



USAID
FROM THE AMERICAN PEOPLE

ተግባራዊ ጥራት ለህዝብ
ተግባራዊ ጥራት ለህዝብ
ለ ግብርና ጥራት



المملكة المغربية
وزارة التربية الوطنية
والتكوين المهني

USAID/MOROCCO READING FOR SUCCESS SMALL-SCALE EXPERIMENTATION ACTIVITY (RFS-SSE)

EARLY GRADE READING ASSESSMENT BASELINE REPORT

Contract No. AID-608-TO-15-00002

Submission Date: 6 May 2016

EARLY GRADE READING ASSESSMENT

BASELINE REPORT

RFS-SSE

CONTRACTED UNDER ORDER NO. AID-608-TO-15-00002

DISCLAIMER

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

ACKNOWLEDGEMENTS

The School-to-School International and Chemonics International teams would like to thank USAID/Morocco for their support. Particularly, we wish to acknowledge the important contribution of Mariam Britel-Swift from the Education office at USAID Morocco, her and her team have provided continuous advice and valuable support throughout the preparation and conduct of the study.

We would also like to recognize the staff from the Ministry of Education and Vocational Training and namely the National Center for Evaluation and Exam (CNEE) led by Mr. Mohammed Sassi Director of the Center, Mr Ahmed Chaibi, head of the evaluation division, Mr. Said Bouderga head of the examinations division and Mrs. Hafida Kach manager at the center. The CNEE has played an important role at all the different stages of the study: sampling, tools development tools, identifying data collectors and development of the data analysis plan.

We also thank the regional and provincial offices (AREFs) who provided all the necessary support through the mobilization of school inspectors who have effectively contributed to the development of the EGRA test and SSME questionnaires as well as inspectors and counselors who led data collection with great professionalism.

The following persons played a major role in the process:

El Ghordaf Abdelaziz	(Chemonics International)
Selma Talha Jebri	(Chemonics International)
Candace Debnam	(School-to-School International)
Maryvonne Denauw	(School-to-School International)
Selene Rangel	(School-to-School International)
Michel Rousseau	(School-to-School International)
Claire Wasserman	(School-to-School International)

CONTENTS

List of Tables and Figures.....	3
Acronyms.....	6
Chapter 1: Introduction.....	7
1.1 The Reading for Success—Small-Scale Experimentation (RFS-SSE) Activity.....	7
1.2 Evaluation Design.....	7
1.3 What are EGRA and SSME?.....	8
Chapter 2: Methodology.....	9
2.1 RFS-SSE Sampling Design.....	9
2.2 School and Student Sample.....	9
2.3 EGRA/SSME Tool Development.....	10
2.4 Data Collection.....	13
2.5 Quality of Assessment Tools.....	13
2.6 Data Analysis.....	14
Chapter 3: EGRA Results.....	15
3.1 EGRA Results by Subtask.....	15
3.2 EGRA Results by Gender.....	23
Chapter 4: Contextual Variables and EGRA Performance.....	24
4.1 Student Questionnaire.....	24
4.2 Teacher Questionnaires.....	35
4.3 School Director Questionnaires.....	39
4.4 Parent Questionnaire.....	44
4.5 Classroom Inventory Chart.....	45
Chapter 5: Conclusion and Recommendations.....	48
Annexes.....	49

LIST OF TABLES AND FIGURES

Table 1. Number of Students by AREF (Académie régionale d'éducation et de formation) and by Province	10
Table 2. Number of students by sex and group.....	10
Table 3. Number of students by area and by group	10
Table 4. EGRA Subtasks	11
Table 5. Index of internal consistency by subtask and for the entire EGRA	14
Table 6. Results of Experimental Group by Indicator	15
Table 7. MEan Scores, Phonemic Awareness.....	16
Table 8. Proportion of students by number of correct responses by group, Phonemic Awareness ..	17
Table 9. Mean Scores, Syllable Identification	17
Table 10. Mean Scores, Nonword Reading	19
Table 11. Mean Scores, Oral Reading Fluency	20
Table 12. Proportion of Students By Numberof Correct Responses By Group, reading comprehension	21
Table 13. Mean Scores, Reading Comprehension	21
Table 14. Proportion of Students By Numberof Correct Responses By Group, listening comprehension.....	21
Table 15. Mean Score, Llistening Comprehension.....	22
Table 16. Correlation matrix between the number of correct responses for each of the EGRA subtasks	22
Table 17. Performance per subtask by type of school location.....	23
Table 18. Performance by subtask by Gender	23
Tableau 19. Performance By Subtask According to the Language Spoken at HOme	24
Table 20. Performance by subtask according to the repeat status of the student	25
Table 21. Means of transportation for getting to school	25
Table 22. Person who accompanies the students on the way to school	25
Table 23. Reaction of the teacher to a correct response in the exercise book.....	26
Table 24. Reaction of the teacher to an incorrect answer in the exercise book	26
Table 25. Performance by subtask according to the reaction of the teacher to an incorrect response in the exercise book	26
Table 26. Person who provides help for the student in completing assignments at home	26
Table 27. Performance by subtask according to whether the student had breakfast/lunch before coming to school.....	27
Table 28. Performance by subtask according to whether the student has a lunchroom at school	27
Table 29. Performance by subtask according to whether the student had been absent from school or not	28
Table 30. Reasons given by the student for having been absent from school.....	28
Table 31. Performance by subtask according to whether the student had been late to school or not	29
Table 32. Reasons given by the student for arriving late to school	29
Table 33. Performance by subtask according to whether the student said that no one in the household knew how to read.....	30
Table 34. Performance by subtask according to whether the student said his/her father knew how to read	30
Table 35. Performance by subtask according to whether the student said his/her mother knew how to read	31
Table 36. Performance by Subtask According to whether the Student Said his/her Mother and/or Father Knew How to Read.....	31

Table 37. Performance by subtask according to whether the student had a book in Arabic at home	32
Table 38. Performance by subtask according to whether the student has books/magazines/newspapers at home	32
Table 39. Performance by subtask according to whether someone reads with the student at home	33
Table 40. Number of students claiming that the following resources were available at home	33
Table 41. Correlation between the number of resources at home and the performance on each subtask	34
Table 42. Performance by subtask according to whether the student had water at home	34
Table 43. Source of the water available at home	34
Table 44. Persons living with student in SAME Household	35
Table 45. Performance by subtask according to whether the student had to do chores outside of school hours	35
Table 46. Types of Chores Assigned to Students	35
Table 47. Performance of the students by subtask, according to the sex of the teacher	36
Table 48. Performance of students by subtask according to whether the teacher Updates the Lesson Plans or not	37
Table 49. How often the lesson plans are updated	37
Table 50. Individual from whom the teacher requested help	37
Table 51. How often the inspector evaluates the teacher	38
Table 52. Areas in which the inspector gave advice	38
Table 53. Methods used to assess student skills	38
Table 54. Ways oral and written assessment is used	38
Table 55. Type of security problem	39
Table 56. Correlation between the number of years of experience as school director and the performance on each subtask	39
Table 57. Number of days absent by the school directors during the last month	40
Table 58. Grades taught in the school	40
Table 59. Number of days closed outside of school vacations	41
Table 60. Number of students in the school	41
Table 61. Correlation between the number of students in the school and performance on each subtask	41
Table 62. Number of teachers in the school*	41
Table 63. Actions taken when a teacher is absent	42
Table 64. How often the lesson plans are checked	42
Table 65. How often the work of the teacher in the classroom is evaluated	42
Table 66. How often students visited the library	42
Table 67. How often the parents' association holds meetings	43
Table 68. Parents' associations' areas of activity	43
Table 69. Reasons for the teaching inspector's visit	43
Table 70. Number of minutes required to get to school	44
Table 71. Reason for students' absenteeism	45
Table 72. Number of students in the class	45
Table 73. Number of students who had a reading book	45
Table 74. Equipment available in the classroom	46
Table 75. Performance by subtask according to whether a teacher's guide in Arabic was available or not in the classroom	46
Table 76. Other books/booklets available and accessible in the classroom	47
Table 77. Organization of classroom space	47

Figure 1. Study Design by Evaluation Cycle.....	8
Figure 2. Distribution of Correct Responses By Group, Phonemic Awareness	16
Figure 3. Distribution of Correct Responses by group, Syllable Identification.....	18
Figure 4. Distribution of Correct Responses By Group, Nonword Reading.....	19
Figure 5. Distribution of Correct Responses By Group, Oral Reading Fluency.....	20

ACRONYMS

AREF	Académie Régionale d'Education et de Formation
CFRN	Centre Des Formations et Des Rencontres Nationales (Center for Trainings and National Assemblies)
CNEE	National Center for Evaluations and Exams
CSO	Civil Society Organizations
CWPM	Correct Words Per Minute
EDO	Education District Officers
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
EMIS	Education Management Information Systems
MOE	Ministry of Education and Vocational Training
MSA	Modern Standard Arabic
PFC	Provincial Field Coordinators
RFS-SSE	Reading for Success-Small Scale Experimentation
SSME	Snapshot of School Management Effectiveness
STS	School-to-School International
USAID	United States Agency for International Development

CHAPTER 1: INTRODUCTION

1.1 The Reading for Success—Small-Scale Experimentation (RFS-SSE) Activity

The Reading for Success—Small-Scale Experimentation (RFS-SSE) activity is the active component of a broader USAID initiative. The RFS-SSE is designed to reflect ongoing collaboration between USAID/Morocco and the Ministry of Education and Vocational Training (MOE) to improve reading instruction in Morocco. Conceived as a learning activity, the RFS-SSE will develop an evidence base of effective approaches to improve reading skills in targeted primary schools. RFS-SSE comes at a time when the MOE is developing a 15-year education reform called Vision 2030 and a set of medium-term activities for the period 2015-2020. The reform will address a key weakness in the Moroccan educational system: poor reading skills at the primary level. The effort in the RFS-SSE intervention will help to bridge the reform of existing curriculum to new curriculum by providing data and evidence for the envisioned changes.

The RFS-SSE activity is designed to test activities and to strengthen stakeholder engagement to improve reading skills in the early grades. It is the second phase of the larger RFS process that spans from analysis to national implementation of the Reading for Success (RFS) Program. RFS-SSE is using previous research and related activities carried out in Morocco and other countries and draw from lessons learned and evidence to test two reading interventions: new approach to Arabic reading lessons and summer enrichment activities. Reading lessons are being developed and tested and results will feed into planned MOE curriculum reform linked to the Interim Measures for 2015-2020. These efforts have been undertaken to better understand technical assistance needed for implementing activities that improve reading instruction in formal primary schools and test the effectiveness of reading materials. Similarly, the project will engage civil society organizations (CSOs) in the second summer of the project and support and test their efforts to reduce reading loss over the summer months through enrichment activities.

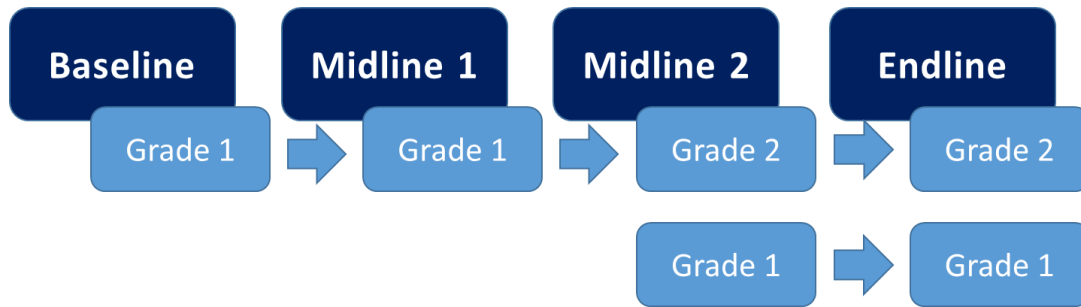
1.2 Evaluation Design

The intent of the RFS-SSE evaluation design is to determine gains in reading across four time points among students in RFS-SSE intervention schools as compared to non-intervention schools. Four evaluations are planned at intermittent time points throughout the life of the project.

In order to assess the impact of the RFS-SSE reading lessons, a longitudinal evaluation design (including equal representation of girls and boys in both urban and rural schools) was selected. As shown in

Figure 1, the evaluation design includes reading assessments of a cohort of Grade 1 students through the end of Grade 2 at four time points (Baseline, Midline 1, Midline 2, Endline) and a second cohort of Grade 1 students at two time points (Midline 2 and Endline). At baseline, all students were assessed with the same EGRA instrument that was developed to test students at a Grade 1 or Grade 2 level. Because the students were selected at random, it is possible to generalize the results to all students in the treatment group in the eight provinces where the reading lessons will be implemented.

FIGURE 1. STUDY DESIGN BY EVALUATION CYCLE



The baseline evaluation was conducted in January 2016. This report presents the findings of the baseline data collection.

1.3 What are EGRA and SSME?

The Early Grade Reading Assessment (EGRA) was developed with funding from USAID and the World Bank. This assessment measures student performance on the most basic foundational skills required for fluency in reading. EGRA assesses the skills needed for reading acquisition. Although many students are not yet fluent readers in the early grades, the EGRA allows us to capture what students, even the “nonreaders”, *can* do and where they are in the developmental path to becoming fluent readers.

EGRA is an individually administered, oral assessment that requires approximately 20 minutes per student. The test is administered to one student at a time by an enumerator (not the teacher) in a location that is outside of the classroom. The enumerator begins by explaining the assessment to the student and asking if the student agrees to participate. Consent is always optional and no student is required to take the assessment. The enumerator creates a relaxed environment for the student and assures the student that the assessment is not used for a grade. The enumerator then begins by asking the student questions aloud and having the student respond aloud. For certain subtasks, the enumerator places a paper stimulus in front of the student containing letters or words, and the enumerator asks the student specific questions about the stimuli. The EGRA administered for the RFS-SSE baseline consisted of six subtasks which are described in more detail below.

Globally, policymakers have used the results of EGRA to develop reforms to improve reading performance of students in the early grades. Moreover, these data have been used by educationists to establish student performance standards so that teachers understand expectations for students at each stage of reading development and can tailor instruction appropriately.

