



**USAID**  
FROM THE AMERICAN PEOPLE



# LIBERIA TEACHER TRAINING PROGRAM: ENDLINE ASSESSMENT OF THE IMPACT OF EARLY GRADE READING AND MATHEMATICS INTERVENTIONS

August 2015

Liberia Teacher Training Program:

# ENDLINE ASSESSMENT OF THE IMPACT OF EARLY GRADE READING AND MATHEMATICS INTERVENTIONS

***Prepared for***

USAID/Liberia  
Ministry of Education  
Republic of Liberia

and

FHI 360 Headquarters  
2224 E. N.C. Highway 54  
Durham, NC 27713 USA

***Prepared by***

Simon King, Medina Korda, Lee Nordstrum, Susan Edwards  
RTI International  
3040 Cornwallis Road  
Research Triangle Park, NC 27709 USA

RTI International is a registered trademark and a trade name of Research Triangle Institute.

This publication was produced for review by the United States Agency for International Development and by FHI 360. It was prepared by RTI International for the Liberia Teacher Training Program.

This development of this document was made possible by the support of the American people through the United States Agency for International Development. The contents are the sole responsibility of RTI International and FHI 360, and do not necessarily reflect the views of USAID or the United States Government.

# ACKNOWLEDGMENTS

The authors want to acknowledge with appreciation and respect for the contributions and support of education staff from United States Agency for International Development (USAID)/Liberia, the Liberian Ministry of Education (MOE), and many other counterparts in the implementation of the Liberia Teacher Training Program Phase II (LTTP II) final early grade reading and mathematics assessment in grades 1, 2, and 3.

At USAID/Liberia, we would like to thank Simone Brown for her leadership and skillful coordination of the efforts of the MOE, and Miriam White for her long-term support and supervision of early grade learning efforts in Liberia. We would also like to thank Felicia Doe-Sumah, Assistant Minister for Basic and Secondary Education, and her MOE colleagues who were instrumental in the successful implementation of the Early Grade Reading Assessment (EGRA), the Early Grade Mathematics Assessment (EGMA), and the Snapshot of School Management Effectiveness (SSME) and in mobilizing the Ministry's support, from start to finish. The success of this assessment owes a debt to the MOE staff who participated in training and directly administered the instruments as trained assessors. We would also like to thank the more than two dozen Liberian colleagues who, having served as assessors on all prior assessments dating back to the USAID EGRA Plus: Liberia project from 2008 to 2011, brought valuable field experience to the final data collection.

We would sincerely like to thank the LTTP II senior leadership, Michael Blundell (Chief of Party) in particular, for swift mobilization of resources in support of this assessment and for timely decision making. We would like to pay special thanks to RTI International's Finance, Operations, and Human Resources staff who worked tirelessly to ensure that the assessment teams were deployed for data collection as scheduled. Scott Kipp and Mathew Finholt-Daniel (RTI's Project Managers), with Geanjay Roberts (Liberia Teacher Training Program's [LTTP's] Assessment Coordinator), led the implementation of this assessment. We are all deeply indebted to our colleagues from the Reading and Mathematics Team who trained assessors, worked long hours to ensure that the data collection commenced on time, and provided other support during the entire program.

The project was backstopped by Medina Korda (RTI's Home Office Project Manager) and Roxanne Zerbonia (RTI's Home Office Project Coordinator). Drs. Peggy Dubeck and Yasmin Sitabkhan (RTI's reading experts) developed the reading assessment tools, and Dr. Sitabkhan developed the mathematics assessment tools. RTI

statisticians Simon King and Susan Edwards provided tireless support in cleaning and analyzing the data. Special thanks go to RTI's Information and Communication Technology Team, who supported the development and adaptation of the electronic instruments and the final compilation of the database.

The authors of this report were Simon King, Medina Korda, Lee Nordstrum, and Susan Edwards. The report was reviewed by Dr. Yasmin Sitabkhan, Adolphus Toe, Trokon Wayne, Geanjay Roberts, and Patience Suah. Simon King coordinated the report-writing process. We would also like to thank Erin Newton for editing this report and Felice Sinno-Lai and Gail Hayes for formatting it.

Finally, we would like to thank all of the principals, teachers, and pupils who participated in this study and to the program overall. To all those who have not been mentioned here, we would like to sincerely thank you all for your contributions to this assessment and to the program implementation overall.

# CONTENTS

	PAGE
Acknowledgments .....	iii
List of Figures .....	vii
List of Tables .....	ix
Abbreviations.....	xi
Executive Summary .....	1
1 Introduction: Background and Purpose of the Study.....	4
1.1 Country Profile.....	4
1.2 Project Description .....	6
1.2.1 Goals and Expected Results .....	6
1.2.2 Net Instructional Time Realized by LTTP.....	7
1.3 Description of Intervention Components .....	3
1.3.1 Teacher and Pupil Materials .....	3
1.3.2 Training.....	3
1.3.3 Coaching.....	4
1.3.4 Assessment .....	5
2 Measuring Impact: The LTTP Research Design .....	7
2.1 History and Summary of the Design.....	7
2.2 Research Questions.....	9
2.3 Sampling.....	9
2.4 Data 12	
2.5 Assessment Instruments .....	13
2.5.1 EGRA Instrument.....	14
2.5.2 EGMA Instrument .....	14
2.5.3 SSME Instruments.....	14
3 Main Results.....	16
3.1 Notes on the Results for Cohort 1 and Cohort 2 Pupils .....	16
3.2 Instrument Reliability and Validity .....	17
3.3 EGRA Results .....	17
3.3.1 Reading Benchmarks .....	17
3.3.2 EGRA Instrument for Liberia .....	18

3.3.3	Analysis of EGRA Results .....	19
3.3.4	SUMMARY OF EGRA RESULTS .....	28
3.4	EGMA Results .....	29
3.4.1	EGMA Instruments for Liberia .....	29
3.4.2	Analysis of EGMA Results.....	30
3.5	SSME Results .....	36
3.5.1	SSME Instruments for Liberia.....	37
3.5.2	Findings from Pupil Interviews.....	37
3.5.3	Findings from Teacher Interviews .....	40
3.5.4	Findings from Head Teacher/Principal Interviews.....	49
3.5.5	Factors Associated with Pupil Performance .....	52
3.5.6	Summary of SSME Findings .....	62
4	Challenges.....	64
4.1	Impact of Ebola.....	64
4.2	Systemic Challenges.....	65
5	Recommendations.....	68
	References .....	72
	Annex A: Technical Details about Sample Design and Weighting Procedures .....	1
	Annex B: Technical Details About Subtask Equating.....	1
	Annex C: EGRA, EGMA, and SSME Instruments for the Liberia Endline Assessment.....	1
	Annex D: Gender Analysis of Scores, Cohort 2 .....	1
	Annex E: Other Statistical Analyses—Detail .....	1
	Annex F: Technical Details on Instrument Reliability and Validity Testing .....	1
	Annex G: Summary of Benchmarking Exercise in Liberia .....	1
	Annex H: Intraclass Correlation Coefficients and Standard Deviations .....	1

# LIST OF FIGURES

PAGE

Figure ES1. Average oral reading fluency, grade 1.....	1
Figure ES2. Average oral reading fluency, grade 2.....	2
Figure ES3. Average oral reading fluency, grade 3.....	2
Figure 1. Age distribution of pupils, endline sample.....	12
Figure 2. The change in distribution between Cohort 1 and Cohort 2 for correct letters per minute, grade 2.....	25
Figure 3. The change in distribution between Cohort 1 and Cohort 2 for correct invented words per minute, grade 2.....	27
Figure 4. The change in distribution between Cohort 1 and Cohort 2 for oral reading fluency words per minute, grade 2.....	28
Figure 5. The change in distribution between Cohort 1 and Cohort 2 for number identification per minute, grade 2.....	34
Figure 6. The change in distribution between Cohort 1 and Cohort 2 for quantitative comparison problems, grade 2.....	36
Figure 7. Reading practices at home of interviewed children (endline). ....	39
Figure 8. The classroom activities related to reading among interviewed children.....	40
Figure 9. Teacher characteristics and classroom resources.....	45
Figure 10. Years of experience as reported by the teachers.....	46
Figure 11. Teacher training and support in reading and mathematics instruction.....	47
Figure 12. Teachers' lesson plans on the day of the assessment.....	48
Figure 13. How teachers measure pupil progress.....	49
Figure 14. Characteristics of head teachers and schools.....	51
Figure 15. Head teachers' expectations regarding when all pupils should be able to read fluently.....	52
Figure D-1. Gap in oral reading fluency scores, grade 2 in Cohort 2, by gender (correct words per minute).....	D-4

Figure D-2. Gap in scores for number identification problems per minute, grade 2 in Cohort 2, by gender.....	D-5
Figure E-1. Change in distribution for oral reading fluency.....	E-2
Figure E-2. Change in distribution for correct letters per minute. ....	E-3
Figure E-3. Change in distribution for correct words per minute.....	E-5
Figure E-4. Change in distribution for correct invented words per minute.....	E-6
Figure E-5. Change in distribution for listening comprehension score.....	E-7
Figure E-6. Change in distribution for correct addition problems per minute.....	E-8
Figure E-7. Change in distribution for correct subtraction problems per minute.....	E-9
Figure E-8. Change in distribution for correct number identification problems per minute.....	E-10
Figure E-9. Change in distribution for missing number score. ....	E-11
Figure E-10. Change in distribution for quantitative comparison score. ....	E-12
Figure E-11. Change in distribution for word problem score.....	E-13

# LIST OF TABLES

	PAGE
Table 1. Summary of LTTP Interventions, by Subject, Cohort, and Academic Year .....	1
Table 2. Implementation of Reading and Mathematics Programs: Cohorts, Sample Sizes, and Schedule of Systematic Assessments, by School Year .....	8
Table 3. Final Sample Frequencies.....	11
Table 4. Subtasks of the (English) EGRA Instrument in Liberia .....	18
Table 5. Summary of Oral Reading Fluency Comparing Cohort 2 and Cohort 1, by Grade (Correct Words per Minute) .....	19
Table 6. Summary of Correct Letters Per Minute Comparing Cohort 2 and Cohort 1, by Grade .....	20
Table 7. Summary of Correct Familiar Words Per Minute Comparing Cohort 2 and Cohort 1, by Grade .....	21
Table 8. Summary of Correct Invented Words Per Minute Comparing Cohort 2 and Cohort 1, by Grade .....	22
Table 9. Summary of Listening Comprehension Percent Correct Comparing Cohort 2 and Cohort 1, by Grade .....	23
Table 10. Summary of Reading Comprehension Percentage of Correct Words Comparing Cohort 2 and Cohort 1, by Grade.....	23
Table 11. Summary of Initial Sounds Percent Correct Comparing Cohort 2 and Cohort 1, by Grade .....	24
Table 12. Subtasks of the EGMA instrument in Liberia .....	30
Table 13. Summary of Correct Addition Problems per Minute Comparing Cohort 2 and Cohort 1, by Grade .....	31
Table 14. Summary of Correct Subtraction Problems per Minute Comparing Cohort 2 and Cohort 1, by Grade .....	31
Table 15. Summary of Correct Number Identification Problems per Minute Comparing Cohort 2 and Cohort 1, by Grade .....	32
Table 16. Summary of Quantity Discrimination Problems Percent Correct Comparing Cohort 2 and Cohort 1, by Grade .....	33

Table 17. Summary of Word Problems Percent Correct Comparing Cohort 2 and Cohort 1, by Grade .....	33
Table 18. Summary of SSME Instruments for Liberia LTTP II .....	37
Table 19. Results from Pupil Interviews: Midterm and Endline.....	37
Table 20. Results from Teacher Interviews: Midterm and Endline .....	40
Table 21. Results from Head Teacher Interviews: Midterm and Endline.....	49
Table 22. School Characteristics and Reading Performance (Endline).....	53
Table 23. School Resources and Reading Performance (Endline) .....	57
Table 24. Pupil-Reported Teacher Practices and Reading Performance (Endline).....	59
Table 25. Home Reading Practices and Reading Performance (Endline).....	60
Table 26. Cohort 2 Schools' Progress on Lesson Plans at End of 2014/2015 Academic Year .....	65
Table B-1. Mean scores for reading fluency and comprehension, baseline and midterm.....	B-1
Table D-1. Change in Boys' and Girls' Reading Scores, Cohort 2, by Subtask and Grade.....	D-2
Table F-1. Cronbach's Alpha Coefficients for Reading Subtasks .....	F-2
Table F-2. Cronbach's Alpha Coefficients for Mathematics Subtasks .....	F-2
Table H-1. ICC and Standard Deviation Values for Oral Reading Fluency, by Grade.....	H-1

# ABBREVIATIONS

clpm	correct letters per minute
COEFF	model coefficient
cwpm	correct words per minute
DID	difference-in-difference
EdData II	Education Data for Decision Making
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
G1, G2, G3	grades 1, 2, 3
ICC	intraclass correlation coefficient
INSET	in-service education and training
LTTA	Liberia Teacher Training Program
LTTA II	Liberia Teacher Training Program Phase II
MOE	Ministry of Education
ORF	oral reading fluency
OYSS	Ordinary Yet Significant Series
PD	professional development
PPS	probability proportional to size
PTA	Parent-Teacher Association
RTTI	Rural Teacher Training Institute
SD	standard deviation
SE	standard error
SSME	Snapshot of School Management Effectiveness
USAID	United States Agency for International Development

This page intentionally left blank.

# EXECUTIVE SUMMARY

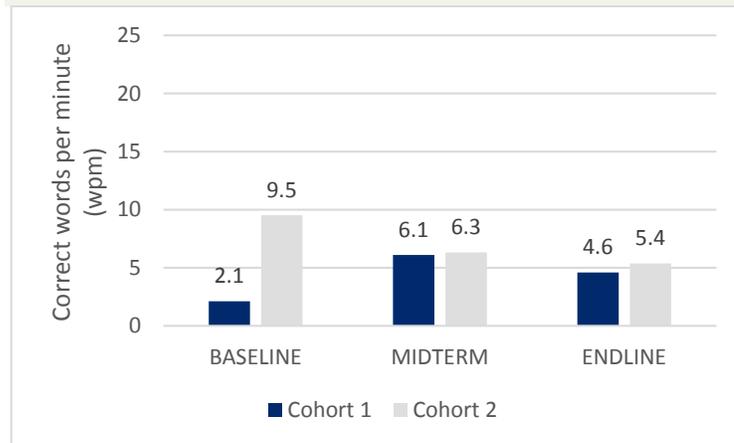
This report addresses the main research question: In Cohort 1 and Cohort 2 schools, were the pupils who participated in the United States Agency for International Development's Liberia Teacher Training Program (LTTP) Phase II (LTTP II) reading and mathematics interventions now achieving better results?

Additionally, this report also addresses the following research questions:

- How well were the reading and math programs implemented in Cohort 1 and Cohort 2 schools?
- What pupil, family, and school factors are associated with the variations in performance of pupils at different schools?

Because the Cohort 2 schools became LTTP intervention schools and the Cohort 1 schools stopped receiving direct support between the midterm and endline evaluations, it was decided to focus the main reporting on changes that had occurred between these two time points in addition to reporting the baseline results.

