou pare pou kòmanse?

☐ Mwen te resevwa konsantmanan epi e elèv la dakò pou kòmanse: **wi**

(Si elèv la pa bannou konsantman l – li pa dakò pou patisipe, di l mèsì epi ale jwenn yon lòt elèv).

1. Date du test :	Date : _____ Mois : _____ Année : _____	7. Fonctionnement de l'école	☐ Matin ☐ Après-midi ☐ Les deux
2. Nom de l'enquêteur :		8. Code unique de l'élève :	
3. Nom de l'école :		9. Année d'études de l'élève :	☐ 1 ^e Année

		10. Classe (Section):	
4. Département	☐ Nord ☐ Nord-Est ☐ Sud-Est	11. Date de naissance de l'élève	Mois : _____ Année : _____
5. Commune de localisation		12. Genre de l'élève	☐ fille ☐ garçon
6. Section communale		13. Heure début du test	Heure : _____ Minute : _____
		14. Infirmité visuelle ou auditive de l'élève	☐ Aucune ☐ Infirmité visuelle ☐ Infirmité auditive

K-Seksyon 1 : Konesans <u>non</u> lèt yo	K-Seksyon I	60 segonn
	Si elèv la pa rive bay okenn bon repons nan premye ran an.	Si elèv la pa rive reponn aprè 3 segond, di l «kontinye» pandan w ap montre lòt lèt la.

- ☞ (/) Bare chak ekriti kote elèv la bay move repons.
(O) Fè yon ronn sou ekriti kote elèv la te korije limenm.
(]) Mete yon] aprè dènye ekriti elèv la li

Men yon ti paj ki gen lèt kreyòl. Gade twa egzanp sa yo. Fwa sa a, w ap di mwen non yo. Pa egzanp : [indiquez le « m » dans la ligne des exemples] **Non premyè lèt la se « m ».**

Ann eseye kèk lòt egzanp : Ki lèt sa a? [indiquez le « t » dans le rang des exemples]

[Si l'élève répond correctement, dites-lui :] «**Trè byen ! Non lèt sa a se « t ».**

[Si l'élève ne répond pas correctement, dites-lui :] «**Non lèt la sa a se « t ».**

Ann eseye yon lòt egzanp : Ki lèt sa a? [indiquez le « ou » dans le rang des exemples]

[Si l'élève répond correctement, dites-lui :] «**Trè byen ! Non lèt sa a se « ou ».**

[Si l'élève ne répond pas correctement, dites-lui :] «**Non lèt la a se « ou ».**

Mete dwèt ou anba chak lèt kreyòl w ap li [montrez avec votre doigt]. **Li lèt kreyòl yo liy pa liy** [montrez avec votre doigt]. **Kou m di « Kòmanse », w a li yo byen e vit. Dakò ? Touche premye lèt la. Ou pare ? An nou « Kòmanse ».**

Exemple: m t ou

	1	2	3	4	5	6	7	8	9	10		
--	---	---	---	---	---	---	---	---	---	----	--	--

	d	è	r	E	S	M	s	V	A	N	(10)	
	t	d	T	ch	Z	v	G	N	i	F	(20)	
	an	o	i	ch	f	p	f	L	è	i	(30)	
	s	an	ou	an	en	P	B	k	s	i	(40)	
	b	a	e	ch	b	p	j	en	a	M	(50)	
	i	K	n	A	w	g	n	m	b	L	(60)	
	t	v	t	s	v	W	ou	on	D	p	(70)	
	R	o	z	m	b	ou	n	O	ò	A	(80)	
	J	on	g	O	r	L	F	d	Y	E	(90)	
	i	L	k	on	è	A	r	e	en	e	(100)	

Kantite segond ki rete nan kronomèt la:	
Make nan ti kare a si ekzèsis la te rete poutèt pat gen bon repons nan premye ran an:	

Anfòm! Ou pare pou w fè pwochen aktivite a. Tre byen!

K-Seksyon 2 : Konsyans fonemik	 x	 x
	 Si elèv la pa rive bay okenn bon repons pou 5 premye mo yo.	 Si elèv la pa rive reponn aprè 3 segond, make kaz « pa gen repons ».

(✓) Korèk / pa korèk / pa gen repons ditou

Egzèsis sa a se yon egzèsis pou tande. Mwen pral di w yon mo de fwa epi mwen ta renmen ou di m premye son ou tande nan mo a. Dakò ?

Pa egzanp : mo « soup » la kòmanse ak son « sssssss ».

Ann eseye kèk lòt egzanp : Ki premye son ou tande nan mo « chou » ? « chou » ?

[Si l'élève répond correctement, dites-lui :] « **Trè byen ! Premye son nan mo « chou » se « ch ».**

[Si l'élève ne répond pas correctement, dites-lui :] « **Premye son nan mo « chou » se « ch ».**

Ki premye son ou tande nan mo « gato » ? « gato » ?

[Si l'élève répond correctement, dites-lui :] « **Trè byen ! Premye son nan mo « gato » se « g' ».**

[Si l'élève ne répond pas correctement, dites-lui :] « **Premye son nan mo « gato » se « g' ».**

Ou konprann sa m' mande w pou fè a ? Kounye a mwen pral li kèk lòt mo. M ap li chak mo de fwa. Koute byen epi di m ki premye son ou tande nan mo a. Dakò ?

	Ki premye son ou tande nan mo « ____ » ? « ____ » ?	[Li chak mo 2 fwa]	--	--
tas	/t'/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
lòt	/l l l l l/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
jwèt	/j j j j j/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
poul	/p'/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
bèl	/b'/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
vant	/v v v v v/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
dat	/d'/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
fig	/f f f f f f/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
kay	/k'/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
sik	/s s s s s s/	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
--	--	--	Make nan ti kare a si ekzèsis la te rete poutèt elèv la pat bay okenn bon repons nan 5 premye mo yo:	--

Anfòm! Ou pare pou w fè pwochen aktivite a. Trè byen!

K-Seksyon 3 : Konesans <u>son</u> lèt yo	 K-Seksyon 3	 60 segonn
	 Si elèv la pa rive bay okenn bon repons pou 10 premye lèt yo.	 Si elèv la pa rive reponn aprè 3 segond, di l «kontinye» pandan w ap montre lòt lèt la.

- ☒ (/) Bare chak ekriti kote elèv la bay move repons.
 (O) Fè yon ron sou ekriti kote elèv la te korije limenm.
 (]) Mete yon] aprè dènye ekriti elèv la li.

Men yon paj ki gen lèt kreyòl. Gade twa egzanp sa yo. Fwa sa a, w ap di mwen son lèt yo. Pa egzanp : [indiquez le « M » dans la ligne des exemples] **Son premye lèt la a se « mmmmm ».**

Ann eseye kèk lòt egzanp : Ki son lèt sa a ? [indiquez le « t » dans le rang des exemples]

[Si l'élève répond correctement, dites-lui :] «**Trè byen ! Son lèt sa a se « t' ».**

[Si l'élève ne répond pas correctement, dites-lui :] « **Son lèt la a se « t' ».**

Ann eseye yon lòt egzanp : Ki son lèt sa a ye ? [indiquez le « ou » dans le rang des exemples]

[Si l'élève répond correctement, dites-lui :] «**Trè byen ! Son lèt sa a se « ou ».**

[Si l'élève ne répond pas correctement, dites-lui :] « **Son lèt la a se « ou ».**

Mete dwèt ou anba chak lèt kreyòl w ap bay son yo [montrez avec votre doigt]. **Di son lèt kreyol yo liy pa liy** [montrez avec votre doigt] **Kou m di « Kòmanse », wa di yo byen e vit. Dakò ? Touche premye lèt la. Ou pare? Ann « Kòmanse ».**

Exemple: m t ou

	1	2	3	4	5	6	7	8	9	10		
	e	A	t	N	g	n	j	L	E	en	(10)	
	t	N	ch	e	p	L	T	s	A	d	(20)	
	i	ch	w	S	b	e	f	b	O	en	(30)	
	B	J	b	E	K	G	r	k	an	è	(40)	

	r	M	R	z	F	ou	f	P	o	s	(50)	
	i	ou	v	ou	m	L	on	A	è	ch	(60)	
	v	W	en	v	è	Z	s	d	i	b	(70)	
	a	an	L	d	m	O	i	s	p	k	(80)	
	P	r	M	n	ò	t	A	an	i	Y	(90)	
	o	V	on	e	g	i	D	n	F	on	(100)	
Kantite segonn ki rete nan kwonomèt la:												
Make nan ti kare a si ekzèsis la te rete poutèt pat gen bon repons nan premye liy lan:												

Anfòm! Ou pare pou w fè pwochen aktivite a. Trè byen!

K-Seksyon 4 : Lekti mo timoun yo konnen	 K-Seksyon 4	 60 segonn
	 Si elèv la pa rive bay okenn bon repons nan premye liv lan.	 Si elèv la pa rive reponn aprè 3 segonn, di l « kontinye » pandan w ap montre lòt mo a.

- ☒ (/) Bare chak mo ekriti kote elèv la bay move repons.
☐ (O) Fè yon wonn sou ekriti kote elèv la te korije limenm.
☐ (J) Mete yon J aprè dènye mo elèv la li a.