The Snapshot of School Management Effectiveness (SSME) is an instrument that quickly captures a picture (i.e., “snapshot”) of school management and pedagogic practice in a particular school. The SSME was designed to capture indicators that are believed to affect student learning. The results of SSME surveys can inform school, district, provincial, or national administrators and donors about current practices in their schools and classrooms.

SSME collects information on (1) basic school inputs such as infrastructure, learning materials, teacher and school director characteristics, student characteristics, and parental and community

involvement; and (2) classroom teaching practices, including use of material, instructional content, student teacher interaction, assessment methods, and administrative oversight. When analyzed together, these instruments create a comprehensive picture of a school's learning environment.

CHAPTER 2: METHODOLOGY

This chapter describes the RFS-SSE sampling process, development of the EGRA assessment tool, and analysis of the quality of the tool.

2.1 RFS-SSE Sampling Design

A stratified cluster random sampling method was used in order to assure that (1) An equal number of boys and girls in urban and rural schools be assessed and (2) that the results of the study be generalizable to the entire population of intervention schools in each of the eight delegations selected for intervention. Schools were first stratified by geographic location and urban/rural. Within schools, we then stratified by gender. The detailed steps of this process are described below.

The project is intervening in a total of 90 schools throughout eight provinces. In each of the eight RFS-SSE provinces, five to seven intervention schools were selected to be a part of the baseline study sample. The following is a description of the process used for the selection of schools and students. A database of the schools in each delegation served as the database from which the school sample was pulled. The Ministry of Education's "Massar" (Education Management Information Systems) EMIS database was used for the student sample.

1. All 1415 schools in each of the 8 delegations were stratified based on location (rural/urban).
2. Schools were deselected if they contained multi-level Grade One or Grade Two classrooms.
3. Between 10 and 12 schools were selected per province to form the pool of 90 intervention schools for the RFS-SSE project.
4. Among the 90 intervention schools, five to seven schools were randomly selected in each region based on location to form the sample of treatment schools for the EGRA and SSME study. Therefore, half (i.e., 45) of the intervention schools are included in the baseline study.
5. For the EGRA and SSME study, an additional five to seven schools per province were selected to form the control group. Therefore, another 45 schools make up the control group.
6. Within each school, a sample of 10 Grade One girls and 10 Grade One boys was selected for a total of 20 Grade One students per school. In schools where fewer than 20 Grade One students were enrolled, all Grade One students present on that day were assessed, resulting in a boy-girl ratio of 52:48. In schools with more than 20 enrolled Grade One students, a list of randomly-selected replacement students in each school was also provided by the Ministry of Education.

2.2 School and Student Sample

School sample

Within each targeted sub-province, schools were randomly selected with equal probability from the MOE's EMIS data base. For each selected school, backup schools were selected. In each school, students randomly selected using the MOE's EMIS system were tested. A backup list of students was also generated for each school. Where insufficient pre-selected students were present on the day of testing, students were randomly selected. The final sample included 1,729 students in 90 schools (control and experimental)

Student sample

The sample of students who participated in the baseline came from four AREFs (Académie régionale d'éducation et de formation) and from two provinces per AREF, making a total of eight provinces. In all, 1,729 students participated in the baseline, 871 in the control group and 858 in the experimental group. Each province had a sub-group of between 186 and 241 students. For each of the eight provinces, about half of the students were part of either the control group or the experimental group.

TABLE 1. NUMBER OF STUDENTS BY AREF (ACADÉMIE RÉGIONALE D'ÉDUCATION ET DE FORMATION) AND BY PROVINCE

AREF	Province	Control Group	Experimental Group	Total
Sous Massa	Inezgane-Aït Melloul	121 (50.2%)	120 (49.8%)	241 (100%)
	Tiznit	113 (47.5%)	125 (52.5%)	238 (100 %)
Kénitra	Témara	100 (50%)	100 (50.0%)	200 (100%)
	Kénitra	100 (50.3%)	99 (49.7%)	199 (100 %)
Oriental	Oujda	120 (50.2%)	119 (49.8%)	239 (100%)
	Figuig	118 (52.4%)	107 (47.6%)	225 (100%)
Fès Meknès	Taounate	101 (50.3%)	100 (49.7%)	201 (100%)
	El Hajeb	98 (52.7%)	88 (47.3%)	186 (100%)
Total		871 (50.4%)	858 (49.6%)	1,729 (100%)

During sample design, we aimed to include an equal number of girls and boys in both the control group and the experimental group. The exact portion of girls and boys tested in each group is presented in the table below. The proportion of girls is slightly higher than the boys in both groups. This is because in some schools, there were fewer than 10 boys and 10 girls. In these schools, all students in the class were tested and there may have been more girls than boys.

TABLE 2. NUMBER OF STUDENTS BY SEX AND GROUP

	Girls	Boys	Total
Control Group	460 (52.8%)	411 (47.2%)	871 (100%)
Experimental Group	439 (51.2%)	419 (48.8%)	858 (100%)
Total	899 (52.0%)	830 (48.0%)	1729 (100%)

In addition, the sample should have included an equal portion of students each from rural areas and urban areas. In reality, a few more students came from urban areas (51%) than from rural areas (49%). This distribution was about the same for the control group as for the experimental group.

TABLE 3. NUMBER OF STUDENTS BY AREA AND BY GROUP

	Rural	Urban	Total
Control Group	433 (49.7%)	438 (50.3%)	871 (100%)
Experimental Group	418 (48.7%)	440 (51.3%)	858 (100%)
Total	851 (49.2%)	878 (50.8%)	1729 (100%)

2.3 EGRA/SSME Tool Development

EGRA Tool Development

In order to evaluate students' reading skills, an Early Grade Reading Assessment (EGRA) was developed in Modern Standard Arabic (MSA). The steps of the development process were as follows:

1. Selection of subtasks to be included in the assessment (the competencies to be assessed);
2. Development of the content (items) for each of the subtasks;
3. Piloting of the tools (EGRA and SSME) in schools located in the provinces targeted for intervention;
4. Analysis of pilot data and determination of changes required to improve the tools; and
5. Finalization of the tools.

Selection of EGRA subtasks

The baseline EGRA tool was developed to measure six core reading competencies in Grades 1 and 2. The same subtasks were used for both grades. Some of the subtasks were timed in order to measure the speed with which students were able to identify graphemes (syllable identification), decode nonwords (phonics), and read connected text (fluency). Below in Table 4 is a summary of the EGRA subtasks and the skills they are designed to assess:

TABLE 4. EGRA SUBTASKS

Core Reading Skills	EGRA Subtasks	Subtask Length
1. Phonemic awareness	1. Phonemic Awareness	Untimed
2. Alphabetic knowledge	2. Syllable Identification	Timed (1 minute)
3. Decoding	3. Nonword reading	Timed (1 minute)
4. Oral Reading Fluency	4a. Passage Reading	Timed (1 minute)
5. Reading comprehension	4b. Reading Comprehension	Untimed
6. Listening comprehension	5. Listening Comprehension	Untimed

Phonemic Awareness: On this untimed subtask, students were asked to orally identify a phoneme (the smallest unit of sound in a word) at the beginning of 10 familiar words.

Syllable Identification: On this timed subtask, students were presented with 100 letter-diacritic combinations and asked to read the syllables in one minute.

Nonword Reading: On this timed subtask, students were presented with 50 nonwords (words invented for this exercise) and asked to read as many of the words as possible using decoding skills in one minute.

Passage Reading: On this timed subtask, students were given one minute to read a passage consisting of 59 words.

Reading Comprehension: On this untimed subtask, students were asked comprehension questions based on the same 59 word passage they read aloud in the Passage Reading subtask.

Listening Comprehension: On this untimed subtask, students were first read a story aloud by the enumerators, then asked five questions orally to assess their understanding of the story's meaning.

Selection of SSME tools

The Snapshot for School Management Effectiveness (SSME) tools yield a rich picture of school management practices through direct classroom observation, inventories, and interviews with key actors in school life. For this study, the following tools were selected:

1. Classroom Observation: Administered in the classroom of one Grade 1 teacher per school during a reading lesson.
2. Classroom Inventory: Administered in the same class as the Classroom Observation.
3. School Director Questionnaire: Administered to the Director of each school visited.
4. Teacher Questionnaire: Administered to the teacher who is observed using the Classroom Observation.
5. Student Questionnaire: Administered to each student who is selected for assessment.
6. Parent Questionnaire: Administered to two parents of Grade 1 students per school.

Tool development

The RFS-SSE team convened a five day adaptation workshop from December 7-11, 2015 at the Centre des Formations et des Rencontres Nationales (CFRN) in Rabat, Morocco to develop two equated versions of the EGRA and SSME tools in Modern Standard Arabic (MSA). This in-country adaptation workshop was held at the start of the test modification process for EGRA instruments. It provided an opportunity to build content validity into the instrument by having government officials, curriculum experts, and other relevant groups examine the EGRA subtasks and make judgments about the appropriateness of each item type for measuring the early reading skills of their students, as specified in curriculum statements or other guidelines for learning expectations or standards.

A total of 19 participants were selected and invited by the Centre National de l'Évaluations et des Examens (CNEE) attend the workshop. Participants began by reviewing the existing EGRA tool in MSA from Morocco to become more familiar with the assessment and how it has been used in the past to the Moroccan context. Under the guidance of workshop facilitators with expertise in EGRA adaptation, the participants then drafted new content for all assessment subtasks. The new content was developed to be at the appropriate reading level for Grade 1 and Grade 2 students.

Similarly, the SSME instrument was streamlined to include items that were of interest to the participants and were adapted to the conditions of Moroccan context and to the needs of the RFS-SSE activity.

During the adaptation workshop, each instrument was pre-tested in 3 schools in Rabat. (These schools were not included in the sample used for final assessment.) The SSME instrument was then reviewed in light of the pretesting experience, any phrasing of questions that led to misunderstandings was clarified, and problematic questions were removed or modified.

Tool piloting

Following the adaptation workshop, the RFS-SSE team piloted the EGRA and SSME tools to determine their validity and reliability before using them for the baseline evaluation. The pilot was conducted with Grade 1 and Grade 2 students in each of the eight provinces involved in the RFS-SSE activity. A total of 16 students were targeted in each school (8 students in Grade 1 and 8 students in Grade 2). Eight provincial representatives who participated in the adaptation workshop were trained to administer the tools. The tools were administered to students in the standard one-to-one format on tablets using Tangerine electronic data capture software. During EGRA tool piloting, the provincial representatives also administered the SSME tools in the same schools. In all, the tools were piloted with 276 Grade 1 and 2 students in 16 schools.

Tool finalization

Following the pilot, the EGRA assessments were then put through rigorous item-level psychometric analyses (using the Rasch model), which helped to identify items that were too difficult or easy, as well as items that were redundant. For EGRA, three reading passages and three listening passages were piloted, and one of each was selected to be included in the final assessment (see [Annex A](#) for

results). The results of this analysis were shared with key CNEE members during a workshop on December 23, 2015 in Rabat. The purpose of this workshop was to review the quality of the assessment tools, as well as to finalize the instrument prior to operational data collection. The EGRA assessment was found to be valid and reliable (see Section 2.6 for more information).

2.4 Data Collection

RFS-SSE conducted a seven-day enumerator training from January 4-10, 2016 in Rabat prior to the baseline data collection. The training participants were selected by the MOE and were equally sourced from each of the eight provinces of the RFS-SSE activity. The enumerators were trained to administer all six EGRA subtasks on tablets using the electronic data capture application Tangerine. A sub-group of candidates were also trained to administer all six SSME tools using Tangerine and to supervise data collection teams.

During enumerator training, the eight provincial representatives who had participated in the adaptation workshop and administered the EGRA and SSME tools pilot acted as co-facilitators and provincial team leaders. The eight provincial representatives were also trained to be Provincial Field Coordinators (PFCs) for operational data collection.

Enumerators were trained on the specific contents of each of the EGRA subtasks, proper administration protocols for each subtask, and the use of tablets and the Tangerine application, among other topics. Enumerator performance was monitored regularly throughout training by the two lead facilitators. Throughout training, facilitators led two sessions to measure consistency of scoring across all enumerators. Enumerator consistency during training ranged from 96-99% agreement depending on the subtask.

Enumerators also experienced two sessions of real-life practice in neighboring schools. During these practice sessions, teams of enumerators practiced administering the EGRA tools with Grade 1 students and the eight PFCs and the training facilitators as well as key members of the CNEE and the RFS-SSE team oversaw the enumerators' performance.

Data Collection

Data collection occurred from January 15-22, 2016 and the team was able to reach approximately 98% of the target number of students. To ensure the quality of data collection, a trained supervisor oversaw each data collection team of two EGRA enumerators. At the end of each day, supervisors and enumerators discussed progress and problems encountered that day. The supervisors verified that each enumerator had completed the correct number of assessments at each school and ensured that a daily data collection report was completed. Each team visited one school per day and were able to reach all schools as planned. The teams uploaded the results to a cloud database each evening of data collection. Those data were reviewed and tallied nightly by the STS EGRA Coordinator and discussed with the PFCs.

2.5 Quality of Assessment Tools

An analysis of the reliability of the six EGRA subtasks allows us to be certain of the validity of the conclusions drawn from the performance of the students on each of the subtasks.

The first analysis consisted of verifying the accuracy or the reproducibility of the scores coming from each of the subtasks. The Cronbach Alpha coefficient allows us to verify the internal consistency of the various subtasks. This index varies between 0 and 1, where a value near 1 indicates that the performance of the students can be easily generalized to how they would perform on the same subtask but with different items. In other words, each subtask represents a sample of items drawn from the total set of possible items. For example, the subtask of phonemic awareness is made up of 10 items,

but the total number of sounds that are possible to recognize is much larger. A Cronbach Alpha coefficient value near 1 allows us to affirm that the performance of a student would have been similar even with a sample of 10 different sounds. For several subtasks (phonemic awareness, syllable identification, and nonword reading), the values of the Cronbach Alpha coefficient are very high. The two comprehension subtasks (reading and listening comprehension) have lower values. The limited number of items in each of these comprehension subtasks may explain these lower values.