**Figure ES1. Average oral reading fluency, grade 1.**



The midterm assessment was conducted in May and June of 2013, and the endline assessment was administered in June 2015. However, due to the Ebola outbreak, the government ordered the schools in Liberia to be closed from

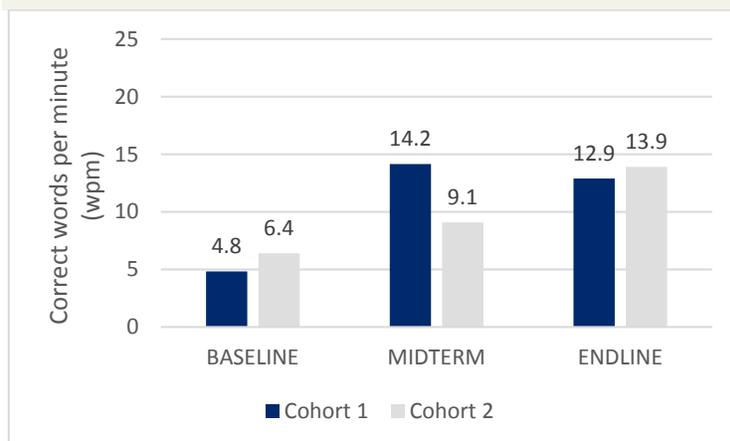
September 2014 until February 2015. It is a simple conclusion that these school closures had a devastating impact on early grade education, including the LTTP intervention. It is most evident in the results for grade 1 (**Figure ES1**), where the grade 1 pupils assessed in the endline had received a very limited amount of school-based instructional time compared with the

previous midterm assessment. As a result, average oral reading scores decreased in all schools. The Cohort 1 schools dropped from 6.1 to 4.6 correct words per minute (cwpm), and the Cohort 2 schools that benefitted from the LTTP intervention also fell from 6.3 to 5.4 cwpm.

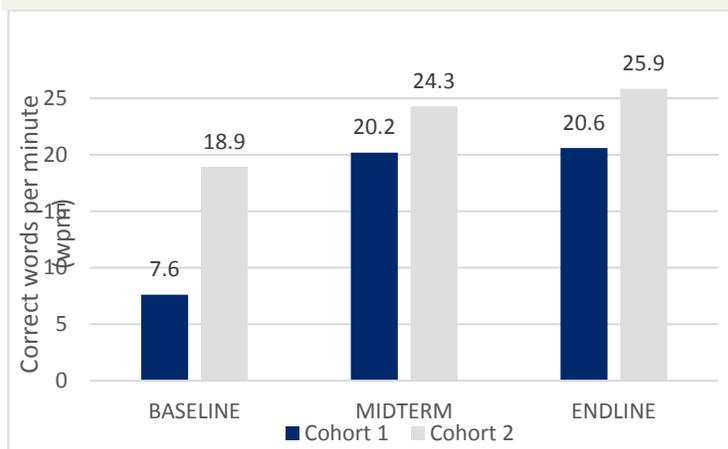
The results for grade 2 are slightly different.

Although the average oral reading fluency for Cohort 1 fell from 14.2 to 12.9 cwpm, the average correct words per minute for the Cohort 2 schools receiving the LTTP intervention increased from 9.1 to 13.9 cwpm. This increase for Cohort 2 is most likely explained by the fact that although these pupils received little classroom instruction time during grade 2, they received a full year of instruction and benefited from the

**Figure ES2. Average oral reading fluency, grade 2.**



**Figure ES3. Average oral reading fluency, grade 3.**



LTTP intervention in grade 1 during the previous year.

The results for grade 3 are a little different. The results for Cohort 1 were almost unchanged, seeing a slight increase from 20.2 to 20.6 cwpm. Cohort 2 fared a little better, seeing an increase in average oral reading fluency from 24.3 to 25.9 cwpm. The causes of the

grade 3 results are a little less intuitive, although, similar to the grade 2 pupils, the Cohort 2 pupils did benefit from a full year of instruction in grade 2 under the LTTP intervention. In-depth details regarding the early grade reading and mathematics assessments are presented in Section 3 of this report.

Having the support from parents or guardians is a critical factor in the success of a child's education. It was reported that only 50% of children are read aloud to at home and 50% of children said that they read aloud to someone at home.

Schools participating in the LTTP intervention seemed relatively well resourced; 94% of teachers in Cohort 1 schools and 93% of teachers in Cohort 2 schools had teachers' guides during the endline assessment. In addition, more than 80% of teachers in Cohort 2 schools had completed reading- or math-specific training and support during the past year. Teacher support is also important. During the assessments, it was reported that pupils in schools where classroom observations occurred daily by school administrators tended to read 6.4 cwpm faster than their counterparts in schools where observations occurred less often (every two weeks).

Although the Ebola outbreak created many challenges for the LTTP intervention, it was noted that pupils who practice letter sounds with their teachers in the classrooms tend to read 5.5 cwpm faster than their peers who do not practice letter sounds in their respective classrooms.

While the effectiveness of the LTTP2 intervention is difficult to determine, it should be mentioned that the pupil scores are low and remain so; for pupils to be able to read fluently with comprehension, the word per minute reading fluency needs to be much higher, particularly in grades 1 and 2 where generally if some pupils can identify a few words, they are usually not able to connect them in the passage to add comprehension.

# 1 INTRODUCTION: BACKGROUND AND PURPOSE OF THE STUDY

## 1.1 COUNTRY PROFILE

Liberia is Africa's oldest republic and one of the least developed and poorest countries on the continent. With some of the lowest Human Development Index indicators in the world, approximately 64% of Liberians live below the poverty line, and 48% live in extreme poverty (International Monetary Fund, 2008; United Nations Development Programme, 2013). Life expectancy is low, and nearly half of the population does not have sufficient food for a healthy life. Development in Liberia has been undermined by a history of unequal distribution of resources and opportunity among Liberia's varied communities and the dire civil war that destroyed much of the country's physical, economic, and social infrastructure between 1989 and 2003. Recovery since 2003 has been uneven, and the country exhibits many of the features of post-conflict fragility.

In addition, the outbreak of Ebola in West Africa and the devastating effect it has had in Liberia is a cause of great concern.

Located on the West Coast of Africa, with a land mass of 38,000 square miles, Liberia is bounded by the Republics of Guinea to the northeast, Sierra Leone to the northwest, and Cote d'Ivoire to the southeast. The present population of Liberia is more than 3.5 million, composed of 19 ethnic communities, including "Americo-Liberian" black freed men and women, former slaves from the Americas who settled in Liberia beginning in 1821. This former settler population constitutes approximately 5% of the present population. Since the landing of the first group of settlers in 1822, the political and economic hegemony of the Americo-Liberian community and the resistance of the indigenous population have been the defining characteristics of Liberia's history.

Tensions surrounding this hegemony came to a climax when Master Sergeant Samuel K. Doe overthrew the Tolbert Government during a military coup on April 12, 1980, and became the first indigenous Liberian President. Since then, the country has moved through cycles of violent conflict, elections, and relative stability; and back to violent conflict. These conflicts burgeoned in 1989, leading to Liberia's first civil war (1989–1996), followed by the election of Charles Taylor as President in 1997. The

second civil war (1999–2000) escalated into the third most destructive war (2000–2003) and pitted anti-Taylor forces against the regime, with extensive destabilizing impact on neighboring countries. One-third of the population of Liberia was displaced, and another third became refugees in the neighboring countries of Guinea, Sierra Leone, Ivory Coast, and Ghana.

Clearly, the past two decades of the twentieth century were the most turbulent period in Liberia's history. The Accra Comprehensive Peace Agreement was signed on August 30, 2003, bringing an end to the wars, the stationing of 14,000 United Nations Peacekeeping Mission troops (United Nations Mission in Liberia), and the installation of the National Transitional Government of Liberia. In 2005, national elections were held, and the winner and current President, Madame Ellen Johnson Sirleaf, was installed in January 2006 as the first woman President in Africa.

With the end of the civil war in 2003 and the elections of 2005, Liberia started on the long path to reconstruction. Infrastructure had been destroyed; the population was dispersed; the capital city, Monrovia, was destroyed; and the systems of government, including education, had collapsed. For example, by the end of the war in 2003, the three Rural Teacher Training Institutes (RTTIs) had not trained any teachers for nearly 20 years, and the institutes were in ruins, as were most of the country's schools.<sup>1</sup>

Since 2005, a number of education reforms aimed at restoring the educational infrastructure and reconstructing schools to accommodate the increasing population of school-aged children have been championed by the government. The overarching goal of the education emergency response in 2005 was to enroll as many children, as quickly as possible, into school to handle quality-of-education issues in the future. As a result, the number of pupils enrolled in schools increased significantly; however, the issues of educational quality and efficiency at all levels of the system remain a major challenge.

In June 2008, the World Bank sponsored the first-ever reading assessment in Liberia by using what was then the newly developed Early Grade Reading Assessment (EGRA) instrument. EGRA had been adapted for the country earlier that month with input from Liberian reading experts and the Ministry of Education (MOE). The analysis report that was based on the data collected clarified that rather than being applied in order to evaluate an intervention, this EGRA “was used to underwrite, with local data, suggestions as to how to improve reading that could then be tested with an intervention project which has an experimental nature” (Crouch & Korda, 2008). Several sequential interventions from the United States for International Development

---

<sup>1</sup> Under the Phase I of the Liberia Teacher Training Program (LTTP), two RTTIs at Zorzor and Kakata were reopened in 2006, followed by the RTTI at Webbo in 2010.

(USAID) would soon follow—including the Liberia Teacher Training Program Phase II (LTTP II)—as described in Subsection 1.2.1 of this report.

The key challenges articulated in the Education Sector Plan for Liberia (March 2010) are the following:

- Lack of coherent policies and appropriate education laws tailored to current and future directions of education
- Limited capacity at all levels of the system
- Weak structure and systems of educational governance, management, and accountability
- Nonexistent or inaccurate education data for informed decision
- Inadequate infrastructure for schools or teacher professional development (PD)
- Excessive number of untrained and unqualified teachers (more than 60% of the teaching workforce)
- High turnover in leadership
- Weak or inadequate linkages among the various levels of the system: national, county, district, school, and classroom.

## 1.2 PROJECT DESCRIPTION

### 1.2.1 GOALS AND EXPECTED RESULTS

The LTTP II is a partnership between FHI 360 and RTI International to provide support to the central MOE. The overarching goal of LTTP II is to enhance pupils' learning in general, and reading proficiency in particular; establish a functional teacher PD system; and strengthen the MOE's capacity to manage such a system. The LTTP II was originally designed to work in nine counties: Grand Gedeh, Grand Kru, Lofa, Maryland, Montserrado, Nimba, River Cess, River Gee, and Sinoe. In 2011 and 2012, because of changes in USAID policies, the number of counties was reduced to five (i.e., Bong, Lofa, Margibi, Montserrado, and Nimba), which USAID identifies as a "development corridor," containing a majority of the Liberian population.

The LTTP II interventions target reforms in the following three areas:

- Result 1: MOE, County Education Office, District Education Office, and RTTI capacity strengthened to plan, manage, and monitor educational services.

- Result 2: Improved teacher policies and procedures for teacher recruitment, training, deployment, and career development.
- Result 3: Improved teacher training programs and reading and mathematics delivery systems.

The three result areas previously mentioned constitute an integrated design of mutually reinforcing and necessary components. Success in the activities in Result 1 and Result 2 (management and policy) are critical for the success of activities in Result 3. The focus of the analyses described in this report was to determine the effectiveness of Result 3. Strengthening of reading and mathematics learning in primary school, particularly in the early grades, is essential for all further pupil learning and for the development of critical-thinking abilities. LTTP II, therefore, has emphasized the development of strengthened early grade reading and mathematics programs in Liberian schools, particularly in grades 1 through 3.

For Result 3, the essential ingredients of the improved instructional programs for reading and mathematics included daily scripted lesson plans, which were in a teacher's guide; activity books for each pupil that were aligned to the teacher's guides; and other materials for teachers and pupils that were designed around those lesson plans. The essential ingredients of these improved programs also included teacher training in reading and mathematics instruction based on the scripted lesson plans; regular coaching and support to help teachers improve their practice; and assessments at regular intervals throughout the school year so that teachers could monitor pupil progress. Section 1.2.2 discusses how each activity was implemented over the last two years in Cohort 1 and Cohort 2 schools.

Two sector-wide and persistent challenges to LTTP II implementation bear mentioning. First, at the school level, large amounts of time are lost when school starts late, ends early, or is closed on the local market day or for other—often capricious—reasons. Time is also lost because of poor attendance by teachers and pupils. Second, at the national level, instability within the MOE persists, with numerous and ongoing changes in key technical and leadership positions during the life of the project.

### 1.2.2 NET INSTRUCTIONAL TIME REALIZED BY LTTP

To fully understand the results for pupils, it is important to carefully examine when teaching began in classrooms and when it ended. Even though, in theory, Cohort 1 and Cohort 2 schools were slated to receive two years of support, the actual time spent on teaching was significantly less because of many challenges (these challenges are discussed in Section 4 of this report). In summary, for reading, pupils in Cohort 1 ultimately received 13 months of direct support for reading, and only four

months of support for mathematics, whereas pupils in Cohort 2 received nearly five and a half months of support in reading and mathematics.

**Table 1** provides an overview of the support provided, with the total time spent on direct work with pupils.

**Table 1. Summary of LTTP Interventions, by Subject, Cohort, and Academic Year**

	COHORT 1		COHORT 2	
	2011/2012	2012/2013	2013/2014	2014/2015
<b>ACTIVITIES</b>				
<b>Reading</b>				
Teacher training	5 days	5 days	5 days	3 days
Teacher's guides	Used the EGRA Plus pilot material: total of 110 lesson plans	New grade 1 (G1) curriculum used in G1 through grade 3 (G3)	G1 curriculum in G1 through G3	Grade-appropriate materials given to each grade
Pupil books	Used the EGRA Plus pilot material: compilation of decodable short stories	New G1 curriculum used in G1 through G3	G1 pupil book with decodable stories and practice sheets	Grade-appropriate pupil books
Supplementary readers	Ordinary Yet Significant Series (OYSS) books		OYSS and WeCare books provided to Cohort 1 and schools	Not applicable
Instructional aid	Pocket chart and letter cards	Pocket chart and letter cards	Not applicable	Not applicable
Coaching support	Each school visited 6–8 times	Each school visited 6–8 times	Each school visited 10–12 times	Each school visited 10–12 times
Teacher's guides	Used the EGRA Plus pilot material: total of 110 lesson plans	New G1 curriculum used in G1 through G3	G1 curriculum in G1 through G3	Grade-appropriate materials given to each grade
<b>Total support time: Reading</b>	<b>5 months</b>	<b>8 months</b>	<b>4 months</b>	<b>1.5 months</b>
<b>Mathematics</b>				
Teacher training	No intervention during Year 1	5 days	5 days	3 days
Teacher's guides	No intervention during Year 1	G1 guides given to grade 2 (G2) and G3 teachers	Grade-appropriate material used for G1; G2 materials provided for G2 and G3	Grade-appropriate material provided to all teachers
Pupil books	No intervention during Year 1	G1 books given to G2 and G3 pupils	Grade-appropriate material used for G1; G2 materials provided for G2 and G3	Grade-appropriate material provided to all pupils
Supplementary materials	No intervention during Year 1	No	None	None
Instructional aid	No intervention during Year 1	No	None	None
<b>Total support time: Mathematics</b>	<b>0 months</b>	<b>4 months</b>	<b>4 months</b>	<b>1.5 months</b>

This page intentionally left blank

## Cohort 1

During Year 1 of implementation (involving only Cohort 1 schools), the curriculum that was piloted as part of USAID’s EGRA Plus: Liberia program (Davidson, Korda, & Collins, 2011; Piper & Korda, 2011; and RTI International, 2011a) was used in all three grades. Because the overall quality of instruction and the resulting pupil performance were poor, LTTP II provided grade 1 materials to all three grades. This approach allowed all of the pupils to “catch up” and build solid foundations of basic skills. During the 2013/2014 school year, grades 2 and 3 used their grade-appropriate materials.

However, the time needed to print and distribute the teacher and pupil materials was longer than anticipated, and the reading materials were not delivered to schools at the August 2011 start of the 2011/2012 school year as planned. Likewise, training of teachers was provided in November and December 2011. Coaches were deployed to their assigned districts in December 2011 and February 2012. Teachers and pupils received the materials in February 2012.

During Year 2 of support, the reading materials again reached schools later than anticipated because heavy rains delayed distribution. Mathematics materials were delivered to schools in November and December 2012, and teachers began using them in January 2013. These delays meant that the midterm assessment was conducted after a shorter implementation period than was originally envisaged. Reading lessons were being implemented for 11 or 12 months, instead of two school years; and mathematics lessons for only five months instead of a full school year.