Men yon paj ki gen mo kreyòl. Gade twa egzanp sa yo. Pa egzanp : [indiquez le mot « mi » avec le doigt] **Premyè mo a se « mi ».**

Ann eseye kèk lòt egzanp : Li mo sa pou mwen. [indiquer le mot « do » dans le rang des exemples]

[Si l'élève répond correctement, dites-lui :] « **Trè byen ! Mo sa a se « do ».**

[Si l'élève ne répond pas correctement, dites-lui :] « **Mo sa a se « do ».**

Ann eseye yon lòt egzanp : Li mo sa pou mwen. [indiquer le mot « twa » dans le rang des exemples]

[Si l'élève répond correctement, dites-lui :] « **Trè byen ! Mo sa a se « twa ».**

[Si l'élève ne répond pas correctement, dites-lui :] « **Mo sa a se « twa ».**

Mete dwèt ou anba chak mo kreyòl w ap li [montrez avec votre doigt]. **Li mo kreyòl yo liy pa liy** [montrez avec votre doigt]. **Kou m di « Kòmanse », wa li yo byen e vit. Dakò ? Touche premye mo a. Ou pare ? Ou pare ? Ann « Kòmanse ».**

Exemple : **mi** **do** **twa**

1	2	3	4	5	
La	pòt	blan	tab	lou	(5)
Moto	mè	flè	jape	lalin	(10)
Pil	pri	rele	epi	yon	(15)
Bòl	vid	wout	papa	ap	(20)
Mwen	fè	sou	Won	chèz	(25)
sa	woz	po	manje	di	(30)

Tou	ti	rat	chat	bwè	(35)
Pen	ble	poul	bon	leswa	(40)
Dat	zanmi	douz	wouj	li	(45)
Bèl	mal	pon	saj	mi	(50)
Kantite segonn ki rete nan kwonomèt la:					
Make nan ti kare a si ekzèsis la te rete poutèt pat gen bon repons nan premye liy lan:					

Anfòm! Ou pare pou w fè pwochen aktivite a. Trè byen!

K-Seksyon 5a : Lekti tèks	 K-Seksyon 3a	 60 segonn
	 Si elèv la pa rive bay okenn bon repons pou 8 premye mo yo.	 Si elèv la pa rive reponn aprè 3 segonn, di l «kontinye» pandan wap montre lòt mo a.

- ☒ (/) Bare chak ekriti kote elèv la bay move repons.
☐ (O) Fè yon ron sou ekriti kote elèv la te korije limenm.
☐ (]) Mete yon] aprè dènye ekriti elèv la li.

Li istwa sa a ki ekri an kreyòl. Li li byen fò, li li vit. Aprè sa, ou pral reponn kèk kesyon sou istwa a. Kòmanse la. [Montrez avec le doigt le premier mot de l'histoire.] **Ou pare ? Ou mèt « Kòmanse ».**

Nan lekòl la, timoun yo renmen jwe foutbòl.	9
Lè konsa, yo fè de kan nan lakou a.	17
Jozèf ak Mak toujou jwe nan menm ekip.	21
Jedi pase, lòt ekip la bat yo twa gòl a zewo. Mak te fache pou sa. Jozèf te bat do Mak,	42
epi li te di Mak : "pa fache, ekip nou an konn genyen tou". Mak dakò. Li ri.	60

Kantite segonn ki rete nan kwonòmèt la:

Make nan ti kare a si ekzèsis la te rete poutèt pat gen bon repons nan premye liy lan:

K-Seksyon 5b : Konpreyansyon tèks ki li a	 x	 x
	 x	 Si elèv la pa rive reponn aprè 10 segonn

Lè elèv la fin li (seksyon 3a), retire tèks la nan menl, poze l kesyon ki anba yo. Poze kesyon ki rive jis kote li te rete nan tèks la.

Pran nòt de tout repons elèv la nan espas ki rezève pou sa. Mete yon « X » nan kaz ki koresponn a chak kesyon.

(✓) Korèk / pa korèk / pa gen repons ditou

Kounye a, ou pral reponn kèk kesyon sou istwa a. Ou mèt reponn an kreyòl oubyen an fransè.

KESYON	REPONS ELÈV LA		
	Korèk	Pa Korèk	Pa gen repons
1. Ki jwèt timoun yo renmen jwe? [foutbòl]			
2. Ki kote yo jwe? [nan lakou a]			
3. Ak kiyès Mak renmen jwe? [ak Josèf]			
4. Pou kisa Jozèf te bat do Mak? [paske Mak te fachelè li te pèdi a]			
5. Kisa ki te fè Mak ri? [paske li panse li ka toujou genyen ankò, paske Josèf te di l ekip nou an konn genyen tou. Pawòl Jozèf di a.]			

Anfòm! Ou pare pou w fè pwochen aktivite a? Trè byen!

PAGE LEFT INTENTIONALLY BLANK

K-Seksyon 6: Konpreyansyon nan koute	 x	 x
	 x	 Si elèv la pa rive reponn aprè 10 segonn

 (✓) Bon repons / pa bon / ou pa gen repons ditou

[Ou gen pou li byen fò yon ti istwa pandan 2 fwa, aprè wa mande elèv la kèk kesyon sou li.]

[Elèv la ka bay repons yo nan yonn ou lòt lang (kreyòl ou byen franse).]

Di elèv la: **Kounye a, mwen pral li yon istwa pou ou de fwa. Aprè sa, ou pral reponn kèk kesyon sou istwa a. Koute lekti a byen.** [Lisez l'histoire deux fois.]

Dakò ? M ap koumanse.

Toto ap sòt lekòl. Lapli a kòmanse tonbe. Toto ap kouri pou lapli. Soulye li sòti nan pye li. Toto pa wè soulye a. Li tonbe kriye. Lè lapli a pase, Toto mache pye atè. Yon machann wè soulye a. Li di : « Toto men soulye w la nan kannal la. » Toto pran soulye a. Li lave li. Li kontan.

[M ap li l yon lòt fwa.]

Kounye a reponn kesyon sa yo. Ou mèt reponn an kreyòl oubyen an franse.

KESYON	Repons Korèk (PA LI POU ELEV LA)	REPONS ELÈV LA		
1. Kijan tan an te ye lè Toto t ap sot lekòl la?	Lapli t ap tonbe	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
2. Ki sa l te pèdi lè l t ap kouri a ?	soulye li	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
3. Ki sa li te fè lè l pa t wè soulye a?	Li te tonbe kriye	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons

4. Kijan Toto te rive jwenn soulye a?	Yon machann ki te montre l li nan kannal la.	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
5. Pou ki sa li lave soulye a?	Paske li te sal, li te gen labou ladann, pou l mete l nan pye l	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons

Anfòm! Ou pare pou w fè pwochen aktivite a. Trè byen!

K-Seksyon 7 : Dikte	 Fèy papye ak kreyon	 x
	 x	 x

(✓) Korèk / pa korèk / pa gen repons ditou

A. Lèt

Bay timoun nan yon kreyon ak yon fèy papye. Pa kite l gade lèt yo. Si timoun nan di : « Mwen pa konnen, » make repons sa a kòm pa gen repons/Li pa ekri anyen.

Mwen pra l di w kèk lèt. Se pou koute m avèk atansyon. Apre chak lèt mwen fin di w, m ap repete l yon lòt fwa pou ou, e w ap ekri lèt ou tande a sou papye a pou mwen. Eske w konprann sa m mande w fè a ? Oke, koute epi ann kòmanse.

[Bay non chak let 2 fwa]			
i	☐ Kòrèk	☐ Pa Korèk	☐ Pa gen repons
b	☐ Kòrèk	☐ Pa Korèk	☐ Pa gen repons
m	☐ Kòrèk	☐ Pa Korèk	☐ Pa gen repons
a	☐ Kòrèk	☐ Pa Korèk	☐ Pa gen repons
L	☐ Kòrèk	☐ Pa Korèk	☐ Pa gen repons

Kounye a, mwen pra l di w kèk mo. Koute m avèk atansyon. Apre chak mo mwen fin di w, m ap repete l yon lòt fwa pou ou, e w ap ekri mo ou tande a sou papye a pou mwen. Eske w konprann sa m mande w fè a ? Oke, koute epi ann kòmanse.

B. Mo

[Li chak mo 2 fwa]

pay	☐ Kòrèk	☐ Pa Korèk	☐ Pa gen repons
fil	☐ Kòrèk	☐ Pa Korèk	☐ Pa gen repons
ten	☐ Kòrèk	☐ Pa Korèk	☐ Pa gen repons

Anfòm! Ou pare pou w fè pwochen aktivite a. Trè byen!

LANGUE FRANÇAISE

Seksyon 1 : Konpreyansyon nan koute	📖 x	🕒 x
	✋ x	🔄 Si elèv la pa rive reponn aprè 10 segonn.

(✓) Korèk / pa korèk / pa gen repons ditou

[Ou gen pou li byen fò yon fraz pandan 2 fwa, aprè wa mande elèv la yon kesyon sou li. Ou swiv etap sa yo pou chak fraz.]

[Elèv la ka bay repons yo nan youn ou lòt lang (kreyòl ou franse).]

[Di elèv la:] **Kounye a, mwen pral li kèk fraz pou ou. M ap li l de fwa. Aprè sa, ou pral reponn yon kesyon sou fraz la. Koute byen. Ou mèt reponn an kreyòl ou an franse. Dakò ? M ap kòmanse.** [Lisez chaque phrase deux fois, puis poser la question y relative.]

Kounye a reponn kesyon sa yo. Ou mèt reponn an kreyòl ou an franse.

Fraz	Kèsyon	Repons (PA LI POU ELEV LA)	REPONS ELÈV LA		
Mario va à la mer à cheval.	Où va Mario ?	à la mer	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
Le petit garçon a mal à la tête.	Qui a mal à la tête ?	Le petit garçon	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons
Valérie se lave les mains avec du savon.	Que fait Valérie avec le savon ?	Lave les mains	☐ Korèk	☐ Pa Korèk	☐ Pa gen repons

Anfòm! Ou pare pou w fè pwochen aktivite a ? Trè byen!