TABLE 5. INDEX OF INTERNAL CONSISTENCY BY SUBTASK AND FOR THE ENTIRE EGRA

Subtasks	Number of Items	Cronbach Alpha
Phonemic Awareness	10	0.966
Syllable Identification	100	0.982
Nonword reading	50	0.951
Reading comprehension	5	0.555
Listening comprehension	5	0.655

For each subtask, an item analysis can also identify problematic items that should be withdrawn from the calculation of the score on the subtask. The analyses are derived from two indices: the difficulty of the item (p) and its discrimination (d). The difficulty index represents the proportion of students that gave a correct response. Ideally, this index should not have a value that is too low (item is too difficult) or too high (item is too easy). As for the discrimination index, it represents the correlation between the score on an item and the total score on the subtask. For this index, the values should not be negative or too close to 0. The results by subtask are presented in Annex A.

For the entire set of subtasks, no item appeared problematic. It should be noted that for the timed subtasks (syllable identification and nonword reading), the last items presented values that were particularly low since only a few of the students were able to get to these items within the one minute allowed. The same phenomenon was observed for the reading comprehension subtask. Given that the comprehension questions were linked to the amount of text read from the reading passage, the last questions were attempted by only a very small number of students. The majority of students did not read the entire passage during the one minute that was allotted and therefore these students were not asked all of the reading comprehension questions.

2.6 Data Analysis

The EGRA psychometrician developed a research plan that included the following components 1) reliability estimates of each subtask, 2) subtask and item statistics, 3) mean and grand mean scores (percent correct scores), 4) data plots, 5) timed and untimed subtask scores, and 6) questionnaire results. The EGRA psychometrician used STATA for the analysis. Descriptive analyses and inferential statistical comparisons were conducted for the student scores and the SSME data.

CHAPTER 3: EGRA RESULTS

The results of the baseline study suggest similar levels of performance between the experimental and control groups. Differences between the two groups' EGRA results were small and not found to be statistically significant for any subtask.

Detailed EGRA results are presented below, including the average score by subtask for all students as well as the average score disaggregated by study group (experimental and control), the proportion of students unable to answer a single item correctly by subtask, a comparison of schools in rural and urban areas, and the average score on each subtask disaggregated by gender.

3.1 EGRA Results by Subtask

The RFS-SSE activity has several key indicators directly related to the EGRA subtasks included on the assessment for Grade 1 students. Table 6 reports the baseline data for the relevant indicators for Intermediate Result 1, "Effectiveness of Arabic reading lessons on student reading competencies tested for Grades 1 and 2."¹

TABLE 6. RESULTS OF EXPERIMENTAL GROUP BY INDICATOR

R.1.1 Indicators	1. Average number syllables, words read correctly by the students per minute (fluency)	2. Percent of zero score students (Students unable to give a single correct answer)	3. Average correct answers to comprehension questions read by the student themselves	4. Percent of students who were unable to give a correct answer after reading a comprehension text (score 0)	5. Average correct answer to comprehension questions of a text listened by the students	6. Percent of students who were unable to give a correct answer after listening to a comprehension text (score 0)
Syllable Identification	23.1	16.2				
Nonword Reading	7.1	40.6				
Oral Reading Fluency	5.8	45.9				
Average number correct answers on reading comprehension			.37			
% zero score on reading comprehension				75.6		
Average number correct answers on listening comprehension					1.12	
% zero score on listening comprehension						46.0

¹ Additional details on the RFS-SSE indicators can be found in the Monitoring, Evaluation and Learning Plan updated in March 2016.

Phonemic awareness

The phonemic awareness subtask is made up of 10 items. Since this subtask was not timed, the students had the opportunity to provide a response to all of the items unless a student was not able to correctly answer any of the first five items, then the subtask was stopped early and the student was moved on to the next subtask. On average, the students responded correctly to 3.2 out of 10 items. The students from the control group had an average of 3.3 correct responses, while those from the experimental group had 3.1 correct responses. This difference, however, was not statistically significant ($t(1727)=1.15$; $p=0.249$). The distribution of the number of correct responses for each of the two groups was in the form of a U, which suggests that a good proportion of the students were not able to give a correct answer and that when a student was in fact able to give a correct answer, he/she was generally able to correctly answer all of the items in this subtask.

TABLE 7. MEAN SCORES, PHONEMIC AWARENESS

	n	Mean Score	Standard Deviation	Range
Control Group	871	3.3	4.1	0 – 10
Experimental Group	858	3.1	4.0	0 – 10
Total	1,729	3.2	4.1	0 – 10

FIGURE 2. DISTRIBUTION OF CORRECT RESPONSES BY GROUP, PHONEMIC AWARENESS

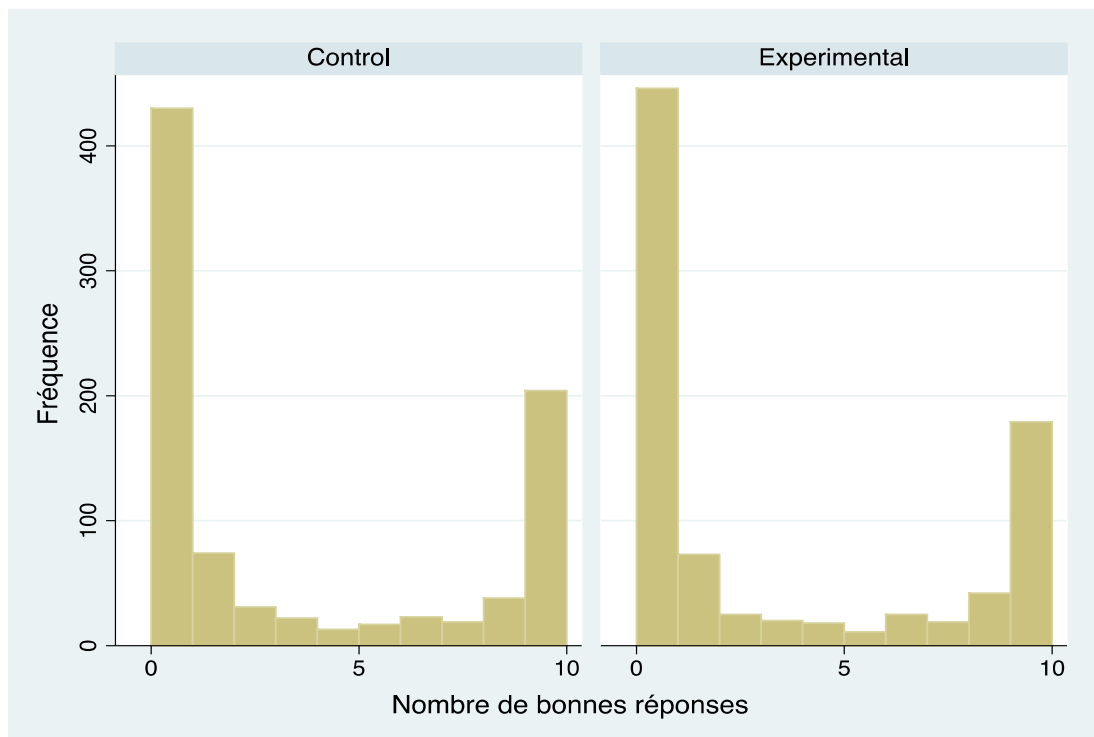


Figure 2 shows that the distribution of students by the number of correct responses was similar for both the control group and the experimental group. About half of the students were unable to give even one correct response in this first subtask (Table 8).

TABLE 8. PROPORTION OF STUDENTS BY NUMBER OF CORRECT RESPONSES BY GROUP, PHONEMIC AWARENESS

Number of Correct Responses	Control Group	Experimental Group	Total
0	430 (49.4%)	446 (51.9%)	876 (50.7%)
1	74 (8.5%)	73 (8.5%)	147 (8.5%)
2	31 (3.6%)	25 (2.9%)	56 (3.2%)
3	22 (2.5%)	20 (2.3%)	42 (2.4%)
4	13 (1.5%)	18 (2.1%)	31 (1.8%)
5	17 (1.9%)	11 (1.3%)	28 (1.6%)
6	23 (2.6%)	25 (2.9%)	48 (4.8%)
7	19 (2.2%)	19 (2.2%)	38 (2.2%)
8	38 (4.4%)	42 (4.9%)	80 (4.6%)
9	74 (8.5%)	75 (8.7%)	149 (8.6%)
10	130 (14.9%)	104 (12.1%)	234 (13.5%)
Total	871 (100%)	858 (100%)	1,729 (100%)

Syllable identification

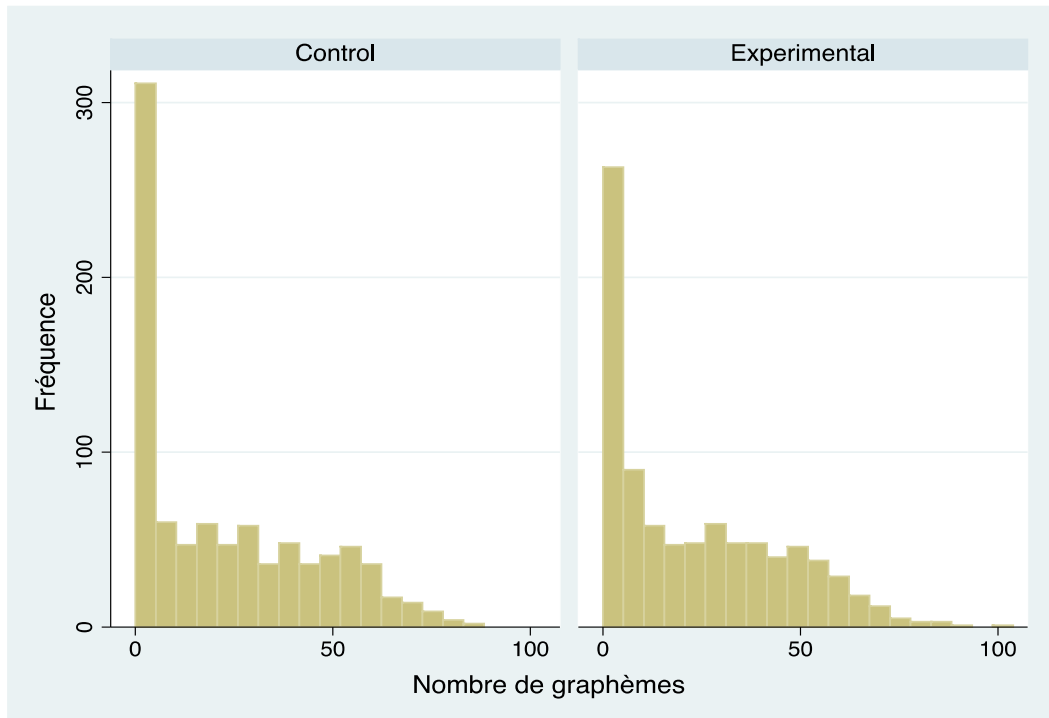
Syllable identification was the second subtask of the EGRA. This subtask was timed and included a total of 100 syllables. If a student did not correctly answer any of the first 10 items, then the subtask was stopped and the student was moved onto the next subtask. In one minute, on average, students were able to correctly read 22.9 syllables. The average performance of the control group (22.9 syllables) was similar to that of the experimental group (23.1 syllables) (see Table 9), and the difference between the two groups was not statistically significant ($t(1727)=-0.18$; $p=0.854$). The distribution of scores on this subtask for each of the groups denotes a positive asymmetry explained by the fact that the majority of the students obtained relatively low scores and that higher performances rarely occurred (see Figure 3).

It should be noted that this is a fluency subtask meaning the scores are presented as number of correct syllables read in one minute. The maximum number of syllables recognized in a minute could go above 100 if a student succeeded in correctly recognizing all 100 syllables of the subtask in *less than* one minute. When this happened ($n=1$), this student's score was adjusted to show the number that he/she would have attained if he/she had taken the entire minute. This occurrence was rare, since 95% of the students correctly recognized 63 or fewer syllables.

TABLE 9. MEAN SCORES, SYLLABLE IDENTIFICATION

	n	Mean Score	Standard Deviation	Range
Control Group	871	22.9	22.2	0 – 87
Experimental Group	858	23.1	21.7	0 – 103.8
Total	1,729	22.9	21.9	0 – 103.8

FIGURE 3. DISTRIBUTION OF CORRECT RESPONSES BY GROUP, SYLLABLE IDENTIFICATION



For the syllable identification subtask, a smaller proportion of students did not succeed in reading a single syllable compared to the almost 50% of students who couldn't provide a correct response on the phonemic awareness subtask. Indeed, only 16.6% (n=287) of the total sample of students were not able to provide a single correct response. In terms of the control and experimental groups, the proportion of students who did not provide a single correct response on the syllable identification subtask was 17.0% (n=148) for the control group and 16.2% (n=139) for the experimental group. These two groups were similar in terms of students who were not able to give a single correct response.