The direct support to Cohort 1 schools ended in May 2013 as planned by the program design. During the last two years of LTTP II implementation (2014 and 2015), Cohort 1 schools did not receive significant support except that LTTP II continued to distribute reading and mathematics books. Additionally, the program organized and facilitated two meetings for Cohort 1 Reading Support Teachers, principals, and District Education Officers from Bong, Lofa, Margibi, Montserrado, and Nimba counties. These meetings focused on reviewing the roles and responsibilities of Reading Support Teachers and principals to ensure that regular reading and mathematics instruction in Cohort 1 schools continued. The meetings also provided an opportunity for the officials to discuss the challenges and opportunities that they face and to review the effective use of reading and mathematics teacher’s guides and pupil activity books.

## Cohort 2

Cohort 2 schools did not receive LTTP II support during the first two years of program implementation. Support to these schools was slated to begin in September 2013 as

per the LTTP II design. However, because of internal restructuring of the intervention, direct support to the schools began in early March 2014 instead, thus allowing for only four months of support. Cohort 2 schools received both reading and mathematics materials and teacher training, school-based support, and other planned support activities. The program faced challenges similar to those of Cohort 1, especially the delayed books distribution for both years (i.e., 2013/2014 and 2014/2015 academic years). The biggest challenge that the program faced regarding the implementation in Cohort 2 schools was the school closings because of the Ebola crisis. Schools were closed between September 2014 and February 2015. Even after the official reopening date, with the gradual actual opening of schools that required LTTP II to wait until schools were safe to open, it took several months to distribute books to schools and to train teachers. Yet, the final assessment had to be completed in May 2015 due to the program closeout.

## 1.3 DESCRIPTION OF INTERVENTION COMPONENTS

### 1.3.1 TEACHER AND PUPIL MATERIALS

Each teacher received two volumes that contained scripted lesson plans for a given grade (materials for grades 1, 2, and 3 were developed). Volume 1 focuses on the first semester and volume 2 on the second semester. In addition to the lesson plans, teachers received letter cards and pocket charts. Each pupil was given a mathematics and a reading activity book for pupils. The content of the books were fully linked to the teacher manuals, allowing pupils to easily follow along during lessons and providing practice exercises to reinforce each day's lesson. Additional reading materials were also provided. Each school received approximately 50 titles per grade to use in their reading rooms. Even with the materials provided by the project, schools and their communities remained, for the most part, print-poor environments.

#### ACCESS TO MATERIALS

“In the Liberian school system, which has a dearth of materials and very limited in-service training opportunities, it was not difficult to gauge the enthusiasm of teachers for the math and reading programs. The early reading and math program provided them with the first training and materials they had received in years.”

—*External program evaluation report*

The cores of the reading and mathematics programs were the sets of scripted lessons developed by LTTP II for reading and math for grades 1, 2, and 3. The mathematics program contains 25 weeks of daily scripted lessons (125 lessons) for each grade and is sequenced in a manner to follow the Liberian curriculum and standards. The reading program consists of 30 weeks of scripted daily lessons (150 lessons) for each grade and aligned with the Liberian curricular expectations. The scripted lessons provide specific instructions for the teachers regarding lesson content

and conduct. Pupil materials were developed so that there would be direct alignment between the content in the teacher's scripted lessons and in the corresponding page of the pupils' books. Additional sets of readers also were provided to each school in Cohort 1 and 2 schools.<sup>2</sup>

### 1.3.2 TRAINING

To take full advantage of these materials, teachers required adequate training. LTTP II provided a two-week workshop (one week for reading, one for mathematics) for teachers at the beginning of the academic year. Follow-up refresher training was held

<sup>2</sup> LTTP would like to thank Jennifer Cooper-Trent for allowing us to use her *Fantastic Phonics* books for Cohort 1. More information about the books is available on the Teach the World to Read Web site at <http://www.teachtheworldtoread.com> (Cooper-Trent, 2015).

at the beginning of the second semester of the same year. These training sessions focused on the use of the instructional materials, especially the scripted lessons.

There was heavy emphasis on using the scripted lessons for practice; teachers were able to review a month's worth of lesson plans.

Two weeks of initial training, however, was a bare minimum. Regular follow up and coaching were necessary to ensure that the teachers learned and internalized the methodologies embodied in the scripted lessons. Additionally, overall, the Liberian teachers themselves had low levels of skill development, in terms of content and pedagogy; therefore, teacher upgrading was needed. High teacher turnover was also an ongoing problem because significant numbers of teachers trained by the LTTP II were later transferred out of project schools.

## ACCLAIM FOR THE INTERVENTION

“The assessment teams’ review of the reading and math program revealed a number of strengths, as well as several weaknesses in the program. On the positive side, both the early grade reading and math programs were widely acclaimed by teachers, principals, and administrators in each of the five counties we visited.”

—*External program evaluation report*

In addition to the training that focused on the specific instructional approach of the LTTP II interventions, the project also requested that teachers who were not fully certified attend the year-long in-service program implemented by the MOE’s RTTIs. This program involves one month of residential training, plus monthly visits. LTTP II supported this initiative by providing training to the in-service trainers who worked with these teachers.

### 1.3.3 COACHING

Teachers were introduced to and got to practice by using the scripted lessons during the trainings offered at the start of the school year and at the beginning of the second semester. However, ongoing coaching played an important role in reinforcing what the teachers learned during those training sessions, helping them implement their lessons, and improving their

instruction on a daily basis. When the EGRA Plus model was taken to scale under LTTP II, coaching was one area in which the approach was compromised, in the interest of lowering costs (i.e., by changing the teacher-to-coach ratio). Under LTTP II, one coach was initially assigned to 12 schools, as compared to one coach for four to

***“Most teachers and principals cited good support and helpful, constructive criticism from the coaches as an important part of the program’s success.”***

—*External program evaluation report*

eight schools under EGRA Plus. This meant many fewer visits per month by a coach to the teachers that he or she supported. LTTP II coaches visited schools once per month, following a systematized coaching model. The LTTP II coaches provided additional training, observed teachers, and assisted them with any challenges that they might have had. Coaches worked less frequently on mathematics-related instruction than on reading because of the limited time available. For Cohort 2, the number of schools was more than halved, allowing for a much smaller coach-to-school ratio, at one coach per six schools. This smaller ratio created opportunities for coaches to provide better support to teachers.

During the last two years of LTTP II program implementation, a different approach to providing coaching support to schools was piloted in 21 schools. Instead of providing one visit per month to schools, the new model required each coach to spend five days with the same school. The program staff developed a day-by-day agenda for coaches that would be implemented with teachers. The agenda included observations, modeling, discussions with teachers in grades 1–3, mini-training sessions after school, assessments, and working with the principal to improve school leadership and to more fully engage parents in school matters. The results of this pilot were remarkable in terms of pupil learning performance, resulting in an average percentage increase of almost 90% on oral reading fluency (ORF) and an approximately 130% increase in comprehension. Some of these lessons learned were incorporated into the

support provided to almost 60% of the Cohort 2 schools.

## SUSTAINABILITY OF COACHING

“The only scheduled and reliable source of instructional supervision in MOE primary schools is being provided by the early grade reading and math coaches funded by the LTTP II. Although limited to supervision of those teachers who have received training in the EGRA and EGMA [Early Grade Mathematics Assessment] methods, field visits by the LTTP II coaches represents a working model that the MOE could emulate.”

—*External program evaluation report*

### 1.3.4 ASSESSMENT

The last component of the interventions concerned regular assessment by teachers of their pupils’ progress. Teachers were trained on how to properly use the EGRA-like

instruments to assess pupil performance on letter knowledge, reading, and comprehension at three points during the year. Teachers were supposed to track pupil performance from one assessment to another, producing individual pupil report cards. Additionally, teachers worked with their principals to develop school report cards that could be discussed with their school’s Parent-Teacher Association (PTA).

Classroom assessment proved to be the most challenging aspect of LTTP II program design and eventually was abandoned in the latter years of the program. Teachers did

not want to conduct these assessments because it was time consuming, and they deemed it to be additional work and therefore demanded to be paid for it. More systematic policy work is required in order to introduce classroom assessment as part of teacher job descriptions, as well as to ensure that teachers have ample time to conduct the assessments. In Cohort 2, no tracking of pupil performance by teachers was completed.

## 2 MEASURING IMPACT: THE LTTP RESEARCH DESIGN

### 2.1 HISTORY AND SUMMARY OF THE DESIGN

From 2008 to 2010, RTI—with the Liberian Education Trust—implemented the EGRA Plus, a Liberia pilot program as a task order within USAID’s Education Data for Decision Making (EdData II) project.<sup>3</sup> Based on the success of the EGRA Plus pilot, USAID decided to include an expanded implementation of the reading and mathematics interventions in its Liberia Teacher Training Program (LTTP).

Of interest to USAID, the Government of Liberia, and the broader community of education stakeholders in Liberia and around the world is whether the same type of impacts that were achieved during the piloting of EGRA Plus could be realized if the program were implemented on a much broader scale. Doing so would imply significantly greater implementation challenges—in terms of simple logistics, but also in terms of maintaining quality across a larger set of actors and responding to the circumstances of a greater cross-section of school–community contexts.

Pupils in EGRA Plus full treatment schools realized an increase in their ORF in English scores by almost three times as much as the level of improvement experienced in control schools. An evaluation of EGRA Plus indicated that its success relied on developing and making available highly scripted instructional materials for teachers and accompanying books for pupils, training teachers on how to use of those materials, providing coaching as a critical component of ongoing support and supervision to teachers, and making use of regular assessments of pupils’ progress.

The LTTP II intervention drew on the EGRA Plus model to introduce similarly structured reading and math programs in grades 1, 2, and 3 to approximately 1,020 schools in four counties (i.e., Bong, Lofa, Montserrado, and Nimba) in a phased approach. Cohort 1, the first to receive support, had 792 schools. During the middle of the 2011/2012 school year, the reading program was introduced in all three grades in these schools. During the middle of the 2012/2013 school year, the mathematics program was introduced in all three grades. Cohort 2, consisting of approximately 330 schools, began participating in the program’s reading and mathematics interventions during the 2013/2014 school year and continued during the 2014/2015 school year.

---

<sup>3</sup> For background about EGRA Plus, see Davidson, Korda, & Collins, 2011; Piper & Korda, 2011; and RTI International, 2011a.

Some changes, although not significant, were made to the intervention approach for supporting the Cohort 2 schools.<sup>4</sup>

To fully investigate the impacts of the reading and mathematics programs, LTTP’s interventions were structured to allow rigorous evaluation. An overview of the impact evaluation approach is presented in **Table 2**.

**Table 2. Implementation of Reading and Mathematics Programs: Cohorts, Sample Sizes, and Schedule of Systematic Assessments, by School Year**

COHORT	2011/2012 SCHOOL YEAR	2012/2013 SCHOOL YEAR	2013/2014 SCHOOL YEAR	2014/2015 SCHOOL YEAR
Cohort 1	792 schools (reading)		792 schools (mathematics)	
Cohort 2	~330 schools (reading and mathematics)			
External Cohort				
Assessment	Baseline May 2011	Midterm May 2013	Final May 2015	
Cohort 1	50	50	50	
Cohort 2	50	50	50	
External Cohort	50	50	50	

- Cohort 1: Schools from the four target counties included in Cohort 1 served as the treatment group for the midterm assessment. These schools stopped receiving LTTP II support after the midterm assessment, but they participated in the endline assessment, as a way to determine whether the gains that were achieved during the treatment were sustained.
- Cohort 2: Schools included in Cohort 2 in the same four counties began to receive treatment after the midterm assessment—thus, during the final two years of the program. Cohort 2 schools served as a control to which the Cohort 1 results were compared. The performance of Cohort 2 schools were to be compared to that of Cohort 1.
- External Cohort: A randomly selected sample of schools outside the four target counties served as another comparator, especially after Cohort 2 began receiving treatment alongside Cohort 1.

<sup>4</sup> The LTTP project methodology is explained in the document *Early Grade Reading and Mathematics Methodology: From Pilot to Scale* (RTI International, n.d.).

## 2.2 RESEARCH QUESTIONS

This report addresses the main research question: In Cohort 1 and Cohort 2 schools, are the pupils who participated in USAID’s LTTP II reading and mathematics interventions now achieving better results?

To answer that question, it was necessary to also pose several other questions, including the following:

- How did performance in the Cohort 1 and Cohort 2 schools compare—at baseline, at midterm, and at final?
- Were the Cohort 1, Cohort 2, and External Cohort schools (and pupils) different in any significant ways?

The assumption underlying this study was that any improvement achieved by pupils in Cohort 1 and Cohort 2 schools would be due to the implementation of the reading and mathematics programs in those schools. Therefore, the following was an additional question of concern: How well were the reading and mathematics programs implemented in Cohort 1 and Cohort 2 schools?

Lastly, when, as assuredly is the case, variations in the performance of pupils across schools manifest themselves, it is useful to investigate the following additional question: What pupil, family, and school factors are associated with the variations in performance of pupils at different schools?

To answer these fundamental evaluation questions, baseline, midterm, and endline assessments were structured to facilitate the necessary comparisons of schools in the three different cohorts. The annexes of this report discuss the sample design and weighting procedures (**Annex A**), subtask equating (**Annex B**), and the instruments used for the endline assessment (**Annex C**). The annexes of this report also discuss the gender analysis of scores for Cohort 2 schools (**Annex D**) and other statistical details (**Annex E**). Lastly, the annexes present technical details about instrument reliability and validity testing (**Annex F**), a summary of the benchmarking exercise in Liberia (**Annex G**), and a discussion of intraclass correlation coefficients and standard deviations (**Annex H**).

## 2.3 SAMPLING

Schools in the four LTTP II counties were randomly assigned to the Cohort 1 and Cohort 2 groupings. These schools were then grouped in clusters of 12 schools based on geographic proximity, which would allow the program to deliver the interventions more efficiently. As illustrated in **Table 1**, during the middle of the 2011/2012 school year, the Cohort 1 schools implemented the reading program, continuing it into the

2012/2013 school year, but experienced delays experienced in receiving materials at the start of that year. The Cohort 1 schools began the mathematics program in January 2013.

Cohort 2 schools began implementing the program during the next school year (2013/2014) and continued to receive treatment through the end of the 2014/2015 school year. LTTP implementation in the Cohort 2 schools suffered severely because of the Ebola crisis, which caused the schools to close—and subsequently stopped the program—for seven months. Except for a small number of schools associated with the RTTIs, schools outside the four target counties did not participate in the program during the lifetime of LTTP II.

For the midterm assessment of LTTP II's impact in Cohort 1 schools, Cohort 2 schools were considered to be a useful comparator because they were in the same counties (and sometimes in the same districts) as the Cohort 1 schools. The schools in the same counties shared similar characteristics and functioned within the same administrative and operational environments. Thus, they provided a useful control for LTTP II, with a comparison of Cohort 1 and Cohort 2 performance at baseline and midterm serving as the means to “isolate” the impact of the reading and mathematics interventions. For Cohort 2 schools, a useful comparator did not exist. Instead, here we report their performance in relation to Cohort 1 because Cohort 1 schools received similar support in the past and very little support during the final two years of the program.

To include data on all three categories of schools, random samples of schools were selected from the Cohort 1, Cohort 2, and External Cohort groups at baseline, at midterm, and again at endline. Based on power calculations conducted at baseline, a total of 150 schools needed to be sampled during each phase of the project, or 50 from each category, as shown in **Table 2**.

For the baseline, midterm, and endline assessments, participating districts within each county were randomly selected. In these instances, schools within a county were then randomly selected based on a probability proportional to each school's enrollment. That is, each school was weighted based on the proportion its pupils represented of the total enrollment in the county. The final sample frequencies are shown in **Table 3**. Note that for the endline assessment, more than 50 schools were selected for both cohorts because data collection occurred during the rainy season and attendance was poor. Therefore, a decision was made to visit more schools to keep the sample size reasonable.