ANNEX C: EVALUATION METHODOLOGY

FINDINGS FROM DESK REVIEW

Evaluation Question	Preliminary Findings	Data Gaps
1) What is the impact, if any, of the ERP intervention on students' reading outcomes in the first and second grades?	None	No access to ERP curriculum or teaching materials
a. How do students exposed to any ERP programming perform when compared to similar students who are not exposed to ERP programming?	Early analyses focus on correct words per minute and automaticity.	The absence of a definitive EMIS system to help in the selection of a control/ comparison group.
b) What is the marginal effect, if any, of ERP implementer on reading outcomes?	None	
c) What is the marginal effect, if any, of type of school (semi-private, public, community, religious, etc.) on reading outcomes?	There are varied results on the effects of public/private schools on reading scores.	The ET lacks definitive lists of schools and their types (public, private, etc.). Data on this can be gathered through data collection.
d) To what extent do environmental or situational factors have an impact on students' ability to learn to read in the ERP target schools?	School management tools exist and can be adapted to the context.	
2) What is the impact, if any, on student learning, of the “enhanced” and “enhanced plus” models carried out by UND? Do students exposed to this model acquire stronger reading skills than their peers in control and other ERP schools? What aspects of the model impact student learning the most?	None	The ET was unable to obtain the curriculum nor in-depth descriptions of the enhanced model's interventions
3) What is the impact, if any, on student learning, of the “pre-reading instruction” model piloted by UNICEF in the pre-K and K levels of a subset of the UNICEF ERP program schools? Do students exposed to the ‘pre-reading instruction’ model acquire stronger reading skills than their peers in control and other ERP schools?	None	Empirical evidence of evaluations of preschool interventions and preschool interventions specifically in Haiti. The list of schools with preschool interventions is yet to be finalized.

Evaluation Question	Preliminary Findings	Data Gaps
4) What is the impact of the ERP intervention, if any, on teachers' classroom practices in early grade reading instruction?	The EGRA tool is well supplemented by objective, quantitative classroom observation data. Template questions exist in various tools and in the EGRA Toolkit.	The ET does not have the ERP curriculum and teaching manuals. Without these, intervention-specific questions cannot be drafted and tools are not as effective not specific as they could be.
a) Do teachers exposed to the 'enhanced' ERP intervention teach reading differently and/or better than their peers in the 'core' ERP schools?	Qualitative data shows that teaching quality and the correct implementation of curriculum by teachers matters.	
b) Do teachers exposed to the ERP intervention teach reading differently and/or better than their peers in the identical or close to identical schools without ERP interventions?	None.	
c) How, if at all, do teachers exposed to the ERP intervention teach reading differently after having received the ERP intervention?	Teacher training is important, but may not be sustained without refresher trainings and ongoing coaching.	
5) What is the impact, if any, of the ERP program on the Ministry's capacity to implement and eventually scale up early grade reading reform? Are the human, material, and financial management capacities at the MENFP sufficiently improved because of the ERP intervention that the Ministry could successfully expand the ERP program nationwide?	Historically, public sector actors may have not played as large a role as they should. Involving these actors is critical for validation of tools and buy-in for the evaluation And use of the results for program remediation. based on data	No data or documents received to date on the role of the MENFP in the ERP program.

SCHOOL SELECTION CRITERIA FOR ERP, BY IMPLEMENTER

Save the Children	AIR	UNICEF	UND
• School is near student's homes and in a safe area • If private, has a valid title and lease for next 5 years • Has at least 100 students who can be enrolled in a scholarship program • A teacher to student ratio that is not higher than 40 • At least one Grade 1 and Grade 2 class and teacher • School works on a double shift • School has latrines and handwashing facilities • School director has a degree from a teaching college and at least 5 years experience • School is willing to exempt up to 10% of children from school fees • School will open its doors to the community at large • School has a code of conduct for teachers • Management promotes respect for rights of the children • Teachers have at least a BACII and 3 years experience • Teachers are regularly present and punctual • School is willing to practice alternatives to corporal punishment	• Has teachers trained from an MENFP teachers college • Has a USAID or other NGO non-education intervention, such as a school feeding program • Is easy to access and can distribute educational materials easily • Students have an adequate learning environment • School promotes preventive health education	• Public school • Includes Grades 1 and 2 • Has "adequate" infrastructure • Has at least 200 students enrolled • Has sufficient well-trained and well-paid teachers • Benefits from other reading-focused programs • Part of the EFACAP network • Is a legally accredited school	• School is "in need of strengthening" • School meets a basic level of functionality • School has minimally competent staff (e.g. teachers who can read) • School staff are interested in participating and commit to the program requirements • At least 10 students enrolled each per Grade 1 and Grade 2

CEM APPROACH

A CEM approach has several advantages over other quasi-experimental approaches.³³ In recent literature, CEM has been shown to be more effective than other common approaches, including propensity score matching (PSM), at reducing model dependence, imbalance, and bias. By using the CEM approach, the evaluation team temporarily ‘coarsen’ the observed variables used for matching into distinct bins/strata. Every combination of coarsened values, for all variables included in the matching, represents a unique stratum, and observations with the same coarsened value on every included variable is placed into the same strata.

Successfully executing CEM, however, requires several additional tasks in terms of sampling. The matching process requires robust school-level data to permit matching of treatment and control schools, assuming enough similar schools exist in a given area. Like many developing countries, Haiti lacks school Census data which can be used as a sample frame for the matching. In its absence, the Haiti ESS team decided to use a partial listing of schools based on the school list available with the four implementing partners. After listing eligible schools, the evaluation team worked to identify a set of selective school-level questions upon which to base the identification of control schools through CEM.

One of the key features of the ERP evaluation is having two treatment arms— core and enhanced, based on the types of intervention adopted by implementing partners. With two treatment arms, the process of creating a “triplet” is easier when the exact treatments are not pre-specified or not already allotted before matching. In that case, matched treatment and control groups can be identified first and then the treatment schools can be allocated to two treatment arms. In a World Bank Study, Wagstaff et al (2014) used a similar approach where they first created two matching groups of communes using CEM, and then randomly assigned matched treatment communes within each stratum to two different treatment arms. The evaluation of ERP is somewhat different in other fronts as well as compared to standard CEM with one treatment and one control group. Because there are two distinctly separate types of intervention planned by four implementing partners across departments, the evaluation requires two separate arms. Moreover, treatment schools for two interventions were already identified based on different criteria at baseline. As a result, the use of CEM with two treatment arms becomes tricky since it needs a common control group which is almost identical to both treatment groups (this is known as a “triplet”). Due to the pre-selection of treatment schools this approach was not feasible.

To overcome these difficulties, ET (i) used basic school level characteristics (such as teachers’ qualifications and experience, access to infrastructure, school management indicators, and student-teachers ratio) that are easily available across control and two treatment groups and (ii) matched two different treatment arms separately with control schools. ET then combined (merged) these matched groups over common control schools so that each control school would correspond to both types treatment schools in each stratum. The team discarded the control schools that matched with one treatment group but did not match with the other.

See King, G., Blackwell, M., Iacus, S., & Porro, G. “Coarsened exact matching in Stata.” (2010); Rosenbaum, PR, and DB Rubin. “The central role of the propensity score in observational studies for causal effects.” *Biometrika* 70, no. 1 (1983): 41-55.

Variables used for CEM matching
School Type: Public School
What is the normal exit age for students at this school?
How many years of experience do you have as the director of this school?
Proportion of trained male teachers more than 25 percent
Proportion of trained female teachers more than 25 percent
Student teachers ratio
Proportion of part time teachers
Student shares text book
Have there been any dropouts since the beginning of the school year?
Language of instruction
Does your school provide any health services to students
School infrastructure: chalkboard
Location: rural- urban

In the process CEM created 4 strata within which all these treatment and control schools were a distributed. The control schools in each stratum correspond to their exact matches from each treatment group.

To ensure that control schools are exact matches to the treatment schools, balancing tests was performed across a set of variables measured at baseline to use for matching.

SHAPLEY DECOMPOSITION

Shapley decomposition, as proposed by Shorrocks (2013), considers the marginal effects of each contributory factors on some output/outcome by eliminating each of them in sequence, and then assigning the average of the respective contributor's marginal contributions to each factor in all possible elimination sequences. This procedure yields an exact additive decomposition of the 'output' into multiple contributions. According to Shorrocks (2013), the procedure can be employed in all areas of applied economics whenever one wishes to assess the relative importance of the explanatory variables Shorrocks (2013).

To find out the relative contributions, or the share of how each driver (groups of school, student, or classroom level characteristic) contributes to overall ORF score, we grouped similar factors (drivers) into several like groups. The ET used a Shapley decomposition method to isolate relative shares of each group of drivers created. For example, Group 1 includes factors related to parental involvement such as parents making sure that students eat before coming to school, if they are aware when the student receives a good grade, or whether the student attended Pre-K/Kindergarten. Similarly, Group 2 includes information about a student's situation related to reading, for example whether the student has time to do homework or study lessons after school, has time to read in class or in a school library, has books other than school books to read at home, or has a Creole exercise book (or books) in school.

Group 3 and 4 are based on drivers associated with classroom observation that capture a teacher's teaching style. Group 3 combines teaching methods, including if a teacher corrects the homework, asks students to spend time in class repeating words aloud, asks learners to recognize letters and say letter names and/or sounds, provides instructions on how to blend sounds and/or decode syllables and words,

encourages students to use the lesson books, and assigns reading for learners to do on their own during school time. Group 4 captures gender equitable teaching practices by combining the drivers of calls on or actively tried to involve a student who was not participating (irrespective of gender), mixes boys and girls together in the seating arrangement, and disciplines boys or girls physically in class.