Nonword reading

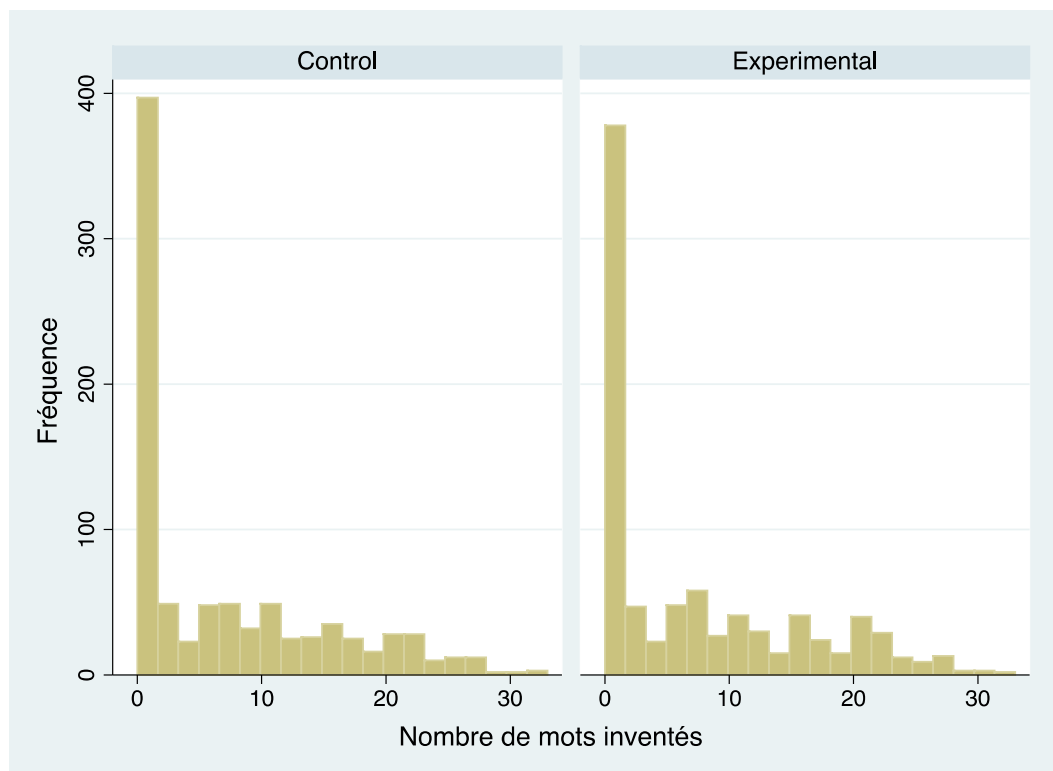
Similar to the syllable identification subtask, the nonword reading subtask was also timed and students were given up to one minute. If a student was unable to correctly read any of the first five nonwords, then the subtask was stopped and the student was moved on to the next subtask. This subtask included a total of 50 nonwords. For this subtask, the student could have read up to 50 total nonwords (or invented words) in one minute. The greatest number of nonwords read correctly was 33. Not a single student was able to read all of the nonwords in one minute. The students were able to correctly read an average of 6.9 nonwords in one minute. The students in the control group read an average of 6.8 words, while those from the experimental group read an average of 7.1 words (see

Table 10). The difference between the two groups was not statistically significant ($t(1727)=-0.84$; $p=0.402$). The distribution of the two groups was a strong positive asymmetry explained by the large proportion of students who were unable to correctly read a single word (see Figure 4).

TABLE 10. MEAN SCORES, NONWORD READING

	n	Average	Standard Deviation	Range
Control Group	871	6.8	8.2	0 – 33
Experimental Group	858	7.1	8.4	0 – 32
Total	1,729	6.9	8.3	0 – 33

FIGURE 4. DISTRIBUTION OF CORRECT RESPONSES BY GROUP, NONWORD READING



Indeed, more than 40.0% of the students were unable to read a single nonword. This proportion was 42.0% (n=366) for the control group and 40.6% (n=348) for the experimental group.

Oral reading fluency

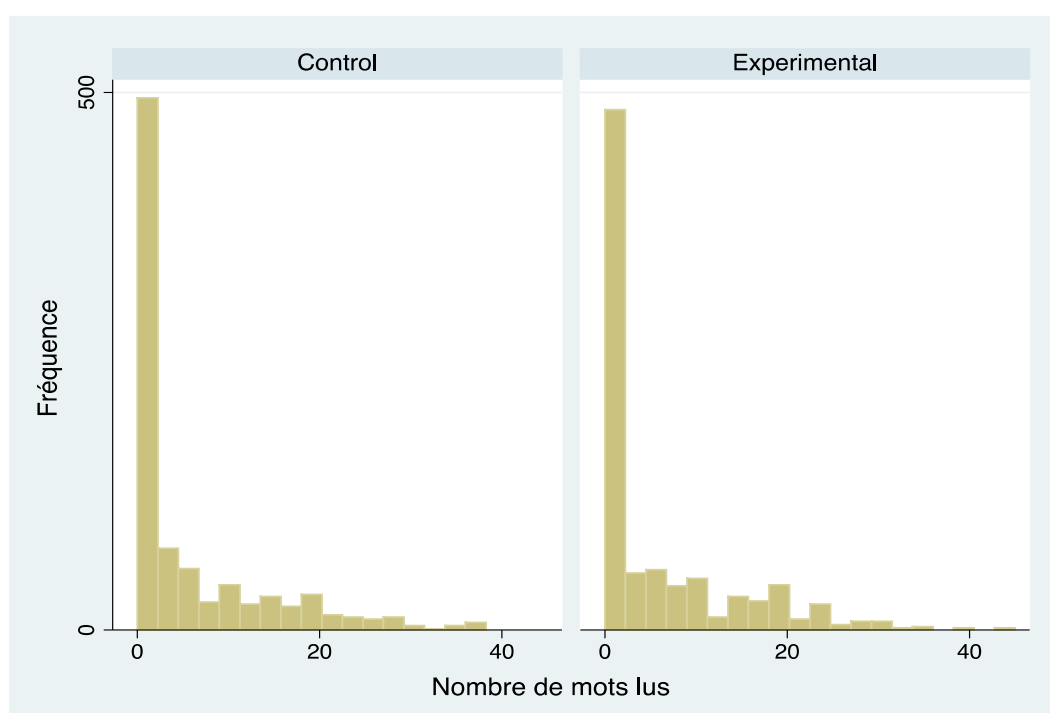
The oral reading passage consisted of 59 words total, and students were given up to one minute to read the entire passage. If a student was unable to correctly read any of the first five words, then the subtask was stopped and the child was moved on to the next subtask. Among all students, the average number of words read correctly (or Correct Words per Minute, CWPM) was 5.6. Students from the control group read, on average, 5.5 CWPM, while those from the experimental group read 5.8 CWPM (see

Table 11). The difference between the two groups was not statistically significant ($t(1727)=-0.669$; $p=0.0504$). No student was able to read all 59 words in a minute. At most, students were able to correctly read 45 words. Similar to the previous subtask, the distribution of scores denoted a strong positive asymmetry (Figure 5). This asymmetry was principally caused by a large proportion of students who could not correctly read one single word.

TABLE 11. MEAN SCORES, ORAL READING FLUENCY

	n	Mean Score	Standard Deviation	Range
Control Group	871	5.5	8.2	0 – 38
Experimental Group	858	5.8	8.3	0 - 45
Total	1,729	5.6	8.3	0 – 45

FIGURE 5. DISTRIBUTION OF CORRECT RESPONSES BY GROUP, ORAL READING FLUENCY



Nearly half of all students (46.2%, n=798) were not able to read a single word correctly in this subtask. The proportion of students not reading a single word was 46.4% (n=404) for the control group and 45.9% (n=394) for the experimental group.

Reading comprehension

This subtask contains a total of five reading comprehension questions about the oral reading passage read by the student in the previous subtask. However, students were not always asked all five of the comprehension questions. Instead, students were asked only the comprehension questions for which they had read far enough in the passage that they could know the answer. For example, the answer to comprehension question #4 may not be revealed in the reading passage until the fourth sentence of that passage. Therefore, if a student only read through the first three sentences, then he/she would only be asked the first three questions. In fact, only one student was able to read the entire passage and was therefore asked all five of the reading comprehension questions and only two students were asked four of the five reading comprehension questions.

The students who were not able to read a single word of the reading passage were not asked any of these questions and a score of 0 was automatically given to them. Fifty-percent of all students (n=863)

were not asked a single comprehension question because they were not able to read far enough into the reading passage to be asked a comprehension question. About three-quarters of the students (76.4%) could not provide a single correct answer to the comprehension questions (this includes the students who were not asked a single comprehension question). This proportion was 77.2% for the control group and 75.6% for the experimental group (see Table 12). On average, students were not able to answer even one question, with an average of 0.37 questions answered correctly (see Table 13). The averages of the students from the two groups were very similar and the difference was not statistically significant ($t(1727)=0,297$; $p=0,767$). At most, students were able to correctly answer a maximum of three questions.

TABLE 12. PROPORTION OF STUDENTS BY NUMBER OF CORRECT RESPONSES BY GROUP, READING COMPREHENSION

Number of Correct Responses	Control Group n(%)	Experimental Group n(%)	Total n(%)
0	672 (77.2%)	649 (75.6%)	1321 (76.4%)
1	111 (12.7%)	127 (14.8%)	238 (13.8%)
2	51 (5.9%)	63 (7.3%)	114 (6.6%)
3	37 (4.3%)	19 (2.21 %)	56 (3.2%)
4	0 (0%)	0 (0%)	0 (0%)
5	0 (0%)	0 (0%)	0 (0%)
Total	871 (100%)	858 (100%)	1,729 (100 %)

TABLE 13. MEAN SCORES, READING COMPREHENSION

	n	Mean Score	Standard Deviation	Range
Control Group	871	0.37	0.78	0 – 3
Experimental Group	858	0.36	0.71	0 – 3
Total	1,729	0.37	0.75	0 – 3

Listening comprehension

Listening comprehension was the last subtask of the EGRA. For this subtask, students were read a short story and then were asked five comprehension questions about the story. The proportion of students who weren't able to answer any of the listening comprehension questions was lower than the reading comprehension subtask but was still rather high. Nearly half of the students (46.4%) did not give a single correct response. This proportion was 46.8% for students from the control group and 46.0% for those from the experimental group (see Table 14). Few students succeeded in providing four correct answers or more. The average number of correct answers was 1.13, and this average was similar in both groups (1.14 in the control group and 1.12 in the experimental group) (see Table 15). The difference was not statistically significant ($t(1727)=0,315$; $p=0,7523$).

TABLE 14. PROPORTION OF STUDENTS BY NUMBER OF CORRECT RESPONSES BY GROUP, LISTENING COMPREHENSION

Number of Correct Responses	Control Group	Experimental Group	Total
0	408 (46.8%)	395 (46.0%)	803 (46.4%)
1	158 (18.1%)	162 (18.9%)	320 (18.5%)

2	147 (16.9%)	162 (18.9%)	309 (17.9%)
3	102 (11.7%)	90 (10.5%)	192 (11.1%)
4	41 (4.7%)	36 (4.2%)	77 (4.5%)
5	15 (1.7%)	13 (1.5%)	28 (1.6%)
Total	871 (100%)	858 (100%)	1,729 (100 %)

TABLE 15. MEAN SCORE, LISTENING COMPREHENSION

	n	Mean Score	Standard Deviation	Range
Control Group	871	1.14	1.34	0 – 5
Experimental Group	858	1.12	1.29	0 – 5
Total	1,729	1.13	1.32	0 – 5

Relationships between the results on the various subtasks

The relationships between the results on the various subtasks of the EGRA were analyzed from the correlations between the scores obtained by the students on each of the subtasks. It is to be noted that all of these correlations were statistically significant ($p < 0.000$). As was expected, these correlations were all positive, indicating that the best students on a subtask were those who performed the best on the other subtasks (see Table 16). In addition, the correlations were particularly strong between the subtasks of syllable identification, nonword reading, and passage reading. Within the set of EGRA subtasks, the listening comprehension subtask presented the weakest correlations in relation to the other subtasks.

TABLE 16. CORRELATION MATRIX BETWEEN THE NUMBER OF CORRECT RESPONSES FOR EACH OF THE EGRA SUBTASKS

	Phonemic Awareness	Syllable Identification	Nonword Reading	Passage Reading	Reading Comprehension	Listening comprehension
Phonemic Awareness	1					
Syllable Identification	0.556	1				
Nonword reading	0.546	0.925	1			
Passage Reading	0.494	0.856	0.908	1		
Reading Comprehension	0.422	0.687	0.726	0.796	1	
Listening Comprehension	0.324	0.399	0.371	0.380	0.447	1

Difference between school location groups

The averages for each of the subtasks were compared as a function of the area in which the school was located. The schools were distributed equally among rural areas and urban areas. Students from urban schools achieved better results on all subtasks compared to students from rural schools. The differences between the averages of the two groups (i.e., urban vs. rural) were all statistically significant. The effect size is a standardized index enabling us to evaluate the size of the differences between the groups. The students from urban schools had higher scores than those from rural areas in syllable identification, nonword reading and reading comprehension (see Table 17). The difference in scores is smaller for the phonemic awareness subtask although the effect size suggests a difference

judged to be medium (threshold is determined using the Cohen suggestion, .2=small, .5=medium, .8=large).

TABLE 17. PERFORMANCE PER SUBTASK BY TYPE OF SCHOOL LOCATION

	Rural		Urban		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.2	3.6	4.1	4.3	<0.000	0.48
Syllable Identification	13.6	17.4	32.0	22.1	<0.000	0.93
Nonword Reading	3.7	6.2	10.1	8.8	<0.000	0.84
Passage Reading	2.6	5.6	8.5	9.3	<0.000	0.76
Reading Comprehension	0.1	0.5	0.6	0.9	<0.000	0.63
Listening Comprehension	0.8	1.1	1.5	1.4	<0.000	0.55

3.2 EGRA Results by Gender

The sample was equally divided between girls and boys. For all subtasks, the boys performed equal to or higher than the girls. However, the difference was statistically significant for only three of the six subtasks: syllable recognition, nonword reading, and passage reading. It should be noted that the values of the effect size for these three subtasks indicate that these differences are, however, quite small. As for the other three subtasks, the averages for girls and boys are practically identical. So, while boys' scores were generally higher overall on the subtasks, the standard deviation between the two groups was rather small (see Table 18). Additionally, the difference between boys and girls was similar regardless of the urbanicity of their school (whether their school is in an urban or rural location), with the exception of the passage reading subtask for which the difference between the genders is larger in urban schools than in rural (see annex C).

TABLE 18. PERFORMANCE BY SUBTASK BY GENDER

	Girls		Boys		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.1	4.0	3.3	4.1	0.290	0.05
Syllable Identification	21.5	21.1	24.5	22.7	0.005	0.14
Nonword Reading	6.3	7.8	7.7	8.7	<0.000	0.17
Passage Reading	5.1	7.8	6.2	8.7	0.006	0.13
Reading Comprehension	0.3	0.7	0.4	0.8	0.145	0.07
Listening Comprehension	1.1	1.3	1.2	1.3	0.225	0.06

CHAPTER 4: CONTEXTUAL VARIABLES AND EGRA PERFORMANCE

4.1 Student Questionnaire

The student's questionnaire dealt with school and family variables that could potentially have an influence on the student's success at school. Descriptive analyses of the SSME data will be presented in this section and the relationships between those data and the performance on the various subtasks of the EGRA will also be studied. Only variables presenting a statistically significant relationship with performance on the EGRA will be presented in a more detailed manner.