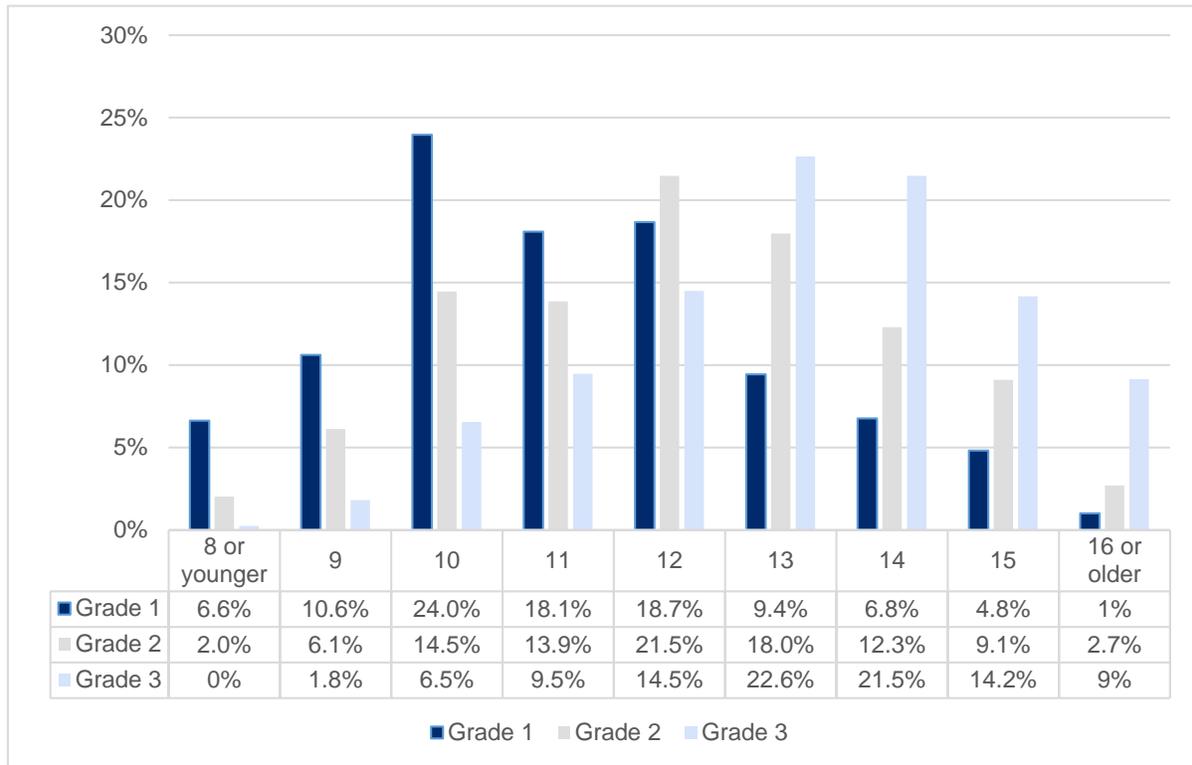
**Table 3. Final Sample Frequencies**

ASSESSMENT	COHORT	NUMBER OF SCHOOLS	NUMBER OF PUPILS			
			GRADE 1	GRADE 2	GRADE 3	OVERALL
Baseline	Cohort 1	63	491	478	466	1,498
	Cohort 2	26	169	153	135	483
	External Cohort	42	351	295	295	983
	OVERALL	131	1,011	926	896	2,964
Midterm	Cohort 1	62	476	457	459	1,454
	Cohort 2	36	281	271	257	845
	External Cohort	45	333	325	306	1,009
	OVERALL	143	1,090	1,053	1,022	3,308
Endline	Cohort 1	55	481	468	452	1,456
	Cohort 2	53	431	416	399	1,299
	External Cohort	50	418	374	373	1,215
	OVERALL	158	1,330	1,258	1,224	3,970

Pupils were selected at random, irrespective of gender. Based on this random selection of pupils, the findings show that 50% (grade 1), 52% (grade 2), and 49% (grade 3) were girls. The ages of the pupils selected for the endline sample are shown in **Figure 1**. Considering the appropriate ages for grade 1 are 6–7 years, grade 2 are 7–8 years, and grade 3 are 8-9 years, most pupils are over the appropriate age for the grade level they are attending. Policy and curriculum recommendations regarding this concern should be based on further research to discover the causes of so many pupils being over age for their respective grade.

More details regarding how the three groups of schools were sampled and weighted for the endline assessment are presented in **Annex A** of this report.

**Figure 1. Age distribution of pupils, endline sample.**



## 2.4 DATA

The EGRA and Early Grade Mathematics Assessment (EGMA) instruments developed for Liberia under EGRA Plus were used to assess pupil performance in reading and mathematics and were further adapted for use under LTTP II. Equivalent versions of the assessment were used at baseline and midterm, with pupils across all three grades being evaluated using the same or equivalent mathematics and reading tests. More information about the equating process is presented in **Annex B** of this report.

For the endline assessment, LTTP II used the same instruments that were used for the baseline assessment, thus eliminating the need for equating between baseline and endline assessment. Detailed descriptions of the EGRA, EGMA, and Snapshot of School Management Effectiveness (SSME) methodologies are available, and the instruments used for this evaluation are presented in **Annex C** of this report. The reading assessment evaluated pupils in seven skill areas: phonemic discrimination,<sup>5</sup>

<sup>5</sup> Several of the midterm and endline EGRA subtasks had slightly different names in the electronic Tangerine® version than had been used in the paper-based instruments at baseline. For simplicity, for our comparative analyses, we use the subtask names from the midterm and endline instruments.

letter name knowledge, familiar word reading, nonword reading or decoding, ORF, oral reading comprehension, and listening comprehension. The mathematics assessment tested pupils on number identification, quantity discrimination, missing number, addition, subtraction, and word problems.

The baseline assessment was conducted in May 2011, well before the reading and mathematics interventions were introduced into Cohort 1 schools. These data provided end-of-grade measures of pupil performance in the three grades. The baseline data presented in this report, however, differ from those contained in the LTTP II *Baseline Assessment Report* from October 2011 (RTI International, 2011b). After careful review of how schools were placed into the three groups (i.e., Cohort 1, Cohort 2, and External Cohort), LTTP II staff and RTI analysts determined that the cohort assignments of 46 schools were no longer accurate and therefore needed to be updated for the midterm analysis. For consistency, the schools' baseline designations as Cohort 1, Cohort 2, or External Cohort were also changed. The reassignment of schools to the appropriate group changed the composition of those groups at both baseline and midterm, thus changing the calculations of weighted mean scores at baseline.

The midterm assessment was conducted from late May to early June 2013, providing end-of-grade points of comparison. The endline assessment was conducted from May through June 2015.

In addition to the pupil reading and mathematics assessments, the endline data collection also surveyed pupils about their families, their home situations, their attendance at school, and the actions of their teachers. Teacher and head teacher questionnaires gathered additional data about each school and its instructional environment (Note: More information about the instruments are presented in **Annex C** of this report).

The resulting set of data contributes to our understanding of the schools in each of the comparison categories, for Cohort 1 and Cohort 2 schools. In that section and throughout this report, all summaries use weighted data and thus are representative of the entire population of schools from which the samples were drawn.

## 2.5 ASSESSMENT INSTRUMENTS

The assessment instruments used at endline were also used for the LTTP II baseline, thereby allowing for a perfect comparison of scores. The assessment instruments used during this study are presented in **Annex C** of this report.

### 2.5.1 EGRA INSTRUMENT

The EGRA is an orally administered instrument that measures the pre-reading and reading skills that serve as a foundation for later reading, which helps to ensure academic success. It takes approximately 15 minutes to administer EGRA to a pupil. EGRA is often combined with a questionnaire to measure a variety of pupil background variables to assist in explaining some of the reading outcome findings. EGRA provides results on the following subtasks: orientation to print, letter name knowledge, phonemic awareness (initial letter sounds), familiar world identification, simple unfamiliar (nonword) decoding, passage reading (ORF), reading comprehension, and listening comprehension.

### 2.5.2 EGMA INSTRUMENT

The interest in an adaptation of an appropriate mathematics assessment in the early grades became strong after the reading assessments revealed low pupil performance on foundational skills. In June 2009, USAID funded a development of EGMA, which was then piloted in Kenya in 2009. Continuing this trend, the World Bank funded an EGMA in Liberia in 2010 as part of the EGRA Plus project. The findings of this assessment indicated that if pupils learned to read, then they would performed better on mathematics—even without an intervention in this subject. Consequently, USAID required LTTP II to intervene and assess pupil reading performance on mathematics in grades 1, 2, and 3. The following EGMA components were used for LTTP II assessments: number identification, quantity discrimination, missing number, addition, subtraction, and word problems.

### 2.5.3 SSME INSTRUMENTS

The SSME consists of a range of instruments that yields a quick, but rigorous and multifaceted, picture of school management and pedagogical practice in a country or region. The SSME was designed to capture indicators of effective schools that past research has shown to affect pupil learning. The resulting data are designed to enable school, district, provincial, or national administrators and donors to learn what is currently occurring in their schools and classrooms and to assess how to make these schools more effective. Building off of the framework for the analysis of effective schools described by Heneveld & Craig (1996), the SSME collects a variety of information: pupil and household characteristics, basic school inputs (e.g., school infrastructure, pedagogical materials, teacher and head teacher characteristics), and classroom teaching and learning processes (e.g., pupil–teacher interaction, assessment techniques). In addition, the EGRA and EGMA components of the study that provide information about the achievement of learning outcomes in reading, writing, and

arithmetic can be associated with the SSME data to gauge whether school- or community-related variables may have had an effect on the assessment outcomes.

Each of the SSME's components is designed to elicit information from a different perspective. For the endline assessment in Liberia, the SSME components chosen were the Student Questionnaire, the Head Teacher Questionnaire, the Teacher Questionnaire, and the School Inventory (Note: The SSME instruments are presented in **Annex C** of this report). The design of the SSME aims to balance the need to include a broad mix of variables with the competing need to create a tool that is as undistruptive to the school day as possible. When combined, the components of the assessment produce a multifaceted and comprehensive picture of a school's learning environment. When the results from multiple schools in a region are compared, then it becomes possible to account for differences in school performance. The four SSME components were administered as follows:

- Student Questionnaire: Administered to each pupil randomly selected for assessment
- Head Teacher Questionnaire: Administered to the head teacher in each school visited
- Teacher Questionnaire: Administered to the teachers whose pupils were selected for assessment
- School Inventory: Administered via observation, school records, and checklists at each school visited.

# 3 MAIN RESULTS

This section of the report presents the findings from the LTTP endline EGRA, EGMA, and SSME administration in the form of comparisons by cohort and grade. Endline performance is also compared to midterm performance via difference-in-difference (DID) analysis and, to the extent feasible, to baseline performance. Section 3.1 elaborates on some differences between the cohorts that required additional steps to ensure valid analysis results.

## 3.1 NOTES ON THE RESULTS FOR COHORT 1 AND COHORT 2 PUPILS

As previously explained in Section 2.1 of this report,

- Pupils in Cohort 2 schools did not receive any specific support, other than regular instruction, between baseline and midterm, but benefited from full USAID intervention between the midterm and endline assessments.
- Pupils in Cohort 1 benefitted from the LTTP II intervention between the baseline and midterm assessments, but did not receive any additional USAID support between the midterm and endline assessments other than materials.

Thus, Cohort 1 in many ways can be viewed as a partial treatment group between midterm and baseline assessments because it is possible that these schools could prolong the benefit of the intervention implemented between baseline and midterm. Additionally, although there were two years between the midterm and endline measurements, the schools were closed for seven months (between September 2014 and February 2015) because of the Ebola crisis; therefore, no instruction occurred during that time.

Section 3.2 briefly discusses instrument reliability and validity (more details are presented in **Annex F** of this report). Sections 3.3 and 3.4 present data for all reading and mathematics subtasks completed by Cohort 1 and Cohort 2 pupils across all three time periods to answer the following analysis questions:

- How did pupils in Cohorts 1 and 2 schools perform during baseline, midterm, and endline assessments?
- How did pupils in Cohort 2 schools perform compared with Cohort 1 pupils between midterm and endline assessments?

These questions formed the core of a DID analysis that we carried out as the most consistent way to report on the effectiveness of the USAID intervention in Cohort 2 (**Table 1**).

*Note that the DID analysis approach is common for treatment-control studies, but these results are not true treatment-control comparisons and should not be interpreted as such.*

As an additional point of interest, comparisons by gender for Cohort 2 are only supplied in **Annex D**; additional technical details appear in the figures in **Annex E**.

## 3.2 INSTRUMENT RELIABILITY AND VALIDITY

Internal consistency is an appropriate and standard classical evaluation approach for cross-sectional data. Cronbach's alpha was 0.84 for the reading subtasks and 0.85 for the mathematics subtasks. Cronbach's alpha should be at least 0.70 for adequacy, and coefficients closer to 1 indicate a good assessment (Aron, Aron, & Coups, 2013). Technical details about instrument reliability and validity testing are presented in **Annex F** of this report.

## 3.3 EGRA RESULTS

### 3.3.1 READING BENCHMARKS

Because readers are examining the EGRA results presented in this section, it may be useful as a point of comparison to be aware of a set of reading-related benchmarks for grade 3 pupils in Liberia that were developed by the MOE.

During March 2014, funding from a task order under the USAID EdData II project and LTTP II made it possible to convene a two-day participatory exercise to determine these benchmarks. Two RTI technical advisors met in Monrovia with representatives of USAID, the MOE, and many other education stakeholders from the donor and nongovernmental organization community to discuss a variety of topics. The topics include the country's reading assessment results to that point, the value of such benchmarks in general, changes that the MOE participants had witnessed because of reading intervention programs, and specific national standards—generated by the attendees—for reading comprehension, ORF, and decoding of nonwords.

The process used and the targets reached by consensus at that meeting were recorded and disseminated in a document titled *Proposing Benchmarks for Early Grade Reading Skills in Liberia*, which has been reproduced for this report as **Annex G**.

### 3.3.2 EGRA INSTRUMENT FOR LIBERIA

The instruments for the May 2015 EGRA and EGMA were the same as those used for the LTTP II 2011 baseline assessment. In 2011, the instruments were adapted for the country context via workshops with the MOE, and input was incorporated from local and international reading and mathematics experts. At that time, the instruments were duly pilot tested and closely examined for item difficulty.

Because the instruments were significantly modified from baseline to midterm, **Annex B** contains reference information about the equating process at that stage.

**Table 4** summarizes the subtasks of the English EGRA designed for Liberia. **Annex C** presents a copy of one form of the student instrument, which combines the reading assessment, the mathematics assessment, and the Student Questionnaire (student context inventory).

**Table 4. Subtasks of the (English) EGRA Instrument in Liberia**

SUBTASK	SKILL	DESCRIPTION— THE CHILD WAS ASKED TO ...
1. Orientation to print	An understanding of directionality of reading print on a page	... indicate where one would begin reading printed text on a page and the direction one would read that text. <i>(Untimed subtask)</i>
2. Letter name knowledge	The ability to produce the name of a letter that is presented in written form	... produce the names of 100 letters presented in written form. Letters were presented in a grid of 10 rows and 10 columns. <i>(Timed subtask)</i>
3. Phonemic awareness (initial letter sounds)	The ability to identify sounds occurring at the beginning of spoken words	... listen to 10 sets of three words read aloud, one of which began with a sound different from the others; and identify which word had the different beginning sound. <i>(Untimed subtask)</i>
4. Familiar word identification	The ability to recognize or decode familiar words	... read aloud from a stimulus sheet as many as possible of the 50 familiar words (5 columns of 10 words each) presented on the sheet. <i>(Timed subtask)</i>
5. Simple invented (nonword) decoding	The ability to decode unfamiliar words	... sound out, or decode, unfamiliar words. To ensure that all words would be unfamiliar words, 50 words without meaning, but following English spelling and grammatical rules, were presented to the child to read. <i>(Timed subtask)</i>
6a. Passage reading (ORF)	The ability to quickly and accurately read connected text on a page	... quickly and accurately read a passage of narrative text of 60 words in length. <i>(Timed subtask)</i>
6b. Reading comprehension	The ability to orally respond to both literal and inferential questions about the ORF passage read	... orally respond to three questions asked about the passage read. <i>(Untimed subtask)</i>
7. Listening comprehension	The ability to comprehend an orally presented story and provide an oral response to question asked.	... orally respond to five questions asked about an orally presented story. <i>(Untimed subtask)</i>

### 3.3.3 ANALYSIS OF EGRA RESULTS

Note that in the Description column of **Table 4**, that four of the reading subtasks were timed (i.e., letter name knowledge, familiar word identification, simple unfamiliar [nonword] decoding, and passage reading [ORF]), thus providing measures regarding how well pupils performed in these skill areas and how automatic each skill had become for them. As previously mentioned, because the endline assessment instruments were the same as those used at baseline, no equating was required for the pupils' scores from the endline assessment. By contrast, the ORF scores for the midterm assessment had been equated to account for differences in the reading passages between the baseline and midterm (see **Annex B**). Baseline, midterm, and endline results for ORF for pupils in Cohort 1 and 2 schools are presented in **Table 5**. Note the dotted lines in **Table 5** that enclose the midterm and endline estimates and DID indicate the focus of this statistical analysis.