To measure training and experience of the principal and teachers, for Group 5 the team put together a principal's years of experience as a school director and his or her years of experience working at the school in question as a teacher. Group 5 also contains other male and female teachers' experience at a school and status of formal training in teaching. Group 6 encompasses all school-level characteristics, such as student to teacher ratio, average numbers of hours taught per day, existence of library, school management's support to students with their school fees, provision of after-school extracurricular activities, existence of a feeding program, active reading club, parent-teacher association, and provision of student kits. While each of these initial six groups has their respective policy implications, the remaining two groups are created to identify how different types of schools (such as private, public, etc.) contributes to student's reading fluency. We also included a final group for implementers as a proxy for factors that are used to identify treatment schools but are not known to the ET. These unobservable characteristics could include socio-economic and geographic factors. Based on this grouping of explanatory variables, the decomposition analysis is done for all students together to find the overall contribution of the various drivers to ORF, and then separately for boys and girls to identify if there is any differential effect on gender.

Group	Theme	Components
1	Parental involvement	Parents making sure that students eat before coming to school, if they are aware when the student receives a good grade, or whether the student attended Pre-K/Kindergarten
2	Student reading	Whether the student has time to do homework or study lessons after school, has time to read in class or in a school library, has books other than school books to read at home, or has a Creole exercise book (or books) in school.
3	Classroom teaching	if a teacher corrects the homework, asks students to spend time in class repeating words aloud, asks learners to recognize letters and say letter names and/or sounds, provides instructions on how to blend sounds and/or decode syllables and words, encourages students to use the lesson books, and assigns reading for learners to do on their own during school time.
4	Gender equity in class	calls on or actively tried to involve a student who was not participating (irrespective of gender), mixes boys and girls together in the seating arrangement, and disciplines boys or girls physically in class.
5	School staff experience	Principal's years of experience as a school director and his or her years of experience working at the school in question as a teacher as well as other male and female teachers' experience at a school and status of formal training in teaching.
6	School-level	student to teacher ratio, average numbers of hours taught per day, existence of library, school management's support to students with their school fees, provision of after-school extracurricular activities, existence of

Group	Theme	Components
		a feeding program, active reading club, parent-teacher association, and provision of student kits.
7	Type of school	Private, public, community, religious, semi-private.
8	Implementer	Unobservable implementer-level factors such as socio-economic status and geographic factors.

ANNEX D: ADDITIONAL FINDINGS

Basic students level information across schools

Indicators	Treatment Core (T1)	Treatment Enhanced (T2)	Control (C)	p-value [T1-C]	p-value [T2-C]
Number of boys and girls in grade 1	37.9	60.1	36.7	0.7	0.0
Number of boys and girls in grade 2	37.8	53.7	33.8	0.29	0.00
Students to teacher ratio	23.8	26.4	20.7	0.29	0.07
Normal entry age for students at this school	4.88	5.21	4.62	0.33	0.02
Normal exit age for students at this school?	13.8	13.2	14.8	0.48	0.24
Average number of hours taught per day	5.33	5.15	5.18	0.26	0.74
% of schools reported dropout since the beginning of the school year.	39%	62%	51%	0.18	0.20
Number of girls dropped out since the beginning of the school year?	6.37	9.49	5.74	0.76	0.16
Number of boys dropped out since the beginning of the school year?	6.68	6.57	7.19	0.84	0.79
% Mixed (co-ed) school	98.3%	93.4%	100.0%	0.32	0.04
% schools with all teaching positions filled by school year start date	85.0%	60.7%	75.4%	0.188	0.08

School level resources and activities, by treatment and comparison

Indicators	Treatment Core (T1)	Treatment Enhanced (T2)	Control (C)	p-value [T1-C]	p-value [T2-C]
A. School Management and Services					
% school have an active parent-teacher association?	84%	77%	72%	0.10	0.54
% school management support students with their school fees?	83%	61%	76%	0.39	0.10
% school have after-school extracurricular activities?	17%	20%	18%	0.84	0.82
% school have a school feeding program	37%	69%	33%	0.66	0.00
% school have a reading club	27%	15%	15%	0.11	1.00
% school distribute student kits	33%	57%	26%	0.40	0.00
% school provide any health services to students	30%	33%	33%	0.74	1.00
% schools visited by a member of the MENFP in the past year to monitor facility	88%	70%	71%	0.02	0.98
B. School Infrastructure					
% schools have computers that students can use or a computer room?	22%	7%	13%	0.22	0.23
% schools have a library	47%	25%	31%	0.08	0.42
% schools have a chalkboard in every classroom	100%	95%	98%	0.32	0.31
% schools benefit from infrastructure renovations	33%	33%	33%	0.95	1.00

% schools benefit from water, sanitation, and hygiene (WASH) intervention	42%	33%	39%	0.80	0.45
% schools have separate toilets for girls and boys	66%	70%	60%	0.52	0.28
% schools have a handwashing station near the girls' latrines	64%	52%	56%	0.51	0.76

Director characteristics

Observed indicators	Treatment Core (T1)	Treatment Enhanced (T2)	Control (C)	p-value [T1-C]	p-value [T2-C]
Years of experience of the principal as a teacher?	20.4	19.5	19.6	0.72	0.94
Years of experience of the principal as a school director	13.7	12.2	12.4	0.46	0.90
Years of experience working at this school	14.5	14.7	14.9	0.83	0.92
Years of experience as the director of this school?	13.2	10.2	11.2	0.23	0.51

Teacher characteristics

Observed indicators	Treatment Core (T1)	Treatment Enhanced (T2)	Control (C)	p-value [T1-C]	p-value [T2-C]
Number of full-time teachers work at this school (30 or more hours a week)	8.82	8.61	7.79	0.13	0.16
Number of full time female teachers	5.03	5.00	4.20	0.13	0.17
Number of part time teachers at the school part-time (less than 30 hours a week)	4.26	3.25	3.90	0.80	0.66
Number of part-time female teachers	0.25	0.10	0.20	0.67	0.37
Number of female teachers with formal teacher training	3.74	3.63	3.30	0.44	0.60
Number of male teachers with formal teacher training	3.84	3.59	3.46	0.70	0.90
i. More than secondary school	3.48	3.09	2.88	0.33	0.74
ii. Secondary school	3.29	2.08	1.74	0.19	0.55
iii. Primary school	1.33	0.04	0.32	0.02	0.21
i. More than secondary school	4.87	3.97	4.91	0.97	0.48
ii. Secondary school	1.32	1.52	1.11	0.47	0.24
iii. Primary school	0.62	0.00	0.20	0.08	0.26

EGRA scores, by grade and treatment

Grade		Core (T1)	Enhanced (T2)	p value
Grade 1	clpm	10.9	5.9	0.00
Grade 1	clspm	6.7	4.5	0.00
Grade 1	cwpm	3.0	1.7	0.00
Grade 1	orf	3.1	3.9	0.25
Grade 1	Creole letter dictation score (out of 5)	3.1	2.5	0.00
Grade 1	Creole word dictation score (out of 3)	0.1	0.0	0.00
Grade 1	Initial sound recognition score (out of 10)	1.4	1.2	0.13
Grade 1	Listening Comprehension score (out of 5)	3.3	3.2	0.10
Grade 1	Reading comprehension score (out of 5)	0.1	0.1	0.34
Grade 2	clpm	26.2	18.7	0.00
Grade 2	clspm	14.3	10.4	0.00
Grade 2	cwpm	13.3	7.8	0.00
Grade 2	orf	16.9	12.0	0.00
Grade 2	Creole letter dictation score (out of 5)	4.2	3.8	0.00
Grade 2	Creole word dictation score (out of 3)	0.7	0.4	0.00
Grade 2	Initial sound recognition score (out of 10)	3.1	1.6	0.00
Grade 2	Listening Comprehension score (out of 5)	3.7	3.7	0.84
Grade 2	Reading comprehension score (out of 5)	0.8	0.5	0.00

EGRA scores, by gender and treatment

		Treatment			P value	
	var	Core (T1)	Enhanced (T2)	Control (C)	[T1-C]	[T2-C]
Boys	clpm	17.16	10.49	12.53	0.00	0.03
Boys	clspm	9.60	6.65	6.80	0.00	0.79
Boys	cwpm	6.70	3.79	4.41	0.00	0.26
Boys	orf	7.62	7.20	5.01	0.00	0.02
Boys	Creole letter dictation score (out of 5)	3.59	3.01	3.14	0.00	0.19
Boys	Creole word dictation score (out of 3)	0.37	0.16	0.17	0.00	0.61
Boys	Initial sound recognition score (out of 10)	2.26	1.34	1.54	0.00	0.27
Boys	Listening Comprehension score (out of 5)	3.50	3.50	3.33	0.03	0.02
Boys	Reading comprehension score (out of 5)	0.40	0.28	0.22	0.00	0.20
Girls	clpm	19.83	13.90	15.62	0.00	0.10
Girls	clspm	11.43	8.11	9.01	0.00	0.19
Girls	cwpm	9.58	5.63	6.95	0.00	0.09
Girls	orf	12.35	8.59	9.56	0.03	0.41
Girls	Creole letter dictation score (out of 5)	3.70	3.32	3.47	0.01	0.11
Girls	Creole word dictation score (out of 3)	0.45	0.25	0.28	0.00	0.55
Girls	Initial sound recognition score (out of 10)	2.30	1.46	1.56	0.00	0.61
Girls	Listening Comprehension score (out of 5)	3.53	3.37	3.40	0.09	0.80
Girls	Reading comprehension score (out of 5)	0.53	0.37	0.36	0.00	0.89