Language spoken at home

In the sample, 76.6% (n=1,314) of students said that Darija was the language spoken at home. There was little difference in performance between the students who spoke Darija and Amazigh at home. The students who spoke Amazigh had a significantly higher average than the students who spoke Darija on the phonemic awareness subtask. However, the difference between the two groups is of little importance because the effect size is small (.16). On the listening comprehension subtask, it was the students who spoke Darija at home that had a higher average. The difference on this subtask was more important than the phonemic awareness subtask with an effect size of .31 (see Table 19).

TABLEAU 19. PERFORMANCE BY SUBTASK ACCORDING TO THE LANGUAGE SPOKEN AT HOME

	Darija		Amazigh		p	Effect size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.1	3.9	3.7	4.3	0.006	0.16
Syllable Identification	22.6	22.0	24.6	21.5	0.104	0.09
Nonword Reading	6.8	8.3	7.5	8.1	0.131	0.08
Passage Reading	5.6	8.3	5.8	8.2	0.774	0.02
Reading Comprehension	0.4	0.8	0.4	0.7	0.957	0.00
Listening Comprehension	1.2	1.3	0.8	1.2	<0.000	0.31

Repeating of the grade

Among the students in the sample group, 13.5% (n=232) said that they were repeating their first grade year. It appears that these students performed, on average, lower on the EGRA subtasks than the students who were not repeating. For the entire set of six subtasks, the difference between the averages of the students from the two groups is statistically significant. The analysis of the effect size suggests that this difference is particularly important in the subtasks of syllable recognition, reading of nonwords, and passage reading (see

Table 20). For these three subtasks, the values of the effect size represent a difference of medium importance. The listening comprehension subtask is the one for which the difference between the two groups was smallest.

TABLE 20. PERFORMANCE BY SUBTASK ACCORDING TO THE REPEAT STATUS OF THE STUDENT

	Non-repeaters		Repeaters		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.4	4.1	1.9	3.3	<0.000	0.36
Syllable Identification	24.6	22.4	13.2	15.5	<0.000	0.53
Nonword Reading	7.6	8.5	3.3	5.5	<0.000	0.53
Passage Reading	6.2	8.6	2.2	4.7	<0.000	0.50
Reading Comprehension	0.4	0.8	0.1	0.4	<0.000	0.39
Listening Comprehension	1.2	1.3	0.9	1.2	<0.000	0.15

Means of transportation for getting to school

The majority of students (89.7%, n=1,543) said that they come to school on foot. Also, it appears that the bus is a means of transportation that is rarely used. Most of the time, the students are accompanied by one of their parents (42.6%, n=732) or one of their siblings (21.9%, n=378). About one student out of five gets to school alone (see Table 21 and Table 22).

TABLE 21. MEANS OF TRANSPORTATION FOR GETTING TO SCHOOL

	n (%)
On foot	1543 (89.7%)
By bicycle	71 (4.1%)
By car	64 (3.7%)
Other	30 (1.7%)
By bus	12 (0.7%)
Total	1,720 (100%)

TABLE 22. PERSON WHO ACCOMPANIES THE STUDENTS ON THE WAY TO SCHOOL

	n (%)
Mother/Father	732 (42.6%)
Sister/Brother	378 (21.9%)
Alone	351 (20.4%)
Friends	201 (11.7%)
Other	59 (3.4%)
Total	1,721 (100%)

Reaction of the teacher to correct and incorrect responses from the student

A large majority of the students (98%) have an exercise book. In the majority of cases, when the student gave a correct answer, the teacher would react positively to the student's work. As for the reaction of the teacher when the student gave an incorrect answer in his/her exercise book, in 62.9%

of the cases (n=1,016), the teacher would correct the errors, and in 37.1% (n=598) of the cases he/she would reprimand the student or would do nothing (see Table 23 and Table 24).

TABLE 23. REACTION OF THE TEACHER TO A CORRECT RESPONSE IN THE EXERCISE BOOK

	n (%)
Reacted positively to the work	1481 (93.5%)
Nothing/Gift/Other	103 (6.5%)
Total	1,584 (100%)

TABLE 24. REACTION OF THE TEACHER TO AN INCORRECT ANSWER IN THE EXERCISE BOOK

	n (%)
Corrects the errors	1016 (62.9%)
Nothing/Reprimands/Other	598 (37.1%)
Total	1,614 (100%)

The performance of the students on the EGRA subtasks appear to be higher when the teacher corrects the student's errors in his/her exercise book (see Table 25). For the entire set of subtasks, the difference between the averages of the two groups is statistically significant, but the values of the effect size suggest that these differences are rather small.

TABLE 25. PERFORMANCE BY SUBTASK ACCORDING TO THE REACTION OF THE TEACHER TO AN INCORRECT RESPONSE IN THE EXERCISE BOOK

	Nothing/Reprimands/Other		Corrects errors		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.8	3.9	3.6	4.2	<0.000	0.21
Syllable Identification	20.4	21.3	25.5	22.3	<0.000	0.23
Nonword Reading	5.9	7.7	7.9	8.6	<0.000	0.25
Passage Reading	4.6	7.6	6.6	8.7	<0.000	0.25
Reading Comprehension	0.3	0.6	0.5	0.8	<0.000	0.27
Listening Comprehension	1.0	1.2	1.3	1.4	<0.000	0.18

Help at home for assignments

Nearly three-quarters of the students (72.4%, n=1,245) said that they received help in doing assignments at home. The performance of students was similar whether they received help at home on their assignments or not. Among the students who said they received help, most of the time the help came from one of their siblings (51.7%, n=643) or one of their parents (42.1%, n=523) (see Table 26).

TABLE 26. PERSON WHO PROVIDES HELP FOR THE STUDENT IN COMPLETING ASSIGNMENTS AT HOME

	n (%)
Brother/Sister	643 (51.7%)
Mother/Father	523 (42.1%)

Other	75 (6.0 %)
Grandparents	2 (0.2%)
Total	1243 (100%)

Nutrition

A large proportion of the students (85.9%, n=1,484) said that they had eaten a meal (breakfast or lunch) before coming to school. For the entire set of EGRA subtasks, the performance of students who said that they had eaten before coming to school was higher than those who said they had eaten nothing (see Table 27). The difference between the averages of the two groups for each of the subtasks was statistically significant. For most of the subtasks, the values of the effect size suggest that this difference is of medium importance. The comprehension subtasks were the two subtasks for which the differences between the two averages was smallest.

TABLE 27. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT HAD BREAKFAST/LUNCH BEFORE COMING TO SCHOOL

	Did not have breakfast/lunch before school		Had breakfast/lunch before school		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.1	3.5	3.4	4.1	<0.000	0.32
Syllable Identification	15.4	20.0	24.2	21.9	<0.000	0.41
Nonword Reading	4.4	7.0	7.4	8.4	<0.000	0.36
Passage Reading	3.3	6.0	6.0	8.5	<0.000	0.33
Reading Comprehension	0.2	0.5	0.4	0.8	<0.000	0.25
Listening Comprehension	0.9	1.2	1.2	1.3	0.001	0.22

The proportion of students having the advantage of a lunchroom at school was 38.5% (n=663). The averages of students who had a lunchroom at school were lower than those who did not have this service for the entire set of EGRA subtasks (see Table 28). The differences between the two groups were particularly significant for the subtasks of syllable recognition, reading and passage reading. Phonemic awareness was the subtask for which the difference was the smallest.

TABLE 28. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT HAS A LUNCHROOM AT SCHOOL

	No lunchroom		Lunchroom		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.4	4.1	2.9	3.9	0.005	0.14
Syllable Identification	25.6	22.8	18.6	19.7	<0.000	0.32
Nonword Reading	7.9	8.7	5.4	7.2	<0.000	0.32
Passage Reading	6.7	9.0	3.9	6.5	<0.000	0.34

Reading Comprehension	0.4	0.8	0.2	0.6	<0.000	0.27
Listening Comprehension	1.3	1.4	0.9	1.2	<0.000	0.26

To the question “Where do you have lunch?”, 17.9% (n=308) of the students said they brought it and 2.2% (38) said that they bought it. The majority of students (79.8%, n=1,366) provided the answer “Other.” An analysis of the students’ responses revealed that a large majority of the students indicated that they ate at home. It seems that the wording of the question might have created confusion for the students.

Absenteeism and tardiness

About 41% (n=703) of the students had been absent during the week preceding the EGRA test. It appears that these students did not do as well on the various subtasks than those who had been present (see Table 29). The difference between the two groups was statistically significant for the entire set of subtasks, with the exception of the listening comprehension subtask. However, these differences appear to be of little importance.

TABLE 29. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT HAD BEEN ABSENT FROM SCHOOL OR NOT

	Not Absent		Absent		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.6	4.2	2.7	3.9	<0.000	0.22
Syllable Identification	25.3	22.2	19.7	21.2	<0.000	0.26
Nonword Reading	7.9	8.5	5.6	7.7	<0.000	0.28
Passage Reading	6.5	8.6	4.5	7.7	<0.000	0.24
Reading Comprehension	0.4	0.8	0.3	0.7	0.005	0.14
Listening Comprehension	1.2	1.3	1.1	1.3	<0.000	0.08

Among the students who said they had been absent from school during the week before the EGRA test, most of the students (67.7%, n=475) claimed it was because they had been ill. The other reasons given were all of marginal proportions (see Table 30). Illness was also the most common reason cited by parents when they were asked why their children missed school (see Table 71).

TABLE 30. REASONS GIVEN BY THE STUDENT FOR HAVING BEEN ABSENT FROM SCHOOL

	n (%)
Illness	475 (67.7%)
Other	162 (23.1%)
Woke up too late	17 (2.4%)
Had to take care of an ill member of the family	16 (2.3%)
Bad weather	15 (2.1%)
Mistreated by the teacher or by other students	13 (1.9%)
School is not important	11 (1.6%)
Had to look after brothers/sisters	8 (1.1%)

Market day	7 (1.0%)
School is too hard	6 (0.9%)
No way to get to school	5 (0.7%)
Had nothing to eat	3 (0.4%)
Did not have a school uniform	3 (0.4%)
School is not safe	2 (0.3%)
Working at home	0 (0%)

A little more than one-third of the students (34.1%, n=582) said they had been late to school during the week preceding the EGRA test. For the entire set of subtasks, the average of the students who had not been late was higher than those of students who had been late (see Table 31). The difference between the two groups was statistically significant for all of the subtasks. However, the values of the effect size suggest that the differences between the averages of the two groups were not very large.

TABLE 31. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT HAD BEEN LATE TO SCHOOL OR NOT

	Not late to school		Late to school		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.4	3.1	2.8	3.9	0.002	0.16
Syllable Identification	24.8	22.6	19.6	20.1	<0.000	0.24
Nonword Reading	7.6	8.6	5.8	7.6	<0.000	0.22
Passage Reading	6.2	8.6	4.7	7.7	<0.000	0.18
Reading Comprehension	0.4	0.8	0.3	0.7	0.003	0.15
Listening Comprehension	1.2	1.3	1.0	1.3	0.014	0.13

Among the students who came late to school, half of them claimed that the reason for their tardiness was that they woke up late (see Table 32).

TABLE 32. REASONS GIVEN BY THE STUDENT FOR ARRIVING LATE TO SCHOOL

	n (%)
Woke up late	291 (50.0%)
Other	159 (27.3%)
Illness	31 (5.3%)
No way to get to school	29 (4.9%)
Didn't have school uniform	17 (2.9%)
Bad weather	16 (2.8%)
Had to look after brothers/sisters	9 (1.6%)
Work at home	6 (1.0%)
Taking care of an ill member of the family	3 (0.5%)
Mistreated by the teacher or other students	0 (0%)

Status of literacy at home

A very small proportion of students (4.3%, n=75) claimed that no one knows how to read at home. Nearly two-thirds of the students (65.9%, n=1,139) said that their father knows how to read, 57.5%

(n=993) said that their mother knows how to read, and 74.9% (n=1,294) said that their brother or sister knows how to read.

The fact that an individual in the household knows how to read is linked to the performance of the student on the EGRA subtasks. Indeed, the average of students who said that one person at home knew how to read was higher for the entire set of subtasks (see Table 33). All the differences in the averages were statistically significant. In addition, the values of the effect size suggest that for the entire set of subtasks, the differences are of medium importance.

TABLE 33. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT SAID THAT NO ONE IN THE HOUSEHOLD KNEW HOW TO READ

	Someone		No one		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.3	4.1	1.7	3.4	0.001	0.38
Syllable Identification	23.5	22.0	11.5	15.2	<0.000	0.55
Nonword Reading	7.2	8.3	2.9	5.4	<0.000	0.52
Passage Reading	5.8	8.3	2.2	5.4	<0.000	0.44
Reading Comprehension	0.4	0.8	0.1	0.3	0.001	0.38
Listening Comprehension	1.1	1.3	0.7	1.1	0.007	0.32

The fact that a parent knows how to read is also related to the performance of the student, and this is true whether it was the father or the mother (see Table 34 and

Table 35). In all cases, the students who claimed that their father or mother knew how to read obtained a higher average on all the subtasks, and the differences were all statistically significant. The effect size of the difference between the two groups is similar if one considers the father or the mother. When it was a question of brothers or sisters, no statistically significant difference was noted between those who said that their siblings knew how to read and those who said that they did not.