**Table 5. Summary of Oral Reading Fluency Comparing Cohort 2 and Cohort 1, by Grade (Correct Words per Minute)**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	2.1	6.1	4.6	0.6	0.1
	2	9.5	6.3	5.4		
2	1	4.8	14.2	12.9	6.1	0.3
	2	6.4	9.1	13.9		
3	1	7.6	20.2	20.6	1.2	0.0
	2	18.9	24.3	25.9		

An **effect size** can be interpreted as a gain in standard deviations. Effect sizes of greater than 0.25 standard deviations are interpreted as a “quantified positive” effect, even though they may not reach statistical significance in a given study (Institute of Education Sciences, 2014). For the purposes of this study, the effect size is calculated as *Cohen’s d*.

The results in **Table 5** for Cohort 1 show that for all grades, many of the gains in mean reading fluency between baseline and midterm—when the cohort received USAID intervention—were lost between midterm and endline when the cohort received no specific intervention; that is, fluency decreased or remained virtually unchanged. As a result, the DIDs are inconclusive as impact evaluation interpretations. If anything, these results highlight the in-country educational challenges and fidelity of treatment for

Cohort 2 because of the Ebola crisis.

For example, in grade 1 and Cohort 1, baseline to midterm results increased from 2.1 to 6.1 correct words per minute (cwpm), but decreased to 4.6 at the endline. Cohort 2, which received the USAID intervention between midterm and endline, was mixed in its results. Grade 1 decreased (from 6.3 to 5.4 cwpm), grade 2 increased (from 9.1 to

13.9 cwpm), and grade 3 had a modest increase (from 24.3 to 25.9 cwpm). All of these mean scores for grade 1 indicate that the pupils, on average, cannot read and are at best identifying familiar words. The greatest DID is seen in grade 2, where Cohort 2 had a 6.1 cwpm gain over Cohort 1 between endline and midterm. However, this figure is inflated because the Cohort 1 mean scores decreased between midterm and endline. The effect size for this DID is 0.3 standard deviations (SD), which would otherwise be interpreted as a positive effect, and the DID is not statistically significant. (Note: A brief discussion of intraclass correlation coefficients and standard deviation is presented in **Annex H** of this report.) The other effect sizes reported were small. An explanation for why grade 2 performed reasonably well compared with grade 1 in Cohort 2 is that grade 2 pupils received at least one year of instruction (as grade 1 pupils) between the midterm and endline measurements, whereas grade 1 pupils received only approximately six weeks of support. Overall, the grade 2 mean scores for reading fluency are still very low, and the pupils are not reading at a rate of correct words per minute in which they are comprehending what they are reading. The scores for grade 3 pupils might have leveled off more because they did not have enough instruction time at a higher level to further increase their scores. The mean scores for grade 3 indicate that some pupils are starting to read with some fluency and comprehension.

The results of the mean letter name knowledge subtask as reported in correct letters per minute (clpm) is presented in **Table 6**.

**Table 6. Summary of Correct Letters Per Minute Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	42.6	56.2	52.0		0.5
	2	57.9	53.6	61.4	12.0*	
2	1	61.0	75.2	71.5		1.0
	2	64.8	54.6	80.5	29.7**	
3	1	72.0	85.4	85.9		0.5
	2	65.2	79.0	92.1	12.6*	

\* $p < 0.05$ , \*\* $p < 0.001$

Although Cohort 1 reported gains between baseline and midterm assessments, the increase in mean letter name knowledge subtask either was flat or slightly decreased between midterm and endline. For example, in grade 3 and Cohort 1, the baseline mean score was 72.0 clpm, the midterm mean score was 85.4 clpm, and endline mean score was 85.9 clpm. In contrast, Cohort 2 reported improvements across all

grades from midterm to endline, which resulted in statistically significant DID comparisons when compared to Cohort 1. However, similar to reading fluency, these gains must be in context. Cohort 1 is not a true comparison group and in some places suffered decreases between midterm and endline. This is demonstrated in grade 2, where **Table 6** shows the greatest DID of 29.7 clpm: Cohort 1 decreased from 75.2 to 71.5 clpm, and Cohort 2 increased from 54.6 to 80.5 clpm.

Although the gains of Cohort 2 alone are encouraging, they are in places not much higher than when the cohort was the control group at baseline. As such, these endline gains cannot be conclusively attributed to the intervention alone. For example, Cohort 2 and grade 3 increased by 13.8 clpm from baseline to midterm (from 65.2 to 79.0 clpm) and increased by 13.1 clpm from midterm to endline (from 79.0 to 92.1 clpm). Although the results for letter knowledge are strong, this subtask is the earliest skill learned out of all of the subtasks and is least related to reading fluency.

The mean results for correct familiar words per minute (see **Table 7**) are unusual for an intervention study. For all grades and cohorts, mean scores increased from baseline to midterm, and then decreased from midterm to endline. The DID analysis therefore shows which cohort decreased the least. In grade 2, Cohort 1 decreased by 8.6 cwpm (from 18.6 to 10.0 cwpm) and Cohort 2 decreased by 1.5 cwpm (from 12.4 to 10.9 cwpm). With the possible exception of grade 3, all of these average pupils' scores are low, indicating that the pupils have had little time reading or being read to by others.

**Table 7. Summary of Correct Familiar Words Per Minute Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	2.4	8.7	3.9		
	2	5.8	9.1	4.8	0.5	0.1
2	1	5.0	18.6	10.0		
	2	7.1	12.4	10.9	7.0*	0.4
3	1	6.6	25.1	16.8		
	2	11.3	28.5	17.2	-2.9	-0.1

\* $p < 0.05$

The invented words subtask measured a pupil’s ability to blend sounds together to form a word, which is considered to be a critical pre-literacy skill. The results for invented words (**Table 8**) paint a picture similar to that for the results for ORF. The results between midterm and endline assessments were very flat or decreased. The only real exception again was in grade 2, Cohort 2, which improved from 1.2 to 3.4 cwpm. This score resulted in a DID of 4.5 cwpm; however, this result must be considered in the context that in Cohort 1, the scores decreased by 2.2 cwpm. The average scores for all of the grades indicate that pupils are struggling to sound out individual letters or groups of letters and cannot form words they have not seen before. As a result, reading a passage or text will be challenging and comprehension of the same text very unlikely.

### SAMPLE OF INVENTED WORDS

loz ep yat zam tob zom ras  
 mon jaf duz tam af ked ig  
 el tig pek dop zac ik

**Table 8. Summary of Correct Invented Words Per Minute Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE–MIDTERM )	EFFECT SIZE
1	1	0.2	0.9	0.5		
	2	0.4	0.7	0.8	0.6	0.2
2	1	0.3	3.2	1.0		
	2	0.9	1.2	3.4	4.5**	0.7
3	1	0.3	4.5	2.5		
	2	1.4	5.0	5.2	2.1	0.2

\* $p < 0.05$ \*\* $p < 0.001$

The listening comprehension subtask measures the ability of pupils to comprehend oral language. Listening comprehension includes the use of comprehension strategies, processing of language, and generation of appropriate replies that are also needed in comprehending written language. Deficiencies in listening comprehension abilities can lead to challenges in comprehension. To test listening comprehension, pupils were read a short passage and asked to answer five questions about the passage. **Table 9** shows consistency in that almost all the mean percent correct scores decreased between the midterm and endline assessments. More than anything, this finding highlights Liberia’s educational challenges resulting from the Ebola health crisis and subsequent school closures.

**Table 9. Summary of Listening Comprehension Percent Correct Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM )	EFFECT SIZE
1	1	25%	57%	39%	14%	0.4
	2	37%	47%	43%		
2	1	33%	63%	51%	14%	0.3
	2	36%	51%	53%		
3	1	44%	69%	63%	-6%	-0.2
	2	47%	73%	61%		

After completing the oral passage reading subtask, pupils were asked a series of comprehension questions based on the passage they just attempted to read. Because pupils were asked only questions that corresponded directly with the text that they were able to read (e.g., a pupil correctly reading 20 words of the reading passage would be given only those questions that corresponded to the first 20 words of the passage), the number of questions asked to individual pupils varied. The results of the reading comprehension subtask are presented in **Table 10**. Although gains were observed in all cohorts and grades between the baseline and midterm assessment, the mean percentage of correct words decreased for all subpopulations between midterm and endline. Although the results were confounded by situational challenges regarding the Ebola crisis, if we compare the listening and comprehension results, it is clear that pupils can comprehend a passage more easily when listening than when they are reading for themselves, highlighting that comprehension is less of an issue compared to fluent reading.

**Table 10. Summary of Reading Comprehension Percentage of Correct Words Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM )	EFFECT SIZE
1	1	2%	7%	3%	-3%	-0.2
	2	12%	10%	3%		
2	1	5%	15%	11%	4%	0.1
	2	9%	11%	10%		
3	1	10%	22%	19%	3%	0.1
	2	20%	24%	23%		

The initial sound discrimination subtask assessed the pupils' phonemic awareness (the ability to explicitly identify and manipulate sounds of language). Phonemic awareness has been found to be one of the most robust predictors of reading acquisition. Phonemic awareness is often used to identify pupils at risk for reading difficulties in the early elementary grades. As shown in **Table 11**, all of the Cohort 1 average percent scores decreased between the midterm and endline assessments, whereas the percent scores of all Cohort 2 pupils increased. In grade 2, Cohort 1 decreased by 3% and Cohort 2 increased by 15%, yielding a statistically significant DID of Cohort 1 versus Cohort 2 of 18%. This subtask has a clear association with the LTTP II curriculum and lesson plans, and as such, the lack of USAID intervention in Cohort 1 between the midterm and endline assessment might explain this, but it is very difficult reach a conclusion about it. The scores for initial sounds reflect the listening comprehension results; pupils hear and respond appropriately, but they are not yet able to make the step to reading text.

**Table 11. Summary of Initial Sounds Percent Correct Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	25%	57%	39%		0.4
	2	37%	47%	43%	14%	
2	1	33%	63%	51%		0.3
	2	36%	51%	53%	14%	
3	1	44%	69%	63%		-0.2
	2	47%	73%	61%	-6%	

Gains in average scores have often been explained by the reductions in the number of pupils scoring zero on specific skills. If these gains were driven only by a reduction in zero scores, then the intervention could be thought of as most successful for pupils without existing literacy skills. Exploring distributional shifts provides information beyond changes in mean and zero scores. Although a DID analysis informs us whether pupil scores have changed *on average*, distributions tell us where those changes in pupil scores have occurred. All else being equal, by comparing

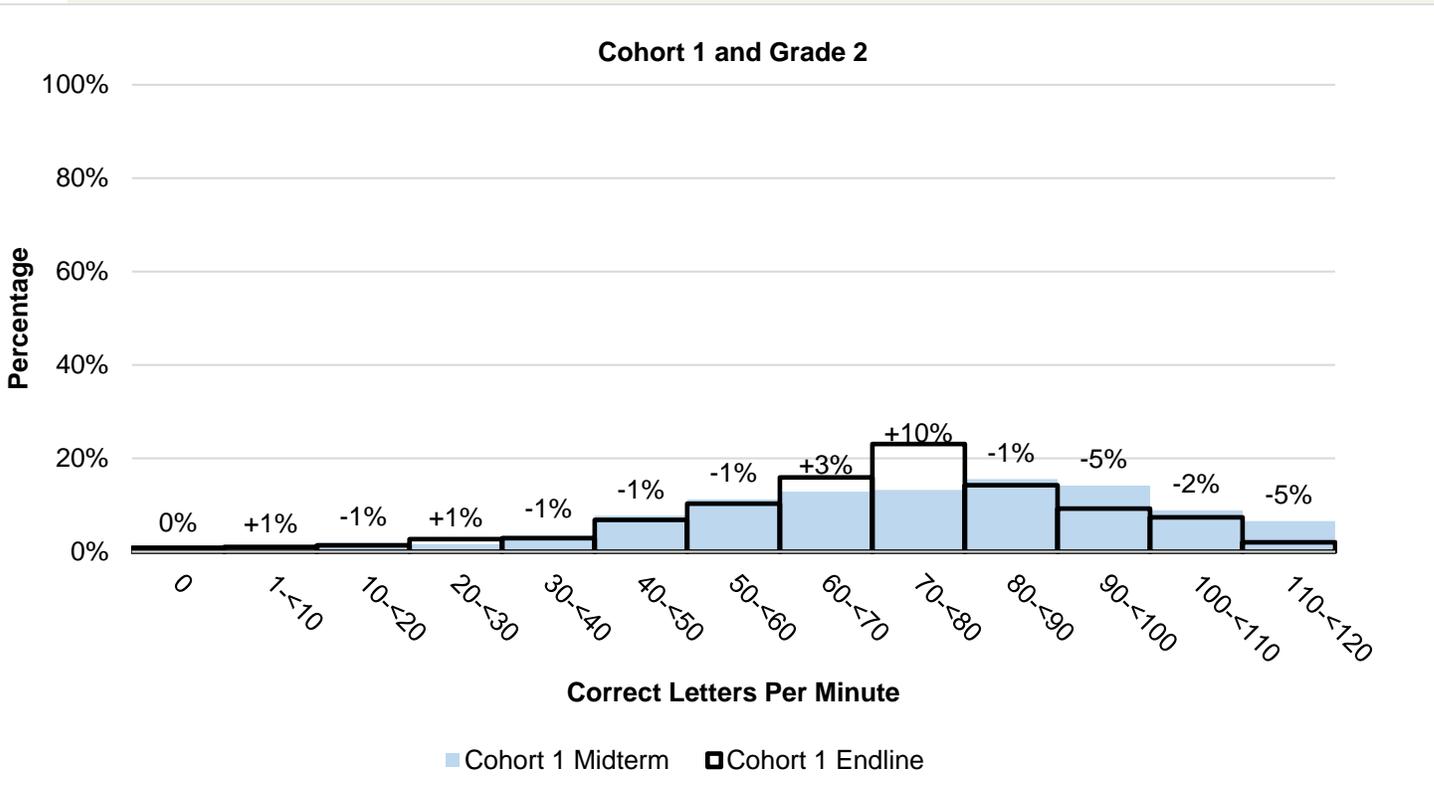
distributional shifts between the midterm and endline assessments, we can see the area(s) that realized the greatest impact from an intervention. **Figure 2** present the shifts in the entire distribution of scores for pupils in grade 2, between midterm and endline assessments, for correct letter identification per minute by cohort. Cohort 1 showed little change, other than a slight