Learning outcomes across treatment groups (Core vs. Enhanced) by type of school

		Core (T1)	Enhanced (T2)	Difference [T2-T1]	p value
Private	clpm	20.5	15.1	-5.42	0.00
Private	clspm	11.4	8.2	-3.13	0.00
Private	cwpm	9.6	6.1	-3.50	0.00
Private	orf	11.8	7.9	-3.89	0.00
Private	Creole letter dictation score (out of 5)	3.8	3.5	-0.28	0.00
Private	Creole word dictation score (out of 3)	0.5	0.3	-0.22	0.00
Private	Initial sound recognition score (out of 10)	2.5	1.6	-0.87	0.00
Private	Listening Comprehension score (out of 5)	3.5	3.4	-0.09	0.21
Private	Reading comprehension score (out of 5)	0.6	0.4	-0.15	0.01
Public	clpm	12.9	10.2	-2.64	0.01
Public	clspm	8.2	6.8	-1.35	0.03
Public	cwpm	4.1	3.8	-0.30	0.65
Public	orf	4.9	7.9	3.01	0.00
Public	Creole letter dictation score (out of 5)	3.3	2.9	-0.35	0.00
Public	Creole word dictation score (out of 3)	0.3	0.2	-0.09	0.05
Public	Initial sound recognition score (out of 10)	1.7	1.2	-0.43	0.05
Public	Listening Comprehension score (out of 5)	3.5	3.4	-0.02	0.80
Public	Reading comprehension score (out of 5)	0.2	0.3	0.06	0.21

ANNEX F: EVALUATION STATEMENT OF WORK

STATEMENT OF WORK: IMPACT EVALUATION OF USAID/HAITI'S EARLY READING PROGRAM

PURPOSE OF THE EVALUATION

The Early Reading Project (ERP) is a two-year intervention to support early grade literacy in the first two grades of primary school in up to approximately 670 public, community, and semi-private Haitian schools. The impact evaluation is intended to measure whether the ERP intervention in these schools has resulted in students in the target schools demonstrating greater skills in reading than their peers of the same grades in non-targeted homologous schools. The primary stakeholders for this evaluation include the Ministry of Education and Professional Training (MENFP), USAID/Haiti, the implementing partners of the Early Reading Project (the LAC/Reads program, UNICEF, UND, Save the Children, and UNOPS), and other donors and entities active in Haiti's education sector (the World Bank, JICA, IDB, World Vision, etc.), and a variety of NGOs and religious affiliations devoted to improving learning in Haitian primary schools.

SUMMARY INFORMATION

Activity Name	USAID Office	Implementer	Cooperative Agreement/ Contract #	TEC	Life of Strategy / Project / Activity	Active Geographic Regions	Mission DO
Early Reading Program (ERP)	Education	Save the Children	N/A	N/A	2017-2019	PAP	4
LACReads	Education	AIR	AID-OAA-A-14-00058	\$5m	2016-2019	TBD	4
Read Haiti: Scaling Early Grade Reading Improvement in Haiti	Education	UND	AID-521-A-16-00007	GDA \$12m	2016-2020	Grand Anse TBC	4
Haiti Wins	Education	UNICEF	GHA-G-00-07-00007	\$14m	2017-2019	North, North-East, South-East	4

BACKGROUND

DESCRIPTION OF THE PROBLEM

According to the national curriculum, Haitian students are intended to study reading in Creole and French for 40 minutes each three out of every five days of the week each school year. However, reading instruction in the early grades in Haitian schools is often cursory, under-resourced, and ineffective. As a result, Haitian children score badly on standardized reading assessments. Few materials properly designed for the instruction of reading in the early years are available, teachers are rarely trained in the techniques that research has shown to be most effective in the teaching of reading, and formative evaluation is almost never properly conducted, so students who are struggling to learn to read with ease are often not identified in time for an appropriate intervention. System-wide, Haiti practices ‘automatic promotion’ in the early primary years, with the result that students who have not learned to read are passed up from grade to grade until the point where their deficits in reading skills force them to fail out or drop out of school.

The MENFP is aware of students’ poor outcomes in reading instruction and of the system-wide challenges that must be addressed to improve those outcomes. In partnership with USAID, and using USAID funding, the MENFP, from 2011-2016, implemented two reading interventions: the TOTALE program and the Ann ALE program. In total, 297 schools were supported in improving reading instruction under these programs, and many organizations present in the education sector in the country acquired copies of the programs student and teacher materials and used them in their own interventions. Unfortunately, however, the USAID-organized support for these programs was not sufficient to ensure fidelity of implementation of the scripted reading lessons in all the 297 targeted schools. As a result, materials did not always arrive on time in all the classrooms, and few teachers received the sustained support they required to apply the program effectively in their teaching.

SUMMARY STRATEGY/PROJECT/ACTIVITY/INTERVENTION TO BE EVALUATED

To address these flaws and strengthen the implementation of evidence-based early grade reading instruction in Haitian schools, USAID in 2017 launched a new program known as the Early Reading Project (ERP). Under the ERP, USAID works with the LAC/READS program and its partners (FONHEP), with UND, Save the Children and with UNICEF to implement a streamlined and enhanced early grade reading program in the early grades of school. There is also a pending award that will be included in year 2. The ERP program has the following three objectives:

1. Improve the MENFP’s capacity to implement and eventually scale reading reform
2. Improve teachers’ ability to teach early grade reading
3. Improve students’ early grade reading outcomes

Partners’ responsibilities for implementing ERP are detailed in the table below:

PARTNER	ROLE	Y1 schools	Y2 schools	Department(s)
TBD ('core' ERP model)	1. Provide teacher and administrator training on the use of the ERP curriculum for 150 public and community schools 2. Provide teacher coaching and monitoring on the use of the ERP curriculum in 150 public and community schools 3. Utilize existing MENFP systems and architecture to the maximum extent possible to transfer capacity in the areas of teacher training and coaching for the use of ERP materials	0	150	TBD
LAC Reads ('core' ERP model)	1. Support the MENFP in organizing and communicating about the ERP program 2. Activate teacher training and coaching networks, via sub-award to FONHEP, to provide training on the use of the ERP curriculum to teachers in 200 majority semi-private, religiously affiliated schools (some public schools are possible) 3. Activate and reinforce teacher training and coaching systems via FONHEP to provide coaching on the use of ERP curriculum to teachers in 200 majority semi-private, religiously affiliated schools	150	200 (+50 and retain same 150)	TBD
UND ('enhanced' ERP model)	1. Design and implement, in 160 schools, program extensions and enhancements focusing on the instruction of reading comprehension and on the use of supplementary program materials 2. Implement USAID and MENFP co-developed and co-designed bilingual, scripted, phonics-based ERP curriculum in 160 majority Catholic (some public) schools 3. Deliver teacher and supervisor training in ERP curriculum use in 160 majority Catholic (some public) schools	160	160	Grand-Anse (TBC)
UNICEF ('enhanced' ERP model)	1. Design and implement, in a subset of the core 120 schools, a program to teach pre-reading skills in the pre-Kindergarten and Kindergarten years (this is still in concept phase) 2. Implement USAID and MENFP co-developed and co-designed bilingual, scripted, phonics-based ERP curriculum in 120 schools 3. Conduct community engagement activities in support of early grade reading, such as summer camps and parent-teacher organizations	120	120	North, North-East, South-East

PARTNER	ROLE	Y1 schools	Y2 schools	Department(s)
Save the Children ('core ERP model)	I. TT from within the SAVE organization's ECP, among a mixed group of schools.	39	39	Port-au-Prince

DEVELOPMENT HYPOTHESIS AND THEORY OF CHANGE

The ERP development hypothesis is that a lack of ministry leadership, a lack of transparency and efficiency on the part of UDCLEM, a habit of associating schools to the experience in a haphazard fashion, a lack of coordination amongst actors and networks, and a lack of strong feedback loops between the school classrooms and the program designers have all made it impossible for early grade reading to gain traction in Haitian schools. Effective implementation of early grade reading programming leading to improved reading outcomes can be achieved if the following conditions are met:

- a. The ministry and UDCLEM receive the needed support and technical guidance;
- b. All primary schools implement the program (even in an experimental stage);
- c. Pre-existing frameworks are leveraged to introduce the materials, teaching techniques and materials, and coaching routines to the schools;
- d. Strong feedback loops are established between all types of schools implementing the program and the design team in Port-au-Prince.

EVALUATION QUESTIONS

This impact evaluation is expected to answer the following questions:

Q1. What is the impact, if any, of the ERP intervention on students' reading outcomes in the first and second grades? Some sub-questions the evaluation team should consider in addressing this question are:

- a. How do students exposed to any ERP programming perform when compared to similar students who are not exposed to ERP programming?
- b. What is the marginal effect, if any, of ERP implementer on reading outcomes?
- c. What is the marginal effect, if any, of type of school (semi-private, public, community, religious, etc.) on reading outcomes?
- d. To what extent do environmental or situational factors have an impact on students' ability to learn to read in the ERP target schools?

Q2. What is the impact, if any, on student learning, of the "enhanced" and "enhanced plus" models carried out by UND? Do students exposed to this model acquire stronger reading skills than their peers in control and other ERP schools? What aspects of the model impact student learning the most?

Q3. What is the impact, if any, on student learning, of the “pre-reading instruction” model piloted by UNICEF in the pre-K and K levels of a subset of the UNICEFERP program schools? Do students exposed to the ‘pre-reading instruction’ model acquire stronger reading skills than their peers in control and other ERP schools? (TBD)

Q4. What is the impact of the ERP intervention, if any, on teachers’ classroom practices in early grade reading instruction?