TABLE 34. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT SAID HIS/HER FATHER KNEW HOW TO READ

	Father cannot read		Father can read		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.5	3.8	3.5	4.2	<0.000	0.26
Syllable Identification	15.9	19.1	26.6	22.4	<0.000	0.50
Nonword Reading	4.3	6.7	8.4	8.7	<0.000	0.50
Passage Reading	3.3	6.4	6.8	8.8	<0.000	0.44
Reading Comprehension	0.2	0.5	0.5	0.8	<0.000	0.40
Listening Comprehension	0.9	1.2	1.2	1.4	<0.000	0.20

TABLE 35. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT SAID HIS/HER MOTHER KNEW HOW TO READ

	Mother cannot read		Mother can read		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.5	3.7	3.7	4.2	<0.000	0.31
Syllable Identification	15.9	18.9	28.1	22.5	<0.000	0.58
Nonword Reading	4.5	6.8	8.8	8.8	<0.000	0.55
Passage Reading	3.3	6.4	7.4	9.0	<0.000	0.50
Reading Comprehension	0.2	0.5	0.5	0.8	<0.000	0.43
Listening Comprehension	0.9	1.2	1.3	1.4	<0.000	0.37

More specifically, the performance of students on the majority of EGRA subtasks was higher when both parents know how to read compared to neither parent or only one parent knowing how to read. Similarly, the performance of students who said that neither parent could read was statistically lower than those students who said that at least one of their parents could read (see Table 36).

TABLE 36. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT SAID HIS/HER MOTHER AND/OR FATHER KNEW HOW TO READ

	Neither parent	Father only	Mother only	Both parents
Phonemic Awareness	2.0*	3.0	3.5	3.8
Syllable Identification	12.7*	19.8	22.8	29.4**
Nonword Reading	3.2*	5.9	7.8	9.4**
Passage Reading	2.4*	4.4	7.6	7.9**
Reading Comprehension	0.1*	0.3	0.3	0.5**
Listening Comprehension	0.8	0.9	1.3**	1.4**

* Average is statistically lower than the other groups

**Average is statistically higher than the other groups

Most of the students (84.7%, n=1,457) who were questioned claimed that there was at least one book in Arabic at home. The students who said they had one book in Arabic at home obtained a higher average on each of the subtasks compared to those who did not have a book in Arabic at home (see Table 37). The difference between the averages on each subtask was statistically significant. The values of the effect size indicate that these differences were of small to medium importance.

TABLE 37. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT HAD A BOOK IN ARABIC AT HOME

	Does not have Arabic book at home		Has Arabic book at home		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.1	3.7	3.4	4.1	<0.000	0.30
Syllable Identification	15.7	19.8	24.3	21.9	<0.000	0.40
Nonword Reading	4.5	7.1	7.4	8.4	<0.000	0.36
Passage Reading	3.3	6.8	6.0	8.4	<0.000	0.33
Reading Comprehension	0.2	0.5	0.4	0.8	<0.000	0.29
Listening Comprehension	0.9	1.2	1.2	1.3	<0.000	0.23

Only 28.7% (n=494) of the students said that they had magazines or newspapers at home. Those who did obtained a better performance on the various EGRA subtasks than those who had no reading material at home (see Table 38). All the differences between the two groups were statistically significant. The values of the effect size suggest that these differences are of small or medium importance. Phonemic awareness had the smallest effect size.

TABLE 38. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT HAS BOOKS/MAGAZINES/NEWSPAPERS AT HOME

	Does not have books/magazines/newspapers at home		Does have books/magazines/newspapers at home		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.9	3.9	3.8	4.2	<0.000	0.20
Syllable Identification	20.0	20.7	30.3	23.2	<0.000	0.48
Nonword Reading	5.8	7.6	9.8	9.2	<0.000	0.49
Passage Reading	4.4	7.2	8.6	9.9	<0.000	0.52
Reading Comprehension	0.3	0.6	0.6	0.9	<0.000	0.45
Listening Comprehension	1.0	1.2	1.4	1.4	<0.000	0.31

Finally, one-third of the students (32.8%, n=566) said that someone reads with them at home. These students obtained a higher average on the entire set of the subtasks than the students who did not read with anyone at home (see Table 39). The difference between the averages of the two groups was statistically significant for most of the subtasks, with the exception of the phonemic awareness subtask. However, the values of the effect size suggest that the differences between the averages were of small importance.

TABLE 39. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER SOMEONE READS WITH THE STUDENT AT HOME

	Someone does not read at home		Someone does read at home		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.1	4.0	3.4	4.2	0.096	0.09
Syllable Identification	21.5	21.5	25.8	22.4	<0.000	0.19
Nonword Reading	6.5	8.1	8.0	8.6	<0.000	0.19
Passage Reading	5.0	7.8	6.9	8.9	<0.000	0.24
Reading Comprehension	0.3	0.7	0.5	0.8	<0.000	0.24
Listening Comprehension	1.1	1.3	1.3	1.4	<0.000	0.18

Family living conditions

Concerning the resources of the households where the students live, the majority have access to electricity (96.4%, n=1,665), to television (95.3%, n=1,647) or to a refrigerator (90.2%, n=1,559). Most of them also have a toilet in the house (79.9%, n=1,380) or a washing machine (63.3%, n=1,094). The resources that were least available were a computer (25.5%, n=440), a car (25.3%, n=437) and Internet service (13.9%, n=241) (see Table 40).

TABLE 40. NUMBER OF STUDENTS CLAIMING THAT THE FOLLOWING RESOURCES WERE AVAILABLE AT HOME

	n (%)
Electricity	1665 (96.4%)
Television	1647 (95.3 %)
Refrigerator	1559 (90.2%)
Toilet inside the house	1380 (79.9%)
Washing machine	1094 (63.3%)
Computer	440 (25.5%)
Car	437 (25.3%)
Toilet outside the house	284 (16.4%)
Latrine	243 (14.1%)
Internet	241 (13.9%)
Don't know/No answer	3 (0.2 %)

The total number of resources available at home is related to the performance of the student on the EGRA subtasks. The positive values of the correlations obtained indicate that the more resources the

student has available to him at home, the higher his/her performance will be on the various subtasks (see

Table 41). All the correlations obtained were statistically significant. The numerical values indicate that the correlation between resources and performances is however rather weak.

TABLE 41. CORRELATION BETWEEN THE NUMBER OF RESOURCES AT HOME AND THE PERFORMANCE ON EACH SUBTASK

	Correlation
Phonemic Awareness	0.236
Syllable Identification	0.338
Nonword Reading	0.326
Passage Reading	0.299
Reading Comprehension	0.284
Listening Comprehension	0.230

The vast majority of students (95.8%, n=1,656) have water at home. Students not having water at home (4.2%, n=72) presented with lower performance averages compared to the other students for the entire set of EGRA subtasks (see Table 42). All the differences between the averages of the two groups were statistically significant. In addition, the values of the effect size suggest that the differences between the two groups were of medium importance. Phonemic awareness had the smallest effect size.

TABLE 42. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT HAD WATER AT HOME

	No water at home		Water at home		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	1.7	3.1	3.3	4.1	0.002	0.38
Syllable Identification	11.0	13.6	23.5	22.1	<0.000	0.57
Nonword Reading (Coding)	2.5	4.0	7.2	8.4	<0.000	0.57
Passage Reading	1.4	3.0	5.8	8.4	<0.000	0.53
Reading Comprehension	0.1	0.2	0.4	0.8	<0.000	0.43
Listening Comprehension	0.4	0.9	1.2	1.3	<0.000	0.58

In most situations, when water is available, it comes from a tap (77.2%, n=1,281). Among the other sources, a well is also a source named more frequently than any others although only 12.9% (n=214) claim that the water available at home comes from this source (see Table 43).

TABLE 43. SOURCE OF THE WATER AVAILABLE AT HOME

	n (%)
Tap	1281 (77.2%)
Well	214 (12.9%)
Cistern	88 (5.3%)
Ain	66 (3.9%)
Other	10 (0.6%)
Total	1659 (100%)

In general, students live in a house with both parents as well as brothers and sisters. In addition, a quarter of the students live in the same house as their grandparents (see Table 44).

TABLE 44. PERSONS LIVING WITH STUDENT IN SAME HOSEHOLD

	n (%)
Father	1703 (98.6%)
Brothers/Sisters	1635 (94.6%)
Mother	1622 (93.9%)
Grandparents	457 (26.5%)
Others	232 (13.4%)

Nearly half of the students (46.7%, n=802) have to do chores outside of school hours. For most of these students, these chores consist of household chores. Students having to work outside of school hours presented with lower averages than those who did not need to work (see Table 45). However, the difference in averages between the two groups was not statistically significant for the passage reading subtask as well as for both reading and listening comprehension subtasks. As for the subtasks showing a statistically significant difference, the values of the effect size suggest that the differences are not very important.

TABLE 45. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER THE STUDENT HAD TO DO CHORES OUTSIDE OF SCHOOL HOURS

	No chores outside the home		Chores outside the home		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.5	4.2	2.9	3.9	0.006	0.13
Syllable Identification	24.2	22.1	21.6	21.7	0.018	0.11
Nonword Reading	7.4	8.4	6.5	8.1	0.016	0.11
Passage Reading	5.9	8.2	5.3	8.3	0.148	0.07
Reading Comprehension	0.4	0.8	0.3	0.7	0.058	0.09
Listening Comprehension	1.2	1.3	1.1	1.3	0.416	0.04

TABLE 46. TYPES OF CHORES ASSIGNED TO STUDENTS

	n (%)
Household	539 (67.9%)
Buying what the household needs	170 (21.4%)
Work in the fields	85 (10.7%)

Total	794
--------------	------------

4.2 Teacher Questionnaires

The teacher's questionnaire dealt with his/her personal characteristics, educational practices, and perception of the school environment. One teacher per school responded to the questionnaire. This section presents a descriptive analysis of the responses given by these teachers and the average performance of the students was also studied. When the relationships were statistically significant, more detailed results are presented.

Characteristics of the teachers

The sample group of teachers questioned (n=90) was made up of 51.1% men (n=46) and 48.9% women (n=44). Students of female teachers performed, on average, on the various EGRA subtasks at a higher level than those with male teachers (see Table 47). The difference between the two groups was statistically significant on all of the subtasks. The value of the effect size suggests that these differences are of great importance. The phonological awareness subtask was the one that presented with the largest effect size.

TABLE 47. PERFORMANCE OF THE STUDENTS BY SUBTASK, ACCORDING TO THE SEX OF THE TEACHER

	Male Teacher		Female Teacher		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.7	1.7	3.8	1.7	0.002	0.68
Syllable Identification	17.7	12.0	28.6	11.5	<0.000	0.92
Nonword Reading	4.9	4.0	9.0	4.3	<0.000	0.97
Passage Reading	3.7	3.5	7.5	4.5	<0.000	0.93
Reading Comprehension	0.2	0.3	0.5	0.4	0.001	0.76
Listening Comprehension	0.9	0.5	1.4	0.5	<0.000	0.99

Concerning the level of training of the teachers, 62.2% (n=56) possessed a high school diploma and 25.6% (n=23) a bachelor's degree. Finally 91% of the teachers stated that they possessed a professional qualification diploma. Training on pedagogy of teaching reading had been completed by 42.2% (n=38) of the teachers. Among these, 76.3% (n=29) had received this training within the framework of their fundamental training and 23.7% (n=9) within the framework of in-service training.

On average, the teachers questioned (n=88) had 10.1 years of experience (standard deviation=8.7). The number of years of experience varied between 1 year and 37 years.

Educational practices

In the area of educational practices, 100% of the teachers (n=89) had a register to indicate the absences of the students, and 98.9% filled out the register daily. The classes were made up of an average of 31.3 students (standard deviation=10.9). The number of students per class varied between 11 and 71. As for grade repeaters, there was on average 3.2 repeaters per class (standard deviation=3.2). This number varied between 0 and 21 repeaters.

Sixty-eight percent (68.5%) of teachers (n=89) said that they update their lesson plans. Among these teachers, about half update them once per year. The students of the teachers who say that they update their lesson plans had a lower average on the EGRA subtasks than those students whose teachers say that they do not update them (see

Table 48). The difference between the averages of the two groups is statistically significant for all the subtasks. The value of the effect size suggests that these differences are of medium to great importance.

TABLE 48. PERFORMANCE OF STUDENTS BY SUBTASK ACCORDING TO WHETHER THE TEACHER UPDATES THE LESSON PLANS OR NOT

	Does not update lesson plans		Updates lesson plans		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	3.9	1.5	2.9	1.8	0.011	0.59
Syllable Identification	29.3	10.7	20.5	12.8	0.002	0.72
Nonword Reading	8.9	3.9	6.2	4.6	0.007	0.63
Passage Reading	7.5	3.7	4.8	4.5	0.008	0.62
Reading Comprehension	0.5	0.3	0.3	0.4	0.015	0.57
Listening Comprehension	1.3	0.5	1.0	0.6	0.034	0.49

TABLE 49. HOW OFTEN THE LESSON PLANS ARE UPDATED

	n (%)
Once per week	17 (27.9%)
Once per month	4 (6.6%)
Once per year	29 (47.5%)
Once every two years	4 (6.6%)
More than every two years	6 (9.8%)
Missing	1 (1.6%)
Total	61 (100%)

Among the various pedagogical tools used in class by the teacher, almost all of the teachers (97.8%, n=88) said they used a textbook, and 82.2% (n=74) said they used a teacher's guide. The least used pedagogical tools were books, newspapers or magazines (30%, n=27) and technological tools (34.4%, n=31). Many teachers also made use of everyday practical tools (65.5%, n=59).

When the teacher needs help, he/she can turn to several different individuals. Most of the teachers (87.8%, n=79) said they received help by informally discussing with other colleagues. A third of the teachers obtained help from the teaching inspector. Finally, four respondents said that they never needed help, and two claimed that there was no one from whom to request help (see Table 50).

TABLE 50. INDIVIDUAL FROM WHOM THE TEACHER REQUESTED HELP

	n (%)
--	-------

Informal discussions with other teachers	79 (87.8%)
The teaching inspector	30 (33.3%)
Organized meetings with other teachers	25 (27.8%)
The school director	19 (21.1%)
I never need any help	4 (4.4%)
There is no one to ask for help	2 (2.2%)

Teacher evaluations carried out by a teaching inspector occurred for most of the teachers once per year (41.1%) or once every two years (30%). Only two teachers said that they had never been evaluated by an inspector (see Table 51). During these visits, the inspector primarily gave advice on teaching methods (81.1%) and assessment (44.4%). One-third of the teachers said that the inspector only carried out administrative evaluations (see Table 52).