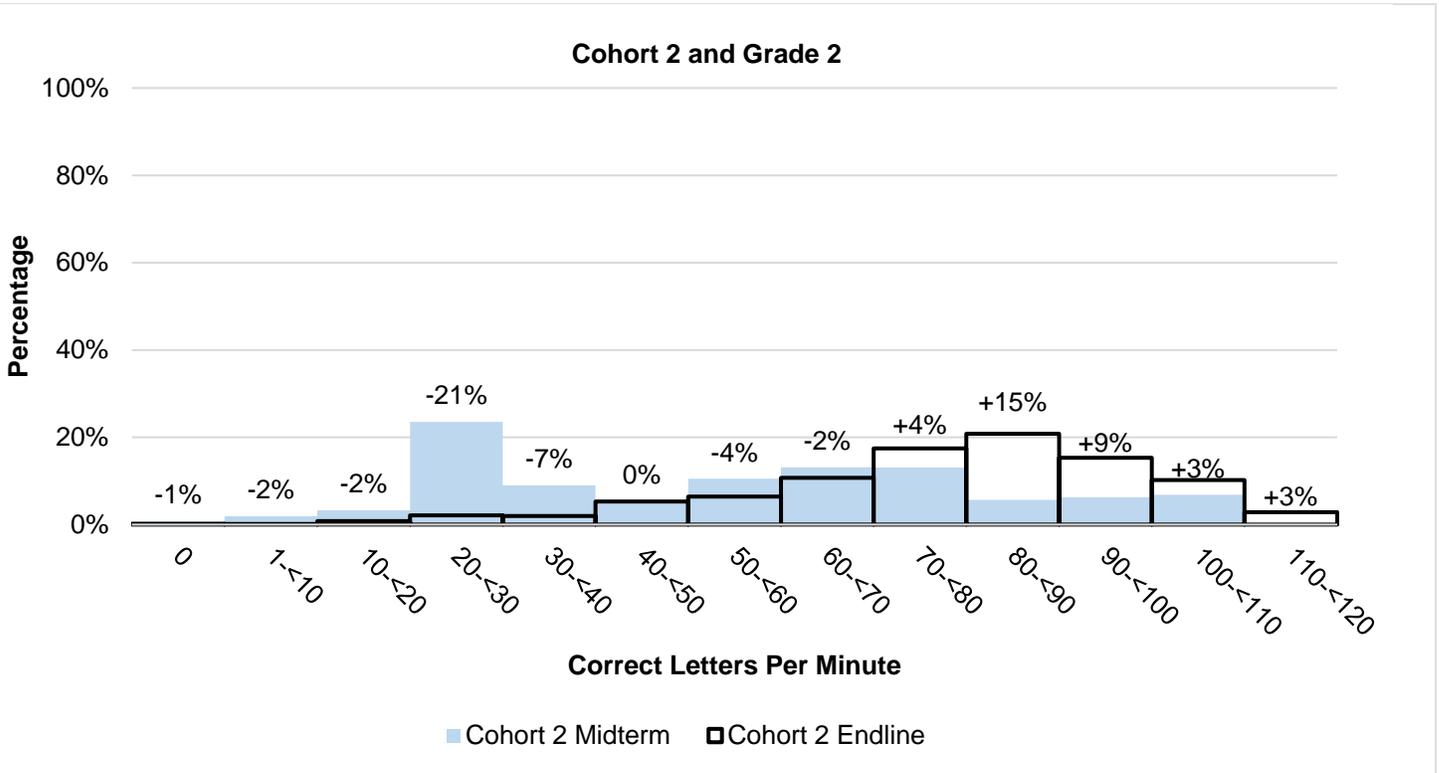
## INITIAL SOUNDS

For initial sound discrimination, pupils were asked to listen to a word such as “tour” and identify the first sound in that word, in this case “/t/.”

decrease at the top end of scores, whereas Cohort 2 showed a general trend of pupils identifying more correct letters per minute at endline than at midterm. It is important to note that all other changes in distributions that are not detailed in this section are included in *Annex E* of this report.

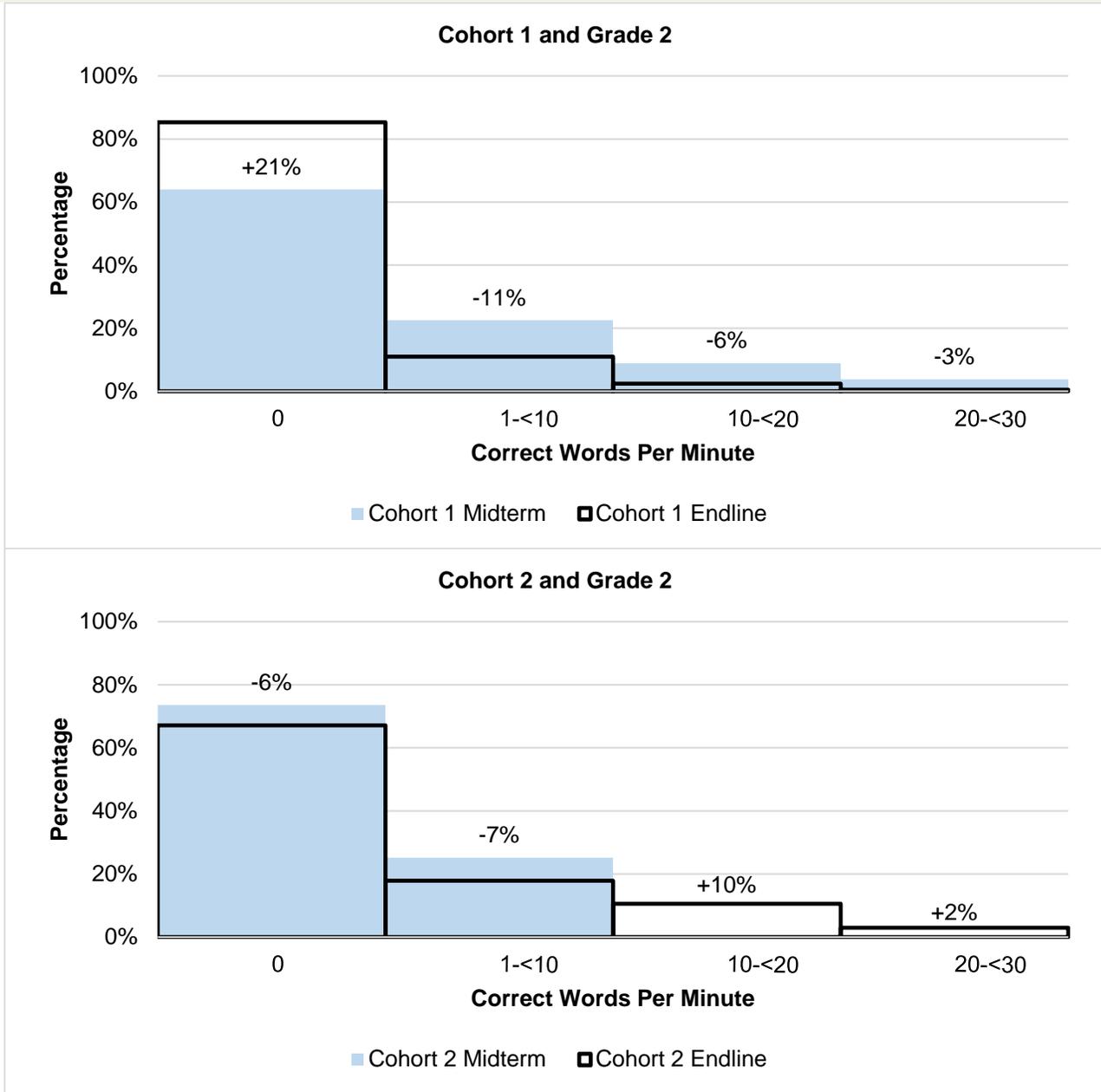
**Figure 2. The change in distribution between Cohort 1 and Cohort 2 for correct letters per minute, grade 2.**





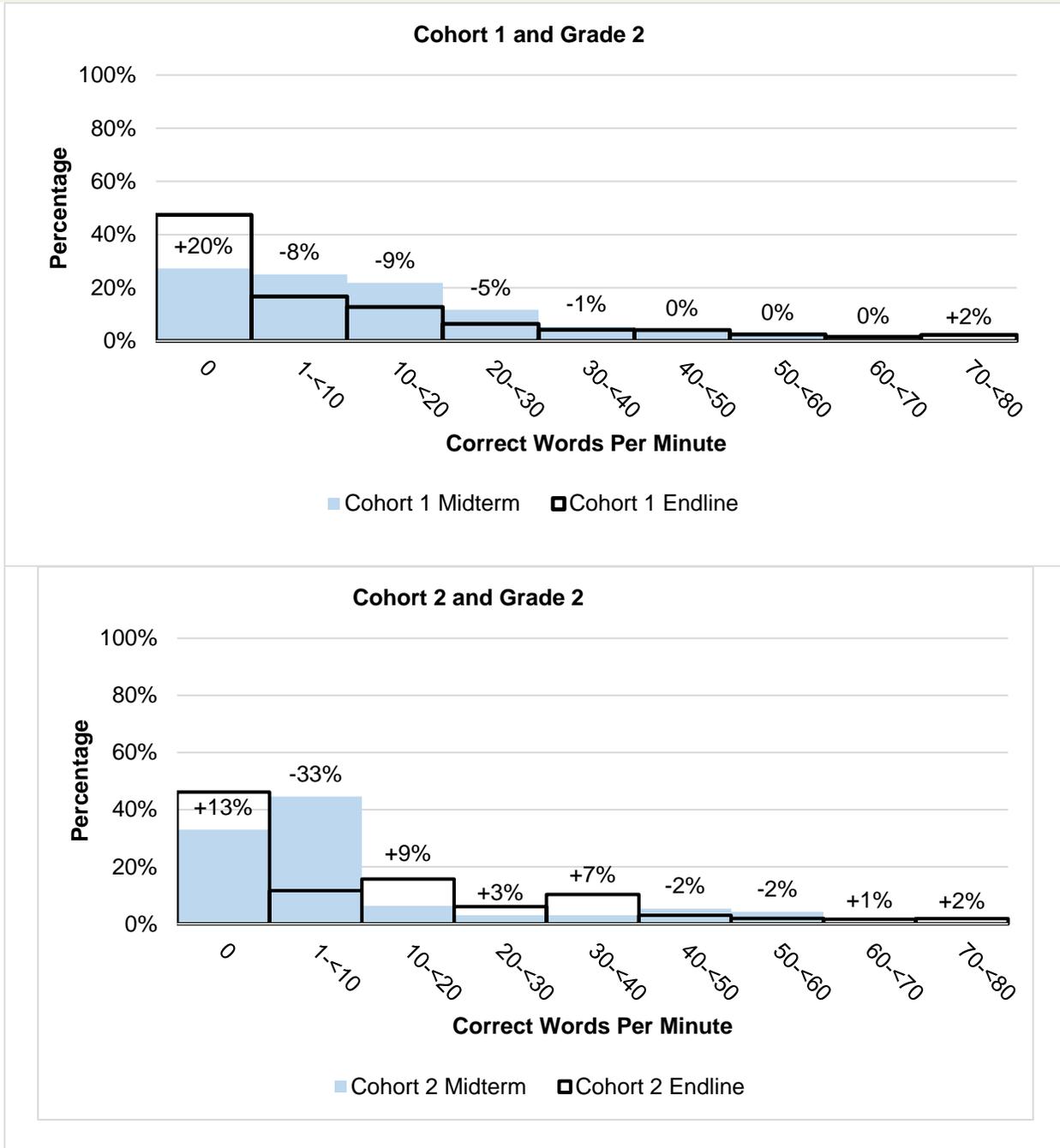
Distributions for invented words for grade 2 are shown in **Figure 3**. Cohort 1 showed an increase in the zero scores, along with a decrease in pupils scoring in the higher rates per minute. In contrast, Cohort 2 showed fewer pupils scoring in the lower rates of zero to 10 cwpm and more pupils scoring in the range of 10 to 30 cwpm. These results, although more descriptive in nature, indicate the potential of the effectiveness of the LTPP reading program, at least in grade 2.

**Figure 3. The change in distribution between Cohort 1 and Cohort 2 for correct invented words per minute, grade 2.**



Letter knowledge and invented word reading are prerequisite skills required for ORF. **Figure 4** shows the changes in distribution for ORF. Cohort 1 showed an increase in zero scores by 20% between the midterm and endline assessments. Cohort 2 showed rather consistent results: a percentage increase of 13% in zero scores, a decrease of 33% for pupils scoring between 1 and 10 cwpm, and an increase in scores for pupils scoring between 10 and 40 cwpm.

**Figure 4. The change in distribution between Cohort 1 and Cohort 2 for oral reading fluency words per minute, grade 2.**



### 3.3.4 SUMMARY OF EGRA RESULTS

All of the endline results have clearly been impacted by the Ebola outbreak. However, because the endline results look at the benefit of the LTTP II intervention program in Cohort 2 since the midterm assessment in 2013, some of the benefits can be seen by

those grades because the pupils had the most relevant support or time on task in developing their reading skills. For example, for Cohort 2, the results for grade 1 are poor in the endline assessment because those pupils entered into schools in 2015 for the first time and were only there a few weeks before being assessed. Grade 2 pupils in the same cohort received LTTP II intervention instruction in grade 1. As a result, most of the improvements for the Cohort 2 LTTP II intervention are evident in this grade. Finally, grade 3 pupils, who received instruction similar to the grade 2 pupils, failed to build on these skills with limited instructional time in grade 3.

The results for the Cohort 1 schools are more concerning. Cohort 1 schools only received limited support from LTTP II since the midterm assessment, and, as a result, it is likely that the mostly lower scores are because the quality of instruction decreased, although the extent of this cannot be determined because of the confounding issue of the Ebola outbreak and interruption in instruction. That being said, none of the pupils' average scores in the Cohort 1 or Cohort 2 schools indicate that pupils are reading with fluency and comprehension. Therefore, many of the scores for foundational skills such as reading invented words are not high enough to enable pupil reading.

### 3.4 EGMA RESULTS

#### 3.4.1 EGMA INSTRUMENTS FOR LIBERIA

**Table 12** summarizes the subtasks of the EGMA designed for Liberia. Pupil achievement was evaluated in several different areas of mathematics skill development. These areas

included number identification, quantity discrimination (greater than and less than), missing number (filling in a sequence of numbers), addition, subtraction, and word problems. For three skill areas (i.e., quantity discrimination, missing number, and word problems), pupil performance was assessed based on the percentage of correct responses. A higher percentage of correct responses indicated that the pupil had more fully developed the skill (because the items in each subtest covered a range of levels of difficulty, pupils who were answering more items correctly were, by definition, handling a greater range of ability levels for a given skill). Three of the mathematics subtests (i.e., number identification, addition, and subtraction) were timed, thus permitting an evaluation of how automatically the pupils performed the skill (as measured by the number of items correctly answered per minute).

#### SAMPLE ADDITION PROBLEMS

4	+	2		8	+	6
16	+	4		5	+	4
10	+	3		2	+	2
2	+	11		3	+	9

In the student instrument presented in **Annex C** of this report, the EGMA instrument follows the reading subtasks.

**Table 12. Subtasks of the EGMA instrument in Liberia**

SUBTASK	SKILL	DESCRIPTION— THE CHILD WAS ASKED TO ...
<b>Subtask that assessed more procedural (recall) type knowledge</b>		
1. Number identification	The ability to identify written numerals.	...say the names of numbers presented on a page with 30 numbers. The numbers ranged from one- to three-digit numbers. ( <i>Timed subtask</i> )
<b>Subtasks that assessed more conceptual (application) type knowledge</b>		
2. Quantity discrimination (number comparison)	This subtask requires the ability to make judgments about magnitude by comparing quantities represented by numbers.	... given two numbers, identify the number that is greater. The number pairs used ranged from a pair of single-digit numbers to five pairs of double-digit numbers and four pairs of three-digit numbers. There were 10 items. ( <i>Untimed subtask</i> )
3. Missing number (number patterns)	This subtask requires the ability to discern and complete number patterns.	... determine the missing number in a pattern of four numbers, one of which is missing. Patterns used included counting forward and backward by ones, twos, fives, tens, and hundreds. There were 10 items. ( <i>Untimed subtask</i> )
4a and 4b. Addition and subtraction level 1	This subtask requires knowledge of and confidence with basic addition and subtraction facts. It is expected that pupils should develop some level of automaticity and fluency with these facts because they need them throughout mathematics.	... mentally solve addition and subtraction problems, with sums and differences $\leq 20$ . The problems ranged from those with only single digits to problems that involved the bridging of the 10. There were 30 items per addition and subtraction subtask. ( <i>Timed subtask</i> )
5. Word problems	This subtask requires the ability to interpret a situation (presented orally to the pupil), make a plan, and solve the problem.	... solve problems presented orally using any strategy that he or she wanted, including the use of paper and pencil and/or counters supplied by the assessor. The focus of this subtask was on assessing a pupil's ability to interpret a situation, make a plan, and solve a problem. Therefore, the numerical values involved in the problem were deliberately small to allow for the targeted skills to be assessed without confounding problems with calculation skills that might otherwise impede performance. The problem situations used were designed to evoke different mathematical situations and operations. There were five items. ( <i>Untimed subtask</i> ).