- a. Do teachers exposed to the ‘enhanced’ ERP intervention teach reading differently and/or better than their peers in the ‘core’ ERP schools?
- b. Do teachers exposed to the ERP intervention teach reading differently and/or better than their peers in the identical or close to identical schools without ERP interventions?
- c. How, if at all, do teachers exposed to the ERP intervention teach reading differently after having received the ERP intervention?

Q5. What is the impact, if any, of the ERP program on the ministry’s capacity to implement and eventually scale up early grade reading reform? Are the human, material, and financial management capacities at the MENFP sufficiently improved because of the ERP intervention that the ministry could successfully expand the ERP program nationwide?

EVALUATION DESIGN AND METHODOLOGY

The ERP impact evaluation will determine the effect of the Haiti ERP program on students reading and learning outcomes and shall compare changes on those outcomes to what would have happened in the absence of the ERP intervention, called counterfactual. Social Impact is encouraged to propose either an experimental or quasi-experimental design to estimate the changes caused by the ERP intervention. It is also expected that Social Impact will complete three or four rounds of data collection, depending on cost considerations dictated by USAID. Options for data collection include:

1. 4 rounds of data collection: a baseline for the first set of schools in November-December 2017, a midline in April-May 2018, a baseline for the second set of schools in October-November 2018 and an endline for all schools in April – May 2019.
2. 3 rounds of data collection: a baseline for the first set of schools in November-December 2017, a baseline for the second set of schools in October-November 2018, and an endline for all schools in April – May 2019.

In all rounds of data collection, the ERP impact must be assessed, at the minimum, through the following seven (7) skills assessed:

- a. Letter-sound mastery (phonetic awareness)
- b. Non-familiar word reading
- c. Familiar word reading

- d. Reading connected text/Oral Reading Fluency (ORF)
- e. Reading comprehension
- f. Writing of a sentence dictated to the student
- g. Writing of a sentence the student generates in response to an open verbal prompt

DELIVERABLES AND REPORTING REQUIREMENTS (BASELINE)

1. **WORK PLAN:** Upon receipt of this Activity Request, Social Impact (SI) shall submit within two weeks a draft work plan to the Contracting Officer's Representative (COR). The work plan will include: (1) the anticipated schedule and logistical arrangements; (2) a list of the members of the evaluation team, delineated by roles and responsibilities with their level of effort; (3) the identification of other required personnel and relevant local subcontractors, their LOE, roles and responsibilities and qualifications; and (4) the deliverable schedule.
2. **EVALUATION DESIGN:** Within two weeks of approval of the work plan, SI must submit to COR an evaluation design (which will become an annex to the Evaluation report). The evaluation design will include: (1) a detailed evaluation design matrix that links the Evaluation Questions in the SOW to data sources, methods, and the data analysis plan; (2) draft questionnaires and other data collection instruments or their main features; (3) the list of potential interviewees and sites to be visited; (4) known limitations to the evaluation design; and (5) a dissemination plan.

USAID/Haiti will take up to 10 business days to review and consolidate comments through the COR. Once the evaluation team receives the consolidated comments on the initial evaluation design and work plan, they are expected to return with a revised evaluation design and work plan within 5 business days.

3. **IN-BRIEFING / INCEPTION REPORT:** Prior undertaking field work, the evaluation team will have an in-briefing with the USAID/Education Team and the Evaluation and Survey Services (ESS) COR to discuss the team's understanding of the assignment, initial assumptions, evaluation questions, methodology, and work plan, and to clarify any questions or logistic needs.
4. **BRIEFING/PRESENTATION:** The evaluation team is expected to hold a final presentation in person to discuss the summary of findings and recommendations to USAID within 20 business days after the conclusion of fieldwork.
5. **DRAFT BASELINE REPORT:** The draft evaluation report should be consistent with the guidance provided in Section IX: **Final Report Format**. The report will address each of the questions identified in the SOW and any other issues the team considers to have a bearing on the objectives of the evaluation. Any such issues can be included in the report only after consultation with USAID. The submission date for the draft evaluation report will be determined in the evaluation work plan. Once the initial draft evaluation report is submitted, the Education Office will have 10 working business days in which to review and comment on the initial draft, after which point the ESS COR will submit the consolidated comments to the

evaluation team. The evaluation team will then be asked to submit a revised final draft report 5 business days hence, and again the Education Office will review and send comments on this final draft report within 5 business days of its submission.

6. **FINAL BASELINE REPORT:** The evaluation team will be asked to take no more than 15 business days to respond/incorporate the final comments from the Education Office. The evaluation team leader will then submit the final report to the COR. All project data and records shall be submitted in full and should be in electronic form in easily readable format, organized, and documented for use by those not fully familiar with the intervention or evaluation, and owned by USAID.

EVALUATION TEAM COMPOSITION

The Evaluation Team shall include a Team Leader (TL) and an Assistant Team Leader (ATL). The two positions are considered Key Personnel and essential to the work being performed.

The Team Leader (TL) is ultimately responsible for the overall management of the evaluation team and the final products, in conformity with this Scope of Work. The TL must be an experienced evaluation expert, with a documented track record of 10 years' experience in the field of evaluation. The TL shall also have demonstrated experience in evaluating early grade reading projects. The TL shall have a master's degree in educational statistics, psychometrics, or other related education field, and will possess excellent writing and interpersonal skills. The Team Leader will be responsible for planning the evaluation, coordinating the implementation of the evaluation, assigning evaluation responsibilities and tasks, and authoring the final evaluation report. He/she must be fluent in English and French. Haitian Creole is highly desirable.

The ATL shall have a master's degree in education or statistics to ensure that all areas of expertise required for the evaluation are effectively covered. He/she shall be an expert in early grade reading instruction and analysis, with at least 7 years of experience in testing reading results in young children. Work experience with USAID will be an asset and fluency in English or French is required. Haitian Creole is highly desirable.

EVALUATION SCHEDULE

TASK	ESTIMATED DUE DATE (Aug, 2017-Nov 2019; 28 months)
Inception Report	September, 2017
Workplan	September, 2017
Scoping Trip	September, 2017
Evaluation Design	October, 2017
USAID review of the work plan, inception report and the evaluation design	October, 2017
Assessor and supervisor training	November, 2017
Pilot assessments and instrument finalization	November, 2017
Data collection and analysis	November-December, 2017
Preliminary Results Report for 2017	March, 2018
USAID Review of draft report	April, 2018
Optional midline data collection (as dictated by USAID)	April-May 2018
Final report for 2017 with USAID comments inserted	May, 2018
National Level Dissemination Workshop for 2017 results	June, 2018
Departmental Level Dissemination Workshops for 2017 results	July-August, 2018
Baseline data collection for the second set of schools	October-November 2018
Assessor and supervisor training (second round of data collection)	March, 2019
Endline data collection and analysis	April-May, 2019
Preliminary Results Report for 2019	July, 2019
USAID Review of draft report	August, 2019
Final report for 2019	September, 2019
USAID review of final report	October, 2019
Final report for 2019 with USAID comments inserted	October, 2019
National Level Dissemination Workshop for 2019 results	November, 2019
Summary report of consultations held to publicize the early grade reading and writing outcomes from the ERP program	November, 2019

FINAL REPORT FORMAT

The evaluation final report should include an abstract; executive summary; background of the local context and the strategies/projects/activities being evaluated; the evaluation purpose and main evaluation questions; the methodology or methodologies; the limitations to the evaluation; findings, conclusions, and recommendations. For more detail, see “How-To Note: Preparing Evaluation Reports” and **ADS 201MAH, USAID EVALUATION REPORT REQUIREMENTS**. An optional evaluation report **TEMPLATE IS AVAILABLE IN THE EVALUATION TOOLKIT**.

The executive summary should be 2–5 pages in length and summarize the purpose, background of the project being evaluated, main evaluation questions, methods, findings, conclusions, and recommendations and lessons learned (if applicable).

The evaluation methodology shall be explained in the report in detail. Limitations to the evaluation shall be disclosed in the report, with particular attention to the limitations associated with the evaluation methodology (e.g., selection bias, recall bias, unobservable differences between comparator groups, etc.).

The annexes to the report shall include:

- The Evaluation SOW;
- Any statements of difference regarding significant unresolved differences of opinion by funders, implementers, and/or members of the evaluation team;
- All data collection and analysis tools used in conducting the evaluation, such as questionnaires, checklists, and discussion guides;
- All sources of information, properly identified and listed; and
- Signed disclosure of conflict of interest forms for all evaluation team members, either attesting to a lack of conflicts of interest or describing existing conflicts of.
- Any “statements of difference” regarding significant unresolved differences of opinion by funders, implementers, and/or members of the evaluation team.
- Summary information about evaluation team members, including qualifications, experience, and role on the team.

In accordance with ADS 201, the contractor will make the final evaluation reports publicly available through the Development Experience Clearinghouse within three months of the evaluation’s conclusion.

OTHER REQUIREMENTS

All quantitative data collected by the evaluation team must be provided in machine-readable, non-proprietary formats as required by USAID’s Open Data policy (see ADS 579). The data should be organized and fully documented for use by those not fully familiar with the project or the evaluation. USAID will retain ownership of the survey and all datasets developed.

All modifications to the required elements of the SOW of the contract/agreement, whether in technical requirements, evaluation questions, evaluation team composition, methodology, or timeline, need to be agreed upon in writing by the COR. Any revisions should be updated in the SOW that is included as an annex to the Evaluation Report.