TABLE 51. HOW OFTEN THE INSPECTOR EVALUATES THE TEACHER

	n (%)
Never	2 (2.2%)
Once per month	5 (5.6%)
Once per semester	19 (21.1%)
Once per year	37 (41.1%)
Once every two years or less often	27 (30.0%)
Total	90 (100%)

TABLE 52. AREAS IN WHICH THE INSPECTOR GAVE ADVICE

	n (%)
Teaching methods	73 (81.1%)
Assessment methods	40 (44.4%)
Does not visit me/administrative evaluation	30 (33.3%)
Respect for students	19 (21.1 %)
Other	2 (2.2%)

The methods used to assess the skills acquired by the students were varied. Most of the teachers used oral assessments (95.6%) or written assessments (86.7%). About half of the teachers also used homework (48.9%). Finally, the semester final assessments were used by 60% of the teachers (see Table 53). Oral assessments were used in various ways by the teachers. About half of them said they used this type of assessment to grade the students, to assess their comprehension of the subject or to design educational activities. The most popular use (80.0%) was however to better adapt the course to the needs of the students (see Table 54). Written assessments were also used for the same purpose with 72.2% (n=65) of teachers saying that they use written assessments to better adapt the course to the needs of the students. Additionally, written assessments are also used to grade students and to design education activities, among other uses (see Table 54).

TABLE 53. METHODS USED TO ASSESS STUDENT SKILLS

	n (%)
Oral assessments	86 (95.6%)
Written assessments	78 (86.7%)

Semester final assessment	54 (60.0%)
Assignments done at home	44 (48.9%)
Project work	17 (18.9%)
Other	2 (2.2%)

TABLE 54. WAYS ORAL AND WRITTEN ASSESSMENT IS USED

	Oral Assessment	Written Assessment
	n (%)	n (%)
To better adapt the course to the needs of the students	72 (80.0%)	65 (72.2%)
To grade students	51 (56.7%)	64 (71.1%)
To assess their comprehension of the subject	48 (53.3%)	45 (50.0%)
To design educational activities	46 (51.1%)	49 (54.4%)
Other	3 (3.3%)	1 (1.1%)

School environment

Concerning the involvement of parents in the educational process, only 31.5% (n=28) of the teachers thought that parents were involved. Only 13.3% (n=12) report problems with safety at the school. The problem most often reported was violence towards school personnel (see Table 55).

TABLE 55. TYPE OF SECURITY PROBLEM

	n (%)
Violence towards the school staff	7 (58.3%)
Theft	3 (25.0%)
Vandalism	3 (25.0%)
Violence towards students	2 (16.7%)
Other	0 (0%)

4.3 School Director Questionnaires

The school director of the schools participating in the study were also invited to respond to a questionnaire. The questions dealt with their personal characteristics, the characteristics of the school and of the school environment. The data was available for 88 of the 90 schools in the study. The analyses presented illustrate the responses given to the various questions. As was the case for the other questionnaires, the responses from the school directors were examined in relation to students' performance. When this relationship was statistically significant, more detailed results are presented.

Personal characteristics of the school directors

Among the 88 respondents, only 5 (5.7%, n=5) were women. All the respondents filled the role of school director in their school (there were no assistant school directors surveyed). The respondents (n=88) reported an average of 12.2 years of experience (standard deviation=5.5). Responses to this question varied between 1 year and 23 years of experience. The average performance of the students from each school was compared to the number of years of experience of the school director. All the correlations between the number of years of experience and the performance on a subtask are statistically significant, with the exception of the first subtask (see Table 56). The positive

correlations indicate that the more years of experience that a school director had, the higher the results that the students obtained on the EGRA subtasks. However, we must note that the correlations observed are relatively weak.

TABLE 56. CORRELATION BETWEEN THE NUMBER OF YEARS OF EXPERIENCE AS SCHOOL DIRECTOR AND THE PERFORMANCE ON EACH SUBTASK

	Correlation
Phonemic Awareness	0.179
Syllable Identification	0.317
Nonword Reading	0.283
Passage Reading	0.297
Reading Comprehension	0.325
Listening Comprehension	0.356

The majority of the school directors had completed secondary education (33.0, n=29%) or had obtained a bachelor’s degree (58.0%, n=51). Only one teacher had obtained a master’s degree or equivalent and a few (7.9%, n=7) had other degrees. The majority of respondents (90.9%, n=80) had completed training in school management. Finally, 79.8% (n=67) of the school directors stated that they had not been absent from school for not even one day during the last month (see Table 57).

TABLE 57. NUMBER OF DAYS ABSENT BY THE SCHOOL DIRECTORS DURING THE LAST MONTH

Number of days absent	n (%)
0	67 (79.8%)
1	6 (7.1%)
2	4 (4.8%)
3	2 (2.4%)
4	2 (2.4%)
8	3 (3.6%)
Total	88 (100%)

Characteristics of the school environment

Almost all of the schools include Grade 1 through Grade 6. Only a little more than 10% of the schools also have seventh and eighth grades (see Table 58). Nearly two-thirds of the schools (61.4%) had not been closed outside of school vacations, and the schools that were closed were closed for only one or two days (see Table 59). The schools had an average enrollment of 346.4 students. The number of students in the school varied between 58 and 1130. On the average, there were a few more boys than girls in the schools (see Table 60).

The correlations between the total number of students and performance on the EGRA subtasks was statistically significant for passage reading, and the reading and listening comprehension subtasks (see Table 61). The positive value of these correlations indicate that the higher the number of students enrolled in a school, the higher the students’ average performance. It must, however, be noted that the correlation between student performance and the number of students is relatively weak. The school directors reported that the schools had an average of 11.7 teachers (see Table 62). The total number of teachers varied between 3 and 33. On average, the schools had a few more women teachers than men.

TABLE 58. GRADES TAUGHT IN THE SCHOOL

Grades	n (%)
Grade 1	87 (98.9%)

Grade 2	87 (98.9%)
Grade 3	87 (98.9%)
Grade 4	87 (98.9%)
Grade 5	87 (98.9%)
Grade 6	88 (100%)
Grade 7	10 (11.4%)
Grade 8	11 (12.5%)

TABLE 59. NUMBER OF DAYS CLOSED OUTSIDE OF SCHOOL VACATIONS

	n (%)
0 days	54 (62.8%)
1 day	26 (30.2%)
2 days	6 (7.0%)
Total	86 (100%)

TABLE 60. NUMBER OF STUDENTS IN THE SCHOOL

	n	Average	Standard deviation	Range
Number of boys	88	182.2	116.3	24 – 575
Number of girls	88	164.3	110.8	25 – 555
Total number	88	346.4	225.6	58 - 1130

TABLE 61. CORRELATION BETWEEN THE NUMBER OF STUDENTS IN THE SCHOOL AND PERFORMANCE ON EACH SUBTASK

	Correlation
Phonemic Awareness	0.035
Syllable Identification	0.206
Nonword Reading	0.207
Passage Reading	0.262
Reading Comprehension	0.289
Listening Comprehension	0.270

TABLE 62. NUMBER OF TEACHERS IN THE SCHOOL*

	n	Average	Standard deviation	Range
Men	84	5.5	4.5	0 – 29
Women	84	6.2	5.1	0 – 28
Total number	84	11.7	6.6	3 - 33

*Four school directors claimed to have more than 150 teachers in their school. Their answers were not included in this table.

Most of the time, whenever a teacher was absent, the school director would distribute the students among the other classes (69.3%, n=61). In certain cases (13.6%, n=12), students were sent home (see

Table 63). The absence of teachers was noted in the registers of 96.9% (n=85) of the schools studied. The teachers' lesson plans were checked by 84.1% (n=74) of the school directors. Those who check these lesson plans did it most of the time once per month (47.7%, n=42) or once per week (25%, n=22) (see Table 64). A slightly smaller proportion (87.5%, n=77) of the school directors evaluate the work of the teacher in the class through checking the lesson plans. School directors evaluating work in the classroom did it in general once per month (50.0, n=44%) or once per week (20.5%, n=18) (see Table 65).

TABLE 63. ACTIONS TAKEN WHEN A TEACHER IS ABSENT

	n (%)
Students are distributed into the other classes	61 (69.3%)
Students are sent home	12 (13.6%)
Substitute teacher is sought	8 (9.1%)
The school director takes charge of the students	6 (6.8%)
Other	1 (1.1%)
Total	88 (100%)

TABLE 64. HOW OFTEN THE LESSON PLANS ARE CHECKED

	n (%)
Every day	2 (2.3%)
Once per week	22 (25%)
Once per month	42 (47.7%)
Once per semester	8 (9.1%)
Lesson plans are not checked	14 (15.9%)
Total	88 (100%)

TABLE 65. HOW OFTEN THE WORK OF THE TEACHER IN THE CLASSROOM IS EVALUATED

	n (%)
Every day	3 (3.4%)
Once per week	18 (20.5%)
Once per month	44 (50.0%)
Once per semester	11 (12.5%)
Once per year	1 (1.1%)
Does not evaluate the teacher's work in the classroom	11 (12.5%)
Total	88 (100%)

In dealing with textbooks, nearly a quarter of the schools (23.9%, n=21) did not have enough textbooks at the beginning of the year. Only 29.5% (n=26) of the schools had a library. Among these schools, students borrow books in 80.8% (n=21) of the schools. In addition, when a library was present, it was visited by the students in general once per week (47.6%, n=10) or every day (28.6%, n=6) (see Table 64).

TABLE 66. HOW OFTEN STUDENTS VISITED THE LIBRARY

	n (%)
Once per month	4 (19. %)
Once per week	10 (47.6%)
Every day	6 (28.6%)
Don't know	1 (4.8%)
Total	21 (100%)

A parents' association is present in 84.1% (n=74) of the schools. The parents' association met in general every two or three months (50.0%, n=37) or once per month (21.6%, n=16). It should be noted that 20.3% (n=15) of the school directors stated that their association had not met a single time. In addition, three school directors claimed not to know the frequency of meetings of the parents' association (see Table 67). The parents' associations were particularly active in the discussion of student problems and their solutions (77.0%, n=57), raising funds (50.0%, n=37) and the management of classrooms and school facilities (44.6%, n=33) (see Table 68). Among schools where a parents' association was present, 60.8% (n=45) of the school directors claimed to be satisfied with their contribution.

TABLE 67. HOW OFTEN THE PARENTS' ASSOCIATION HOLDS MEETINGS

	n (%)
Never	15 (20.3%)
Every two or three months	37 (50%)
Once per month	16 (21.6%)
Once per week	3 (4.1%)
Don't know	3 (4.1%)
Total	74 (100%)

TABLE 68. PARENTS' ASSOCIATIONS' AREAS OF ACTIVITY

	n (%)
Discussion of the students' problems and their solutions	57 (77.0%)
Raising money	37 (50.0%)
Management of school classrooms and school facilities	33 (44.6%)
Evaluation of the efforts made to improve the school and its subtasks	26 (35.1%)
Discussion of problems related to the management of the school	25 (33.8%)
Implementation of agreements for the sale or distribution of textbooks	17 (23.0%)
Approval of school policies	14 (18.9%)
Evaluation of the financial status of the school	4 (5.4%)
Discussion of school programs	3 (4.1%)
Other	2 (2.7%)

During the last year, a visit by a teaching inspector as a result of a request was reported by 53.4% (n=47) of the school directors. The inspectors generally observed how the classes were conducted (93.6%, n=44), offered advice on teaching (78.7%, n=37) or on assessment (61.7%, n=29). In addition, they also gave advice on innovations in teaching programs (see Table 69).

TABLE 69. REASONS FOR THE TEACHING INSPECTOR'S VISIT

	n (%)
Sitting in the classroom and observing how it is conducted	44 (93.6%)
Giving teachers advice on teaching	37 (78.7%)
Giving advice on innovations in school programs	34 (72.4%)
Giving advice on student assessment	29 (61.7%)
Checking students' attendance records	27 (57.5%)
Checking students' final exams and the methods of assessing them.	20 (42.6%)
Giving advice on the possibilities of professional advancement	19 (40.4 %)
Giving advice on the organization and the diligence of the students	13 (27.7%)
Checking students' assessment reports	12 (25.5%)
Giving the school director advice on the management of the school	12 (25.5%)
Checking teachers' personnel files	11 (23.4%)
Giving advice on practices dealing with school cleanliness	8 (17.0%)
Making sure that the financial records of the school were available	4 (4.6%)

Finally, 26.1% (n=23) of the schools had a problem with security during the current year. Vandalism (43.5%, n=10), thefts (21.7%, n=5) or violence towards school personnel (17.4%, n=4) were the most common problems. Two school directors (8.7%) noted violence towards students and two school directors (8.7%) noted "other" security problems.

4.4 Parent Questionnaire

The students' parents were also invited to respond to a questionnaire. A total of 172 parents coming from the 90 schools participated in this study. The results presented in this section are drawn from the parent questionnaire. Given the fact that it was not possible to pair the parents to their respective child, it was not possible to compare the parents' responses with the students' performance. Thus, unlike the previous sections, this section will not present any correlation with the performance of the students on the EGRA subtasks. Of the parents who participated in the survey, 53.5% (n=92) of the parents were men and 46.5% (n=80) were women.

On average, students took 13.1 minutes to get to school. The time varied between 1 and 60 minutes (see Table 70).

TABLE 70. NUMBER OF MINUTES REQUIRED TO GET TO SCHOOL

	n	Average	Standard deviation	Range
Number of minutes	172	13.1	12.0	1 – 60

Fifty-nine (59.9, n=103%) of the parents reported that the school had a parents' association. It is to be noted that about one in five parents (19.2%, n=33) said that they were unaware if the school had a parents' association. Among the parents who said that the school had a parents' association, 78.3% (n=65) had the impression that the association contributed to the improvement of the school.

The majority of parents (88.3%, n=143) said that they receive their child's grades. The parents reported that their child was absent from school primarily for reasons of illness (81.9%, n=141) (see Table 71). Illness was also the most common reason cited for missing school when students were asked themselves (see Table 30). Finally, only 8.8% (n=15) of the parents reported that their child had been the victim of violence during the year.