### 3.4.2 ANALYSIS OF EGMA RESULTS

The EGMA analysis follows the same approach as the analysis for the reading subtasks. DID analysis comparing the midterm to endline results for Cohort 1 and

Cohort 2 was conducted, along with exploring the distributional changes in pupil subtask scores.

The addition problems (**Table 13**) did not require pupils to use more conceptual ideas, such as carrying tens, to derive the correct answer. Most of the pupils used a process technique, such as counting with their fingers, to obtain to the correct result.

As with the reading, the results for most subpopulations declined from midterm to endline, with the exception of grade 2 in Cohort 2, which increased from 7.7 to 8.7 correct problems per minute. Additionally, many of the endline results were lower than the baseline scores, underlying a concerning trend particularly for Cohort 2, where the USAID intervention took place. However, again, it is challenging to draw conclusions because of the confounding effect of the challenges that were experienced. The subtraction problems (**Table 14**) again showed exactly the same pattern of results, only with expected lower means of correct problems per minute. Grade 2 in Cohort 2 again showed the only increase, from 7.7 to 8.7 average correct problems per minute from the midterm to endline assessments.

**Table 13. Summary of Correct Addition Problems per Minute Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	6.3	6.2	5.0		
	2	7.7	6.3	4.6	-0.5	-0.1
2	1	8.3	8.6	7.2		
	2	12.1	7.7	8.7	2.4*	0.4
3	1	11.3	10.8	10.2		
	2	11.5	11.4	10.2	-0.5	-0.1

\* $p < 0.05$

**Table 14. Summary of Correct Subtraction Problems per Minute Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	4.9	4.1	2.8		
	2	6.4	4.3	2.8	-0.2	0.0
2	1	5.6	5.9	4.4		
	2	8.2	4.6	6.2	3.0	0.5
3	1	7.4	7.3	6.7		
	2	9.9	7.9	7.2	0.0	0.0

\* $p < 0.05$

The final procedural mathematics subtask was number identification (**Table 15**). In this section, the pupils were required to identify numbers from one to three digits in length. Again, the results showed outcomes similar to those of the addition and subtraction subtasks, where results were flat or declined in all grades and cohorts, with the exception of grade 2 in Cohort 2, which increased by 2.8 correct problems per minute between the midterm and endline assessments. The average scores for number identification were not high regarding the magnitude of the numbers; none of the numbers contained more than three digits.

## NUMBER IDENTIFICATION PROBLEMS

4 10 28 58 807 94 368 30 106  
 17 9 39 14 711 83 423 34 72 245  
 77 187 52 22 19 33 646 12 64  
 49 301

**Table 15. Summary of Correct Number Identification Problems per Minute Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	11.5	12.9	10.9		
	2	13.4	12.9	12.2	1.3	0.2
2	1	16.0	18.3	16.9		
	2	16.0	15.8	18.6	4.2*	0.4
3	1	20.1	22.3	22.5		
	2	17.4	22.4	23.1	0.5	0.0

\* $p < 0.05$

The first conceptual mathematics subtask was quantity discrimination (**Table 16**). In this section, pupils needed to identify the “greater” number from pairs of numbers. These numbers could be one to three digits in length. The results from this section indicated small, modest, or large gains between the midterm and endline assessments, with no average percentage of correct scores decreasing, unlike the procedural subtasks. Interestingly, the gains for Cohort 1 in grades 1 and 3 were higher than the gains for the respective Cohort 2 group. For example, grade 3 in Cohort 2 remained at 71% for both midterm and endline, whereas grade 3 Cohort 1 increased from 66% to 71% between midterm and endline. However, the only statistically significant DID and large effect size reported for quantity discrimination was for grade 2, in which that grade in Cohort 2 increased its mean problems percentage correct by 20 percentage points between midterm and endline, a gain of 19% over Cohort 1.

## SAMPLE WORD PROBLEM

Momo has 2 mangoes. His father gives him 5 more.  
How many does he have now?

**Table 16. Summary of Quantity Discrimination Problems Percent Correct Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE-IN-DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	36%	41%	50%		
	2	39%	42%	47%	-4%	-0.1
2	1	45%	58%	59%		
	2	36%	44%	64%	19%*	0.6
3	1	54%	66%	71%		
	2	52%	71%	71%	-5%	-0.2

\* $p < 0.05$  \*\* $p < 0.001$

The final mathematics subtask involved word problems (**Table 17**). The results indicated a struggle in the pupils' critical thinking and problem-solving skills, which were a curriculum focus. Most of the average-problems-correct percentages declined to the levels observed at baseline. For example in grade 1 in Cohort 1, the average declined from 41% to 31%, very close to the 33% observed at baseline. In addition, grade 1 in Cohort 2 declined from 40% to 28% (the baseline score had been 29%). These findings seem to reflect a lack of time on task in mathematics compared with time spent on reading skills. Teachers also received less support in mathematics as compared with reading.

**Table 17. Summary of Word Problems Percent Correct Comparing Cohort 2 and Cohort 1, by Grade**

GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE IN DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
1	1	33%	41%	31%		

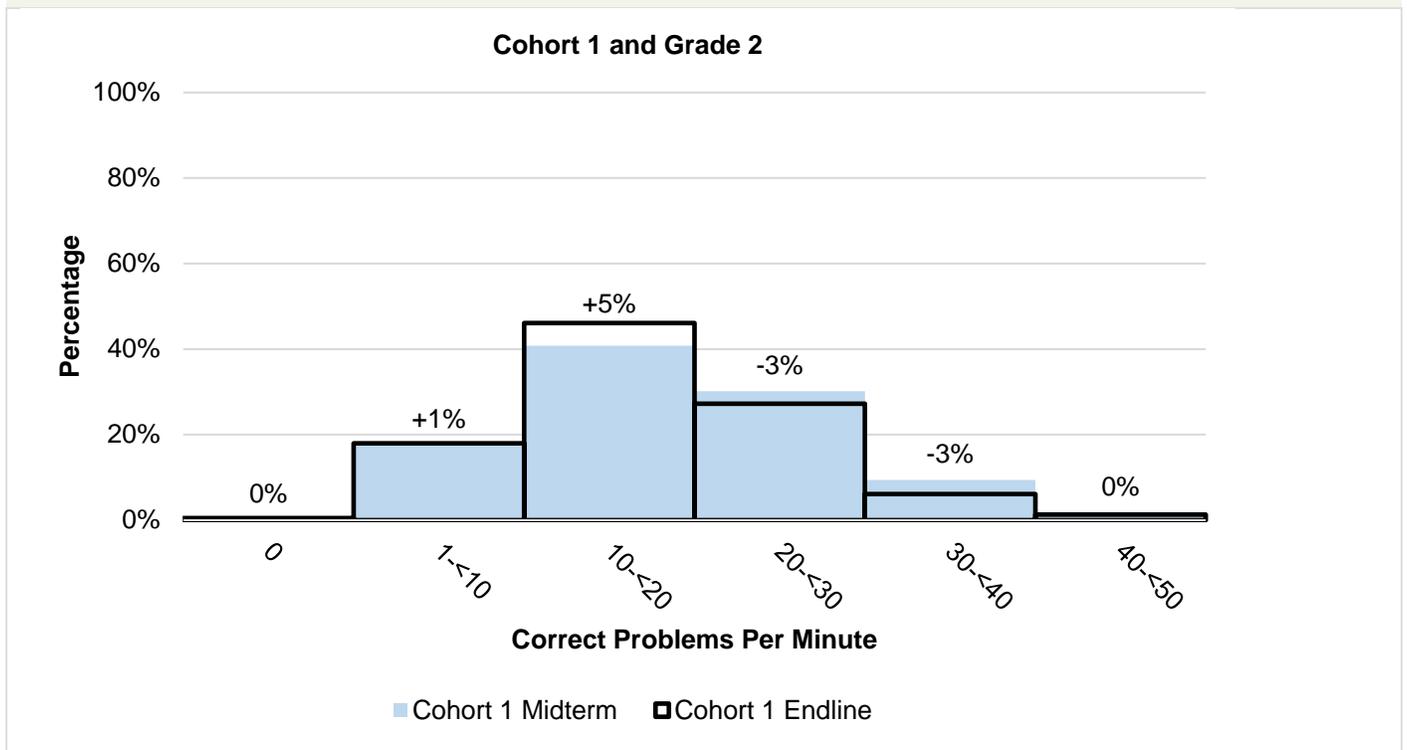
**Table 17. Summary of Word Problems Percent Correct Comparing Cohort 2 and Cohort 1, by Grade**

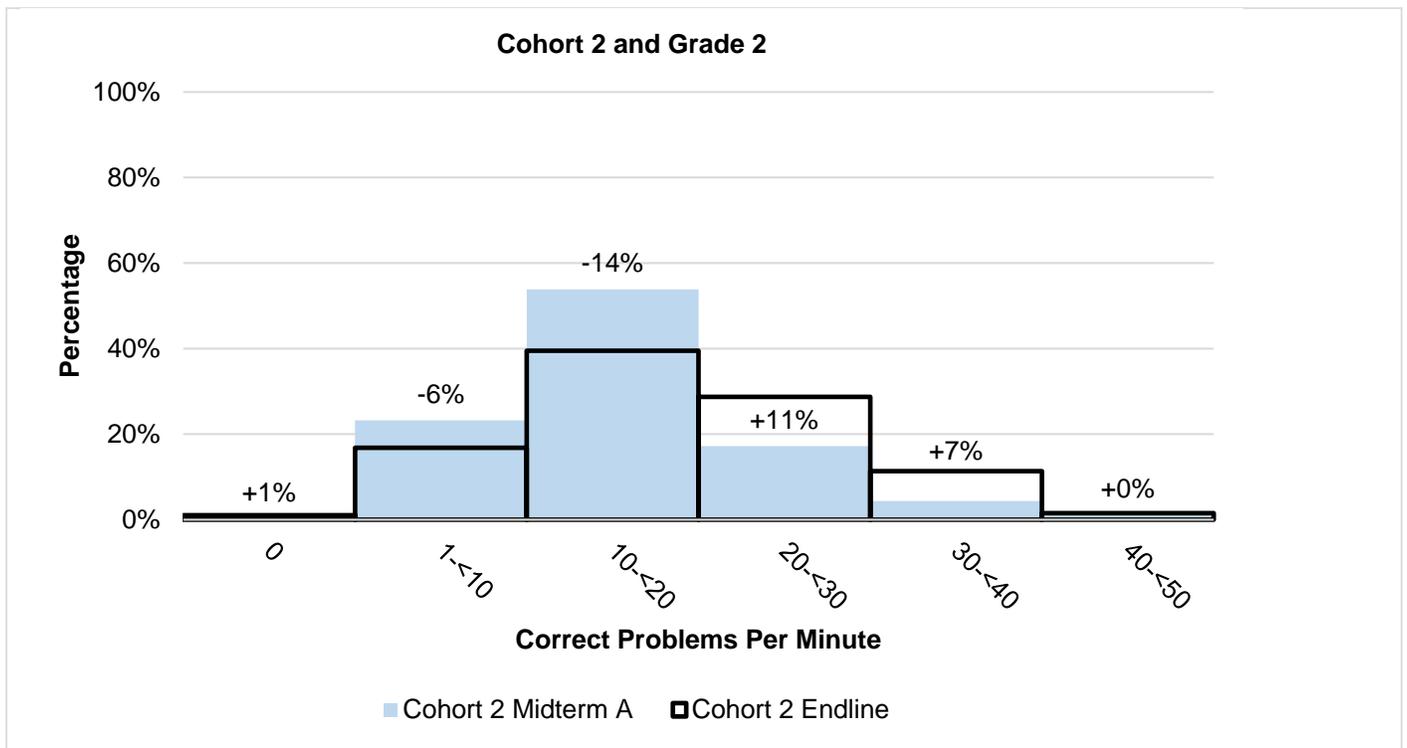
GRADE	COHORT	BASELINE	MIDTERM	ENDLINE	DIFFERENCE IN DIFFERENCE (ENDLINE-MIDTERM)	EFFECT SIZE
	2	29%	40%	28%	-2%	-0.1
2	1	48%	47%	39%		
	2	48%	49%	44%	3%	0.1
3	1	42%	54%	52%		
	2	62%	64%	50%	-12%*	-0.4

\* $p < 0.05$  \*\* $p < 0.001$

**Figure 5** demonstrates exactly this idea for the number identification subtask in grade 2. Although the Cohort 1 graphic in Figure 4 shows that there was very little change, the Cohort 2 graphic indicates a shift from pupils scoring between 1 and 20 correct problems per minute to between 20 and 40 correct problems per minute.

**Figure 5. The change in distribution between Cohort 1 and Cohort 2 for number identification per minute, grade 2.**





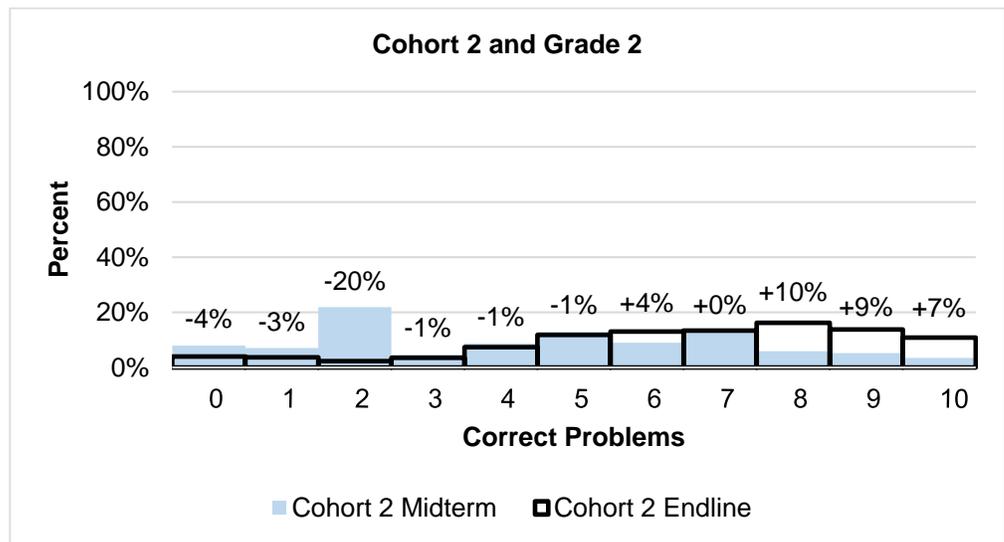
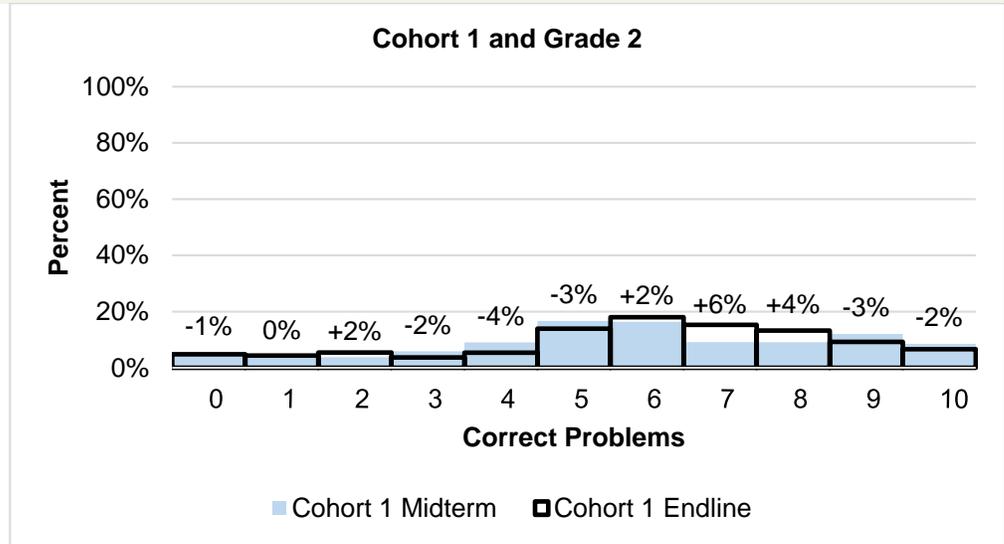
The distributions of pupil scores for the conceptual task of quantity discrimination are shown for grade 2 in **Figure 6**. The shifts in the distributions between midterm and endline for Cohort 1 do not show a distinct pattern; therefore, it is easy to draw the conclusion that the distributions are just displaying differences in natural variability. As shown in **Table 17**, the mean percentage scores for Cohort 1 were approximately the same. In contrast, Cohort 2 had a sharp decline in pupils scoring between 0 and 2 correct problems per minute and a corresponding increase in pupils scoring 8 or higher. Overall, the pupils' scores for mathematics demonstrated less improvement compared with reading for LTTP II, simply explained by the fact that LTTP II was more focused on reading.