ANNEX G: QUALIFICATIONS OF TEAM MEMBERS

PROJECT DIRECTOR/PRINCIPAL INVESTIGATOR, DR. BASAB DASGUPTA, has over 18 years of professional experience in analytical research, designing and working on impact evaluations in international development. Joining SI in 2016 as a Technical Director, he has led international policy evaluations and possesses vast knowledge and experience in education, health, labor and human capital development. Most recently, he served as a Technical Advisor on the Department for International Development (DFID) Educate Girls, End Poverty in Somalia evaluation where he supervised and provided quality control for the end-line evaluation. Prior to SI, Dr. Dasgupta worked for the World Bank for 12 years, where he led and participated in many evaluations and advisory work related to education. For example, Dr. Dasgupta led the evaluation of household income shocks during financial crisis on children's education and health outcomes in five countries from Europe and Central Asia; housing subsidies to urban poor households in South Africa on their children's education outcomes; supported the evaluation program of Fast Track Initiative III (Basic Education) to assess attendance, dropout rate, and learning outcomes from quantitative and qualitative data. Additionally, he worked for the lending project for establishing a National Testing Center for higher education in Tajikistan. He also led an impact evaluation of School Based Management System on student's performance in Indonesia. With these projects, he worked closely with government officials and civil societies. Dr. Dasgupta has dozens of publications in the form of book chapters, articles in academic journals and World Bank policy research working papers. Dr. Dasgupta has a Ph.D. in Economics (2005) from the University of Connecticut.

TEAM LEADER, DR. JOAN MITCHELL-COHEN, is an education specialist and sociologist with over 10 years of experience working with reading and literacy programs both domestically in the US and abroad, particularly in the context of young children. Dr. Mitchell-Cohen completed her Ed. D in Educational Policy, and M. Ed in International Education from the University of Massachusetts. Dr. Mitchell-Cohen is currently Chief of Programming and Evaluation for the U.S. Peace Corps, where she has been working since 2012 as an Education/Literacy specialist. As a Literacy Specialist, Dr. Mitchell-Cohen led early grade reading/literacy efforts in concert with USAID Goal I, designed and developed all childhood literacy sector guidance, project frameworks, and M&E tools for Peace Corps HQ, 66 posts and 20 literacy programs worldwide. She has also designed technical training packages for Peace Corps Volunteers and staff globally on early literacy and early grade reading, Early Childhood Development (ECD), child development, assessment, classroom management, disability and inclusion, community engagement in literacy and early numeracy instruction. Dr. Mitchell-Cohen collaborated with USAID through the Global Education Framework to facilitate USAID and Peace Corps joint programming efforts in early grade reading globally. Through media, events and publications, she raised awareness of literacy and early grade reading within Peace Corps. She has also participated in international efforts to promote and innovate literacy instruction through mobile technology. As a Chief of Programming of Office of Overseas Programming and Training Support of Peace Corps, she provided strategic leadership and direction setting for the Programming Unit comprised of 12 sector specialists that served 8500 Volunteers and 800 programming staff world-wide. She has also worked as a Technical Expert for 5 years for Bernard Van Leer Foundation/Caribbean Child Support Initiative, Bridgetown, Barbados where she developed programming and a programmatic monitoring and evaluation framework, created teaching and learning materials on topics of family literacy, child development and early literacy for practitioners, and policymakers and designed simple parent focused materials for family literacy and

learning for 500+ program participants in four countries. Dr. Mitchel-Cohen is fluent in French and Creole.

ASSISTANT TEAM LEADER, MR. BAUGE MONCOEUR, is a trained engineer and professional statistician. He received his Engineering diploma at the State University of Haiti in 1994. He then obtained two Master's degree in Belgium: the first in demography at the Catholic University of Louvain in 1999, and the second in statistics at the Free University of Brussels in 2005. Since 2006, Mr Bauge has been working in Haiti as an independent consultant, carrying out large-scale surveys and studies. Mr. Bauge has been providing such services to Haitian government entities, international institutions such as the United Nations and the World Bank, and donors such as USAID. Mr Bauge, worked as a data collection coordinator in four of Haiti's departments including North, West, Grande Anse and Nippes, for the mid-term and end-term evaluations of the World Bank's EGRA program. His fields of interest are opinion polls, program evaluation, demographic, mathematic and time series analysis.

PROGRAM MANAGER/EVALUATION SPECIALIST, MR. ANDREW CARMONA, is an M&E professional with over ten years' experience in international development and six years' experience in monitoring and evaluating education, international health, WASH, agriculture, and economic development projects. At Social Impact, Mr. Carmona manages large-scale early grade reading, public financial management, at-risk youth, and leadership sector performance and impact evaluations. He contributes to impact evaluation design utilizing a mix of experimental and quasi-experimental research methods employing both quantitative and qualitative techniques. In addition, he designs, tests, and implements digital instruments using ODK software and SurveyCTO platform, including large-scale household, school-based, public institution, and enterprise surveys, as well as focus group and key informant interview discussion guides. Previously, Mr. Carmona was a Research Activity Manager on the USAID Strengthening Health Outcomes Through the Private Sector Project where he led or participated in evaluations of the private health sector in Senegal, Niger, Benin, and Burundi. Mr. Carmona facilitated over 100 interviews, synthesized data, and contributed to final report findings and recommendations. In Malawi and Benin, Mr. Carmona managed national census studies, which included development of quantitative survey tools, data collection oversight, data analysis, and final report writing. From 2010-2012, Andrew worked on the American Red Cross' Haiti team after the 2010 earthquake as an M&E Program Assistant where he coordinated data collection and indicator tracking for the Red Cross' \$500 million response program. He holds an M.P.A. in Development Practice from Columbia University and speaks French and Spanish proficiently.

PROGRAM ASSISTANT, EUPHONISE LOISEAU, delivers logistical and technical support to Program Managers and Technical Advisors on several of Social Impact's contracts and subcontracts with a variety of government donors. Ms. Loiseau also provides logistical support to project teams, including drafting and editing reports, translating documents, managing and cleaning quantitative and qualitative data, and assisting in the development of survey instruments and protocols. Prior to joining Social Impact, Ms. Loiseau supported the operation of an international security forum, attended by NATO Military officials and diplomats, regarding security concerns in North Africa and the Middle East. She supported research on NATO's growing role in the Middle East, political instability in Mali and Libya, and the effects of military weapon proliferation in the Sahel region. Ms. Loiseau holds a Bachelor of Arts in Political Science from Northeastern University.

LOCAL PARTNER, IFOS. Haiti ESS identified IFOS as their local partner in survey data collection. IFOS has substantial survey data collection expertise in the field of Education. During 2015-16, IFOS conducted midline and end-line surveys for a World Bank assessment of reading skills for Map Li Nèt Ale in Education Pour Tous (EPT) Schools. The survey covered 446 schools located in rural and urban areas in four departments: Grand Anse, Nippes, Ouest and Nord-Ouest on behalf of EPT/MENFP. Recently, IFOS conducted surveys for CLAMS³⁴ assessment— a longitudinal study of Grades 1-3 students to measure progress in their reading and writing skills in Haitian Creole and French in Ann ALE schools in the three USAID strategic development corridors in Haiti. IFOS mobilized a large team of M&E Experts, Education Specialists, language Experts, EGRA trainers and assessors and supervisors to plan, implement and monitor the assessment. IFOS has developed an expertise in electronic surveys and worked in most difficult areas in Haiti. IFOS is working across Haiti through a network of consultants, trainers, and surveyors in Haiti's ten departments. IFOS has a strong team for conducting large surveys in the Education sector. The team leader, Charles Tardieu, is an Education Specialist and holds a Ph.D in Education. The rest of the key personnel from IFOS are well versed with EGRA surveys, especially in French and Creole.

³⁴ CLAMS refers to Clinical Linguistic and Auditory Milestone Scale

ANNEX H: EVALUATION DISSEMINATION PLAN TEMPLATE

Key Milestones	Expected Dates
Statement of Work	
Evaluation Design	
Draft Report	
Final Report	

EVALUATION TITLE:

Audience	Goal	Tool/Medium	Forum	Responsible Party	Timing	Follow-up
<i>Identify stakeholders by asking "Who is likely to be affected by the evaluation and its results? Who is likely to be interested?"</i>	<i>Are we simply pushing out information? Hoping to affect change? Contributing to the knowledge base?</i>	<i>These may include reports, briefs, presentations, blog posts, meetings, facilitated discussions, videos, journal articles, press releases, graphics, emails to listservs</i>	<i>Are there existing networks or venues through which findings should be disseminated? Or will communications be distributed directly to target audiences?</i>	<i>Who is responsible?</i>	<i>Is there a deadline?</i>	<i>Did we achieve our goal? What was the result of the information-sharing? Any observable outcomes?</i>
1:	1:	1:	1:	1:	1:	1:
2:	2:	2:	2:	2:	2:	2:
3:	3:	3:	3:	3:	3:	3:
4:	4:	4:	4:	4:	4:	4:

Source: USAID Bureau for Policy, Planning and Learning, April 2015.

PAGE LEFT INTENTIONALLY BLANK

ANNEX I: DATA QUALITY ASSURANCE MEASURES

For every data collection activity in its portfolio, ESS and its evaluation team embed SI's systematic quality assurance framework called Evaluation Quality, Utilization, and Impact (EQUI®). This system is comprised of a comprehensive set of guidance documents, templates, examples, and check-points. The EQUI® framework aggregates learning from nearly a decade of experience conducting more than 30 IEs and quantitative surveys worldwide and incorporates social science standards and best practices. The system helps evaluation teams efficiently manage survey activities and achieve ESS's high technical quality standards. Our system includes more than ten comprehensive guidance documents for planning data collection and monitoring data quality, to ensure that these activities are carried out with the highest level of integrity.³⁵

To minimize data entry errors and obtain data quickly for verification, the ET utilized electronic data collection tools to collect most of the data. Specifically, IFOS enumerators were equipped with tablets pre-loaded with all quantitative survey instruments. The evaluation team programmed the tools into the tablets using Open Data Kit (ODK) software and tested each of the instruments repeatedly leading up to data collection. For the EGRA tool, the ET used the *Tangerine* software programmed on tablets. The ET vetted, tested, and re-tested the data collection instruments, and then trained, tested, and supervised enumerators. Additionally, the ET programmed the tablets with internal quality checks - skip patterns and logic checks – and drop-down menus to ensure accuracy of certain key fields. The data uploaded to servers directly from each tablet daily and real time review.