TABLE 71. REASON FOR STUDENTS' ABSENTEEISM

	n (%)
Illness	141 (81.9%)
Weather conditions	24 (13.9%)
Does not get up early enough in the morning	11 (6.4%)
Problems related to transportation	11 (6.4%)
Doesn't want to go to school	11 (6.4%)
Funeral	11 (6.4%)
Market day/preparation	6 (3.5%)
Taking care of an ill family member	5 (2.9%)
Taking care of a sibling	2 (1.2%)
Other	2 (1.2%)
Other household chores	1 (0.6%)
Lack of food	0 (0%)
School uniform not ready	0 (0%)

4.5 Classroom Inventory Chart

Among the various tools, a classroom inventory chart had to be filled out by the test administrator. The data for this chart was thus available for the 90 schools. This section presents the distribution of responses to the different items in this instrument. The data collected were also subject to analysis of the relationship of the different items and the performance of the students in the school. Just as in the previous sections, when a correlation was statistically significant, more detailed results will be presented.

The test administrators noted that an average of 6.3 classrooms were functioning during the observation period (standard deviation=3.8). The number of classrooms functioning varied between 1 and 18. This average number is very close to the number of classrooms which would be functioning normally, or 6.6 classrooms. On average, the classes were made up of 27.9 students. The number of students varied between 8 and 54 (see Table 72). The test administrators observed that on average the classes had more boys than girls. The majority of students see med to have a reading book, an exercise book and a pen or pencil (see Table 73). The averages of students possessing each type of these materials was very close to the total number of students in the class.

TABLE 72. NUMBER OF STUDENTS IN THE CLASS

	n	Average	Standard deviation	Range
Boys	90	14.9	5.9	3 – 36
Girls	90	13.0	4.2	2 – 24

Total	90	27.9	8.8	8 – 54
--------------	-----------	-------------	------------	---------------

TABLE 73. NUMBER OF STUDENTS WHO HAD A READING BOOK

	n	Average	Standard deviation	Range
Number of students who had a reading book	89	26.2	9.0	2 - 51
Number of students who had an exercise book	89	26.3	10.4	0 - 58
Number of students who had a pen or pencil	89	25.4	10.8	0 – 58

As for equipment available in the classroom, most of the classrooms had either a blackboard or a whiteboard, pens or pencils, schedules and posters. The presence of teacher’s guides in the classroom was less frequent. Indeed, the teacher’s guides in Arabic was observed in 76.7% (n=69) of the classes and the mathematics guide in 54.4% (n=49) (see Table 74). The availability of the teacher’s guide in Arabic in the classroom was correlated to the students’ performance on the EGRA subtasks. Indeed, when this guide was available, the students obtained better results than those from classes where the guide was not present (see Table 75). The difference between the averages of the two groups was statistically significant for the entire set of subtasks with the exception of phonological awareness. The values of the effect size suggest that these differences were of medium to great importance.

TABLE 74. EQUIPMENT AVAILABLE IN THE CLASSROOM

	n (%)
Blackboard/whiteboard	86 (95.6%)
Pens/pencils	87 (96.7%)
Schedule/Posters	85 (94.4%)
Teacher’s guide in Arabic	69 (76.7%)
Teacher’s guide in mathematics	49 (54.4%)

TABLE 75. PERFORMANCE BY SUBTASK ACCORDING TO WHETHER A TEACHER’S GUIDE IN ARABIC WAS AVAILABLE OR NOT IN THE CLASSROOM

	Teacher’s guide in Arabic not available		Teacher’s guide in Arabic available		p	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation		
Phonemic Awareness	2.6	1.8	3.4	1.8	0.088	0.43
Syllable Identification	15.8	11.1	24.8	12.9	0.005	0.72
Nonword Reading	4.6	3.7	7.6	4.7	0.009	0.66
Passage Reading	3.4	3.2	6.2	4.7	0.014	0.62
Reading Comprehension	0.2	0.2	0.4	0.4	0.023	0.58
Listening Comprehension	0.9	0.5	1.2	0.6	0.036	0.53

The absence of other books or booklets was noted in 63.3% (n=57) of the classrooms (see Table 76). As for material attached to the walls of the classroom, posters were observed in 93.3% (n=84) of the classrooms and student work was observed in 42.2% (n=38) of the classrooms. Finally, the student desks were for the most part organized in rows (81.1%, n=71) (see Table 77).

TABLE 76. OTHER BOOKS/BOOKLETS AVAILABLE AND ACCESSIBLE IN THE CLASSROOM

	n (%)
None	57 (63.3%)
Between 1 and 4	2 (2.2%)
Between 5 and 9	9 (10%)
Between 10 and 19	8 (8.9%)
Between 20 and 39	13 (14.4%)
More than 40	1 (1.1%)
Total	90 (100%)

TABLE 77. ORGANIZATION OF CLASSROOM SPACE

	n (%)
In rows	73 (81.1%)
In small groups	13 (14.4%)
In a circle	2 (2.2%)
Another way	2 (2.2 %)
Total	90 (100%)

CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

The findings of this study will serve as the baseline for three future evaluations. By comparing the results of future evaluations to this baseline study, we will be able to examine the impact of the RFS-SSE activity on the reading performance levels of students who received the intervention, as measured by the EGRA subtasks. A total of 1,729 students participated in the EGRA at baseline, about half of whom attend schools that will be receiving RFS-SSE activities. The other half of students, the control group, will serve as a comparison going forward in order to consider the natural gains in reading performance that students may make without receiving any intervention. The baseline study found that the current performance levels of all students, both in the experimental and control groups were similar, with no statistically significant differences.

On average, students (in both the experimental and control groups) were able to correctly respond to 3.2 out of 10 items on the phonemic awareness subtask. On the syllable identification subtask, students were able to correctly identify 22.9 syllables within one minute, on average. For nonword reading, students, on average, were able to correctly read 6.9 nonwords in one minute. On the reading passage subtask, students read on average at a rate of 5.6 Correct Words per Minute. Students were, on average, not able to correctly answer a single reading comprehension question, with the average number of questions correctly answered being only .37. Finally, on the listening comprehension subtask, students on average were able to correctly respond to 1.13 out of five questions.

The proportion of students who were unable to provide a single correct response on each subtask was often high. On the phonemic awareness subtask, 50.7% of students were not able to correctly respond to even one of the 10 items. The syllable identification subtask had the lowest proportion of zero score students, with only 16.6% of students not being able to correctly identify at least one syllable in a minute. On the nonword reading subtask, 41.3% of students were not able to correctly read a single nonword. When presented with a reading passage, 46.2% of students were not able to read a single word. Linked to the reading passage subtask, the reading comprehension questions also had a high number of zero scores – 76.4% of students were not able to correctly answer a single reading comprehension question (which includes the students who were not asked a single question because they didn't read enough of the passage to be asked a question). Finally, 46.4% of students were not able to correctly answer a single listening comprehension question.

Survey questions administered to students, teachers, parents, and school directors allowed for an analysis of contextual factors that might be correlated with student performance. The results suggest that students who did not repeat Grade 1 performed higher than those students who did. Additionally, it appears that students who had eaten before coming to school scored higher than those students who did not; however, students who said that they had a lunchroom at their school actually performed lower than students without a lunchroom. Having a parent who knows how to read was correlated with higher EGRA scores compared to students who said they do not have a parent who knows how to read. Having a book in Arabic at home; having books, magazines, or newspapers at home; and the teacher having the teacher's guide available in Arabic in the classroom were all correlated with higher EGRA scores compared to students who did not have access to those materials. Finally, the results suggest that students who had female teachers scored higher than those students who had male teachers. Understanding these contextual factors will help to further examine the impact of the RFS-SSE activities during future evaluations.

ANNEXES

ANNEX A: Reliability of the EGRA Subtask Items

Subtask: Phonemics awareness

Table 1. Indices of difficulty and discrimination for the phonemic awareness subtask

Variable	p	d
Item 1	.42	0.70
Item 2	.29	0.83
Item 3	.32	0.81
Item 4	.32	0.87
Item 5	.34	0.89
Item 6	.19	0.68
Item 7	.31	0.89
Item 8	.33	0.91
Item 9	.30	0.90
Item 10	.32	0.92

Subtask: Syllable identification

Table 2. Indices of difficulty and discrimination for syllable identification

Variable	p	d
Item 1	.67	0.62
Item 2	.58	0.71
Item 3	.44	0.70
Item 4	.47	0.56
Item 5	.65	0.64
Item 6	.61	0.70
Item 7	.42	0.74
Item 8	.55	0.71
Item 9	.74	0.53
Item 10	.60	0.70
Item 11	.54	0.73
Item 12	.65	0.64
Item 13	.62	0.69
Item 14	.52	0.75
Item 15	.41	0.76
Item 16	.63	0.67
Item 17	.50	0.75
Item 18	.43	0.74
Item 19	.33	0.69
Item 20	.51	0.78
Item 21	.50	0.72
Item 22	.47	0.72
Item 23	.24	0.61
Item 24	.48	0.73
Item 25	.40	0.75
Item 26	.43	0.77

Item 27	.44	0.76
Item 28	.29	0.71
Item 29	.40	0.78
Item 30	.27	0.72
Item 31	.39	0.81
Item 32	.32	0.75
Item 33	.36	0.82
Item 34	.36	0.82
Item 35	.35	0.81
Item 36	.34	0.82
Item 37	.32	0.83
Item 38	.31	0.81
Item 39	.30	0.82
Item 40	.30	0.81
Item 41	.28	0.80
Item 42	.27	0.79
Item 43	.26	0.79
Item 44	.22	0.75
Item 45	.24	0.77
Item 46	.23	0.77
Item 47	.21	0.75
Item 48	.21	0.76
Item 49	.18	0.72
Item 50	.19	0.74
Item 51	.15	0.66
Item 52	.14	0.65
Item 53	.13	0.65
Item 54	.13	0.66
Item 55	.12	0.65
Item 56	.11	0.63
Item 57	.10	0.59
Item 58	.10	0.61
Item 59	.09	0.59
Item 60	.08	0.56
Item 61	.07	0.52
Item 62	.07	0.52
Item 63	.06	0.50
Item 64	.06	0.48
Item 65	.05	0.46
Item 66	.05	0.44
Item 67	.05	0.44
Item 68	.04	0.44
Item 69	.04	0.42
Item 70	.03	0.40
Item 71	.04	0.43
Item 72	.04	0.41
Item 73	.03	0.40
Item 74	.03	0.39
Item 75	.03	0.38
Item 76	.02	0.36
Item 77	.02	0.35
Item 78	.01	0.31
Item 79	.01	0.31

Item 80	.01	0.31
Item 81	.01	0.30
Item 82	.01	0.26
Item 83	.01	0.28
Item 84	.01	0.27
Item 85	.01	0.25
Item 86	.01	0.24
Item 87	.01	0.22
Item 88	.01	0.22
Item 89	.01	0.22
Item 90	.01	0.19
Item 91	.01	0.20
Item 92	.01	0.20
Item 93	.01	0.20
Item 94	.01	0.20
Item 95	.01	0.20
Item 96	.01	0.18
Item 97	.01	0.17
Item 98	.01	0.17
Item 99	.01	0.17
Item 100	.01	0.15

Subtask: Nonword reading

Table 3. Indices of difficulty and discrimination for the reading of nonwords

Variable	p	d
Item 1	.46	0.70
Item 2	.42	0.68
Item 3	.41	0.78
Item 4	.49	0.69
Item 5	.45	0.73
Item 6	.33	0.71
Item 7	.31	0.64
Item 8	.35	0.68
Item 9	.33	0.69
Item 10	.34	0.74
Item 11	.27	0.66
Item 12	.27	0.70
Item 13	.21	0.69
Item 14	.20	0.69
Item 15	.21	0.72
Item 16	.19	0.72
Item 17	.19	0.73
Item 18	.15	0.71
Item 19	.16	0.73
Item 20	.15	0.70
Item 21	.14	0.69
Item 22	.12	0.65
Item 23	.10	0.61
Item 24	.10	0.61
Item 25	.08	0.57

Item 26	.07	0.53
Item 27	.06	0.52
Item 28	.05	0.48
Item 29	.03	0.41
Item 30	.02	0.34
Item 31	.01	0.27
Item 32	.01	0.25
Item 33	.01	0.22
Item 34	.01	0.19
Item 35	.01	0.19
Item 36	.01	0.18
Item 37	.01	0.16
Item 38	.01	0.13
Item 39	.01	0.12
Item 40	.01	0.10
Item 41	.01	0.10
Item 42	.01	0.08
Item 43	.01	0.07
Item 44	.00	0.00
Item 45	.00	0.00
Item 46	.01	0.03
Item 47	.00	0.03
Item 48	.00	0.00
Item 49	.00	0.00
Item 50	.00	0.00

Subtask: Reading Comprehension

Table 4. Indices of difficulties and discrimination for reading comprehension

Variable	p	d
Item 1	.08	0.46
Item 2	.20	0.50
Item 3	.06	0.52
Item 4	.01	0.05
Item 5	.01	0.05

Subtask: Listening Comprehension

Table 5. Indices of difficulty and discrimination for listening comprehension

Variable	p	d
Item 1	.38	0.47
Item 2	.21	0.48
Item 3	.33	0.49
Item 4	.10	0.38
Item 5	.09	0.28

ANNEX B: Intra-class Coefficient (ICC)

Table 6. Intra-class coefficients by subtask

Subtask Name	ICC
Phonemic Awareness	0.15
Syllable Identification	0.31
Nonword reading	0.28
Passage Reading	0.26
Reading Comprehension	0.18
Listening Comprehension	0.12

ANNEX C: Supplementary Tables

Table 7. Performance by subtask according to the school location and gender of the student

	Rural		Urban		p
	Girls	Boys	Girls	Boys	
Phonemic Awareness	1.9	2.5	4.1	4.1	0.106
Syllable Identification	12.6	14.6	29.9	34.5	0.172
Nonword reading	3.3	4.1	9.1	11.3	0.084
Passage Reading	2.4	2.9	7.6	9.6	0.047
Reading Comprehension	0.1	0.2	0.5	0.6	0.428
Listening Comprehension	0.7	0.8	1.4	1.5	0.917

ANNEX D: SSME Tools

(Chemonics Drive)