The ET used a standard protocol for data and analysis file organization to ensure consistency, efficiency, transparency, and reproducibility. It also required that data collection partners produce data quality assurance reports detailing any challenges faced during fieldwork. Specifically, IFOS supervisors identified low performing enumerators and replaced them before data collection began. Field supervisors and the ET also reviewed all data collected on tablets to ensure that enumerators were administering surveys appropriately. Lastly, IFOS and one local ET member monitored data collection and sent SI's home office ET members weekly team progress reports on the data collection.

³⁵All ESS's dedicated IE staff are trained to use these guidance documents and templates, and learning is continually incorporated back into system improvements. We recognize that every large-scale survey presents unique challenges and we customize all templates to meet specific evaluation needs. Check-lists accompany each guidance document and are reviewed by at least two evaluation team members to ensure compliance with QA requirements. Moreover, by using the system's guidance and templates to collaborate with local data collection partners, EQUI® helps SI fulfil our institutional commitment of building local capacity.

ANNEX J: INTER-RATER RELIABILITY SCORES OF ENUMERATORS

IRR scores of the enumerators who conducted data collection are displayed below. Enumerators had to score a minimum of 85 percent on the IRR test to be eligible for data collection; those highlighted in red were rejected. A total of three IRR tests were held at the enumerator training; these are the results from the final test. Last names of enumerators have been removed for confidentiality purposes.

No	Department	First Name	Group	Sex	Percentage
1	Grand'Anse	Ernst	EGRA	M	92.2%
3	Grand'Anse	Vènèse	EGRA	F	94.4%
4	Grand'Anse	Dady	EGRA	M	96.7%
5	Grand'Anse	Hans William	SP	M	96.7%
7	Grand'Anse	Joane	EGRA	F	96.7%
8	Grand'Anse	Marie Danielle	EGRA	F	95.6%
11	Grand'Anse	Esaïe	EGRA	M	95.6%
12	Grand'Anse	Jocelyn	SP	M	97.8%
13	Grand'Anse	Jackson	EGRA	M	94.4%
14	Grand'Anse	Darline	EGRA	F	97.8%
15	Grand'Anse	Gersand Ronide	EGRA	F	93.3%
16	Grand'Anse	Alex	EGRA	M	95.6%
17	Grand'Anse	Jean Pricien	EGRA	M	98.9%
19	Grand'Anse	Johnny Wickers	EGRA	M	95.6%
20	Grand'Anse	Daniel	SP	M	96.7%
21	Grand'Anse	Marie Erline	EGRA	F	93.3%
22	Grand'Anse	Joe Rosano	EGRA	M	96.7%
23	Grand'Anse	Pierre Marie	SP	M	94.4%
25	Grand'Anse	Jean Richard	EGRA	M	93.3%
27	Grand'Anse	Jean Bénissé D.	SP	M	96.7%
28	Grand'Anse	Jean Rouslé	EGRA	M	94.4%
30	Nord/Nord-est	Yvette	EGRA	F	97.8%
31	Nord/Nord-est	Merline Claude	SP	F	96.7%
33	Nord/Nord-est	Lisemène	EGRA	F	96.7%
35	Nord/Nord-est	Reynald	EGRA	M	97.8%
38	Nord/Nord-est	Sanjuannicomaises	EGRA	M	94.4%
36	Nord/Nord-est	Sainçois	EGRA	M	95.6%
39	Nord/Nord-est	Bendy	SP	M	92.2%
40	Nord/Nord-est	Elise	EGRA	F	95.6%
41	Nord/Nord-est	Monestime	EGRA	M	92.2%
43	Nord/Nord-est	Cassandra	EGRA	F	92.2%
44	Nord/Nord-est	Gladlyn	EGRA	F	91.1%
47	Nord/Nord-est	Mertilus	EGRA	M	94.4%
49	Nord/Nord-est	Vanessa	SP	F	96.7%
48	Nord/Nord-est	Rose-Manie	EGRA	F	95.6%
50	Nord/Nord-est	Claudine	EGRA	F	91.1%
51	Nord/Nord-est	Gitane	SP	F	92.2%
53	Nord/Nord-est	Evintz	EGRA	M	95.6%
54	Nord/Nord-est	Jesula	EGRA	F	95.6%

55	Nord/Nord-est	Linsly	EGRA	M	96.7%
	Nord/Nord-est	Benito	SP	M	
57	Nord/Nord-est	Rechovsky	EGRA	M	96.7%
58	Nord/Nord-est	Edwige	EGRA	F	93.3%
60	Ouest/SE	Rose Audine	SP	F	97.8%
61	Ouest/SE	Marie-Micka	EGRA	F	96.7%
103	Ouest/SE	Edzer	EGRA	M	90.0%
104	Ouest/SE	Louis Salomon	SP	M	97.8%
63	Ouest/SE	Stael	EGRA	F	93.3%
64	Ouest/SE	Chilène Clergé	EGRA	F	83.3%
105	Ouest/SE	Nathalie	EGRA	F	97.8%
106	Ouest/SE	Sauvenel Michel	SP	M	96.7%
107	Ouest/SE	Marcus	EGRA	M	93.3%
110	Ouest/SE	Pierre Olsen	EGRA	M	93.3%
67	Ouest/SE	Sandie	EGRA	F	97.8%
111	Ouest/SE	David	EGRA	M	95.6%
68	Ouest/SE	Menadette	EGRA	F	94.4%
112	Ouest/SE	Dieuna	EGRA	F	93.3%
113	Ouest/SE	Lioud	EGRA	M	95.6%
66	Ouest/SE	Géhu Miliana	EGRA	F	95.6%
115	Ouest/SE	Emilius Fils	EGRA	M	90.0%
117	Ouest/SE	Dimy	SP	M	96.7%
118	Ouest/SE	Sylvio	EGRA	M	85.6%
70	Ouest/SE	Stevensia Sindy	EGRA	F	95.6%
71	Ouest/SE	Avril	EGRA	F	96.7%
72	Ouest/SE	Sony Annie	SP	F	95.6%
120	Ouest/SE	Francklin	EGRA	M	91.1%
73	Sud	Roseberta	EGRA	F	82.2%
79	Sud	Medgine	EGRA	F	96.7%
75	Sud	Mackenson	EGRA	M	84.4%
80	Sud	Wolfie Stael	EGRA	F	94.4%
82	Sud	Moline	EGRA	F	92.2%
96	Sud	Micheline	SP	F	94.4%
85	Sud	Danica	EGRA	F	78.9%
86	Sud	Germine	EGRA	F	95.6%
89	Sud	Delot	EGRA	M	96.7%
90	Sud	Lidenie	SP	F	96.7%
92	Sud	Ledens	SP	M	93.3%
94	Sud	Pierre Max	SP	M	94.4%
93	Sud	Marie Françoise	EGRA	F	92.2%
95	Sud	Valentino Giovanni	EGRA	M	92.2%
77	Sud	Rooloph	EGRA	M	96.7%
78	Sud	Ricker B.	EGRA	M	95.6%
98	Sud	Sherlyne	EGRA	F	95.6%
99	Sud	Imelda Tatiana	EGRA	F	84.4%
100	Sud	Medeline	EGRA	F	96.7%
101	Sud	Cherline	SP	F	93.3%
102	Sud	Denise	EGRA	F	91.1%

ANNEX K: CONSIDERATIONS FOR FUTURE DATA COLLECTION

The ET ran into several problems at baseline which hampered the team's ability to collect high quality data. The points below act as lessons learned from baseline and will be employed at subsequent data collection periods (midline and endline) to mitigate against the risk of low data quality in certain areas.

1. Rearrange important questions in the qualitative FGD and KII guides to focus more specifically on reading barriers (students and parents). Ensure that these questions are up front and not at the end of the guides.
2. Qualitative questions should be piloted to ensure appropriateness for Grade 1 and 2 students. Questions may have been too advanced for these children.
3. Add a question about language of instruction to the Classroom Observation Tool to capture the effects of language in the classroom.
4. Consider adding teacher quantitative questionnaires as a new tool in order to collect additional data on teachers being observed.
5. Add more questions in the Observation Tool on ERP-specific curriculum and teacher training.
6. Fix the ODK error in the Classroom Observation Tool which precluded the ability to note a teacher's level of formal training.
7. At the conclusion of BL data collection, the ET saw high data saturation among qualitative data. By the halfway point (about 30 events out of 56), most themes began repeating. Consider reducing the number of events at midline and/or endline.
8. Only two of 14 MENFP officials interviewed for this evaluation were female. While it is unclear if this is representative of MENFP as a whole, it is important that departmental-level MENFP offices are gender balanced to the extent possible. The ET will work to the extent possible to ensure gender balance in future data collection rounds.

ANNEX L: DISCLOSURE OF ANY CONFLICTS OF INTEREST

NAME	
Title	
Organization	
Evaluation Position?	☐ Team Leader ☐ Team member
Evaluation Award Number (Include project name(s), implementer name(s) and award number(s), if applicable)	
I have real or potential conflicts of interest to disclose	☐ Yes ☐ No
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the 5. operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 6. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 7. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	