### 3.4.3. SUMMARY OF EGMA RESULTS

The endline EGMA shows that the results for grades 1, for the most part, did not really improve compared with the midterm results in 2013. In many areas, the scores actually decreased. This finding is not surprising because we are comparing the 2015 results of pupils who had seven weeks of instruction to those who completed a full year before taking the 2013 midterm assessment. Grade 2 fares better in Cohort 2 schools only where the LTTP II intervention seems to have had a slight impact in spite of the Ebola outbreak. However, the DID results comparing grade 2 in Cohort 1 and Cohort 2 schools flatter the Cohort 2 results somewhat because most of the average scores decreased in Cohort 1 between the midterm and endline assessments. All of the grade 3 results show little improvement and are inconsistent when comparing

Cohort 1 and Cohort 2 because each cohort performs better on different subtasks; therefore, it is difficult to reach a definitive conclusion with grade 3.

**Figure 6. The change in distribution between Cohort 1 and Cohort 2 for quantitative comparison problems, grade 2.**



### 3.5 SSME RESULTS

As previously mentioned, the SSME is designed to provide a comprehensive description of school learning environments in Liberian schools and a portrait of the pupils who populate them. In addition, these instruments can be used to expound upon patterns seen in pupil performance on EGRA and EGMA. Both of these complementary purposes are explored in this section of the report.

### 3.5.1 SSME INSTRUMENTS FOR LIBERIA

**Table 18** summarizes the content of the SSME instruments designed for Liberia. Copies of these instruments are provided in **Annex C** of this report. Note that the SSME portion of the student instrument is titled “Student Context Interview.”

**Table 18. Summary of SSME Instruments for Liberia LTTP II**

INSTRUMENT	LEVEL OF DATA OBTAINED	DESCRIPTION OF DATA COLLECTED
Student Instrument	Student	Student characteristics, household background, access to books, reading habits, and teachers’ practices
Teacher Instrument	Teacher	Teacher characteristics, instructional practices, training, supervision, and academic expectations
Principal Instrument	School	Head teacher characteristics, supervision, resources and materials, teaching corps, and community involvement
School Inventory	School	Background data about the school (e.g., school size, location), enrollment, and student participation in the assessment

### 3.5.2 FINDINGS FROM PUPIL INTERVIEWS

After the administration of the EGRA and EGMA, pupils responded to questions from the Student Questionnaire, which remained largely identical to the instrument used for the midterm assessment. **Table 19** presents selected results from the Student Questionnaire from the midterm and endline for Cohort 1 and Cohort 2 schools, as well as the percentage-point change between the two assessments periods. It is worth re-emphasizing that the midterm and endline assessments were conducted in the same schools (Cohort 1 and Cohort 2), but on different pupils. Thus, the values shown in **Table 19** refer to means from different sets of pupils taken at different times. As with the EGRA and EGMA, because the intervention changed between the baseline, midterm, and endline assessments, the focus in **Table 19** is on the change between the two most recent evaluations.

**Table 19. Results from Pupil Interviews: Midterm and Endline**

INTERVIEW ITEM	MIDTERM (M)		ENDLINE (E)		CHANGE (E – M)	
	COHORT 1	COHORT 2	COHORT 1	COHORT 2	COHORT 1	COHORT 2
Speaks English at home	28%	23%	46%	47%	18%	24%
Parents read/write English	55%	49%	67%	62%	12%	13%

**Table 19. Results from Pupil Interviews: Midterm and Endline**

Parents cannot read/write	34%	38%	Not applicable	Not applicable	Not applicable	Not applicable
Has reading books at home	73%	63%	66%	65%	-7%	2%
Is read aloud to at home	59%	57%	49%	51%	-10%	-6%
Reads aloud at home	61%	61%	49%	50%	-12%	-11%
Takes books home from school	77%	58%	75%	76%	-2%	18%
Has a library at school	37%	31%	29%	23%	-8%	-8%
Teacher reads aloud often	95%	93%	93%	88%	-2%	-5%
Teacher practices letter sounds	93%	72%	86%	91%	-7%	19%
Teacher has pupils read aloud	94%	94%	90%	87%	-4%	-7%
Teacher assigns reading for home	95%	93%	89%	85%	-6%	-8%
Has repeated a grade	70%	69%	74%	78%	4%	9%
Missed school last week	37%	44%	39%	31%	2%	-13%
Eats lunch at school	68%	64%	65%	64%	-3%	0%

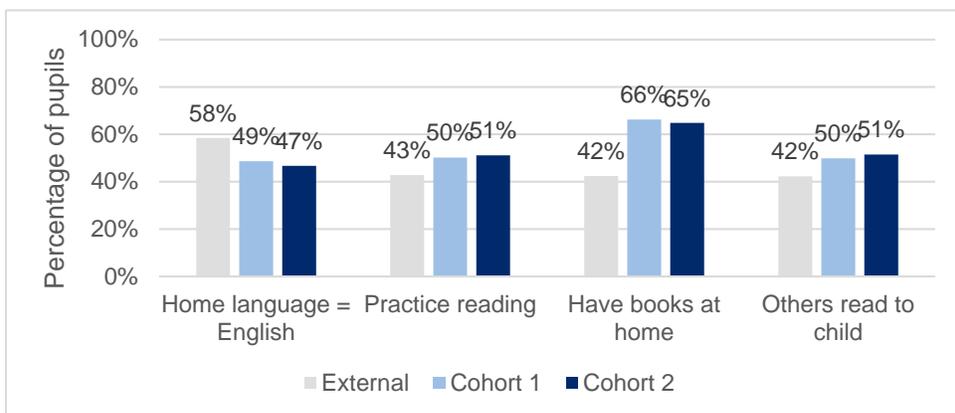
As shown in **Table 19**, several notable changes in pupil background and pupil-reported classroom characteristics have been registered since the midterm assessment. With regard to home language, significantly more pupils in both Cohort 1 and Cohort 2 schools (18 and 24 percentage points more, respectively) reported during the endline assessment that they speak English at home and that it is the language in which their parent(s) read and write (12 and 13 percentage points more, respectively), as compared to the midterm assessment.

Conversely, **Table 19** shows that a lower proportion of pupils reported reading at home at the endline compared to the midterm assessment: only approximately 50% of pupils in both Cohort 1 and 2 schools reported being read to at home or practicing reading at home, compared to approximately 60% at midterm. Reading at home is an extremely important activity in the development of children's reading ability and has been found to predict reading and other academic outcomes; thus, the apparent change in home reading practices from midterm to endline is concerning. Moreover, approximately half of the pupils reportedly were not engaged in these activities at all, although the act of sharing a book or other reading material is a task that does not command a great amount of time on the part of families; however, this could also be due to family disruption because of the Ebola crisis.

Two positive school-level changes were reported by children in Cohort 2 schools: significantly more pupils reported taking books home from school and practicing letter sounds with their teachers at endline as compared to midterm (18 and 19 percentage points more, respectively). This trend erases the gap in reading resources and teacher actions observed at the midterm assessment in these variables between Cohort 1 and 2 schools. Although most pupils reported not having access to books from a library at school, schools in Cohorts 1 and 2 had managed to provide access to other literature that their pupils could borrow (though not all pupils may read these materials at home, as previously discussed).

**Figure 7** compares several of these variables to the respective means found in External Cohort schools. As shown in **Figure 7**, more pupils in External Cohort schools reported having English as their home language. However, pupils in Cohort 1 and Cohort 2 schools reported having books at home, reading to others, and being read to with greater frequency than their peers in External Cohort schools. Although a significant proportion of pupils still did not have books at home or did not practice reading (a finding which should be addressed), the results in **Figure 6** suggest that Cohort 1 and Cohort 2 schools were somewhat successful in getting pupils access to reading materials and in engaging them in reading activities outside of school.

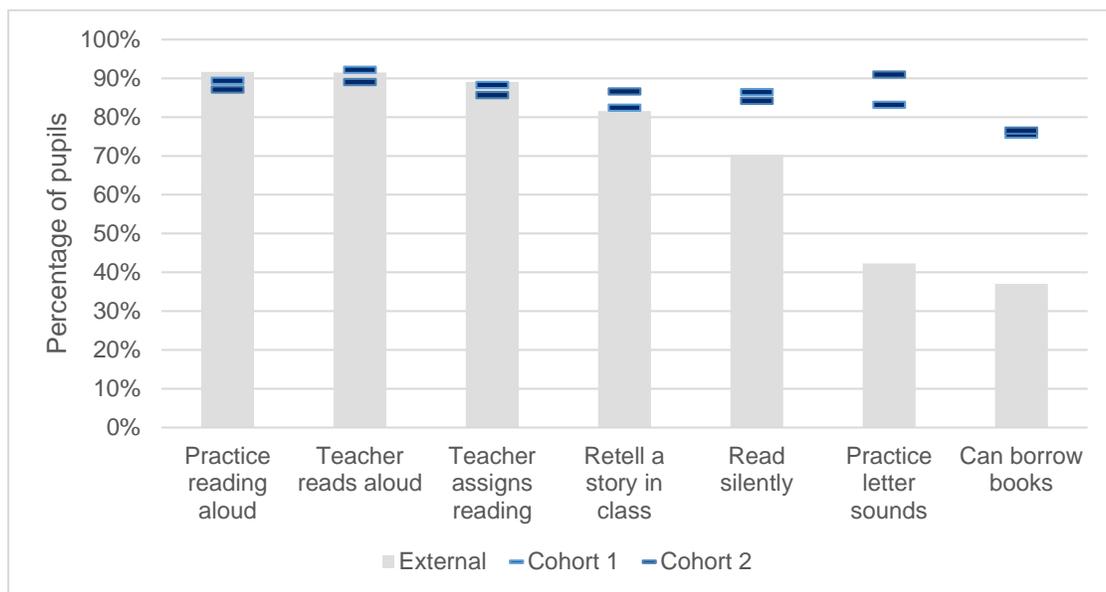
**Figure 7. Reading practices at home of interviewed children (endline).**



**Figure 8** compares pupil-reported classroom activities in Cohort 1 and Cohort 2 schools to External Cohort schools. Although there is little difference between Cohort 1 and Cohort 2 in terms of teachers' actions, and most pupils interviewed at endline reported engaging in these reading activities in class, pupils in External Cohort schools were much less likely to practice reading silently, practice letter sounds, or borrow books than their counterparts in Cohort 1 and Cohort 2 schools. This again

suggests that pupils' access to literacy materials and their engagement in some literacy activities (especially letter sounds) is greater in these schools.

**Figure 8. The classroom activities related to reading among interviewed children.**



### 3.5.3 FINDINGS FROM TEACHER INTERVIEWS

Teachers from grades 1, 2, and 3 were asked to respond to a questionnaire at the endline assessment. On average, 3.9 Cohort 1 teachers and 5.3 Cohort 2 teachers were interviewed per school.

**Table 20** presents means of selected questionnaire items for teachers in Cohort 1 and Cohort 2 schools from midterm and endline evaluations, as well as the percentage point change.

**Table 20. Results from Teacher Interviews: Midterm and Endline**

INTERVIEW ITEM	MIDTERM (M)		ENDLINE (E)		CHANGE (E – M)	
	COHORT 1	COHORT 2	COHORT 1	COHORT 2	COHORT 1	COHORT 2
Is a woman	9%	13%	24%	26%	15%	13%
Has C Certificate	61%	47%	70%	72%	9%	25%

**Table 20. Results from Teacher Interviews: Midterm and Endline**

Has attended in-service education and training in the past year	92%	78%	69%	86%	-23%	8%
Has had training on reading	84%	50%	85%	88%	1%	38%
Has had training on mathematics	71%	47%	66%	83%	-5%	36%
Received support visits on how to teach reading	79%	22%	54%	82%	-25%	60%
Received support visits on how to teach mathematics	73%	73%	37%	74%	-36%	1%
Reports that principal observes classes daily	87%	82%	64%	75%	-23%	-7%
Received an external inspection in past year	88%	74%	75%	75%	-13%	1%
Showed a good example of the day's lesson plan	20%	11%	59%	68%	39%	57%
Uses official curriculum frequently	24%	27%	24%	27%	0%	0%
Has a teacher's guide	90%	87%	94%	92%	4%	5%
In past 5 days, has frequently had pupils						
• Repeat letters/words	32%	21%	24%	29%	-8%	8%
• Sound out unfamiliar words	28%	28%	24%	33%	-4%	5%
• Read aloud	33%	22%	26%	28%	-7%	6%
• Learn meaning of new words frequently	38%	23%	27%	35%	-11%	12%
• Retell a story they read during the week	35%	20%	28%	37%	-7%	17%
• Read on their own in school	37%	33%	36%	23%	-1%	-10%
• Perform a reading assignment at home	47%	21%	36%	36%	-11%	15%
In past 5 days, has						
• Had pupils complete problems at the board	50%	40%	53%	72%	3%	32%
• Had pupils practice problems in partners	48%	32%	48%	50%	0%	18%
• Had pupils perform individual work at their seats	38%	45%	35%	38%	-3%	-7%

**Table 20. Results from Teacher Interviews: Midterm and Endline**

• Called on pupils to answer mathematics questions orally	42%	25%	48%	35%	6%	10%
Believes that at end of year the pupils should						
• Read grade-level stories	63%	61%	67%	70%	4%	9%
• Sound out words	47%	34%	44%	48%	-3%	4%
• Understand stories they read	55%	37%	49%	64%	-6%	27%
• Know letter names	25%	21%	29%	39%	4%	18%
• Perform one-digit addition and subtraction problems	67%	55%	57%	76%	-10%	21%
• Know simple fractions	36%	29%	20%	22%	-16%	-7%
• Tell time	17%	10%	16%	10%	-1%	0%
• Make change	15%	15%	8%	16%	-7%	1%

**Table 20** suggests that the background characteristics of teachers interviewed may have changed from midterm to endline: significantly more teachers were women and held a C certificate. Nevertheless, as before, most teachers were men (three out of four). With regard to training, most teachers reported attending some form of in-service education and training (INSET) during the past academic year (69% in Cohort 1 schools and 86% in Cohort 2 schools), although this was a substantial reduction in Cohort 1 schools from the midterm assessment (when the rate was 92%). Most teachers in both Cohort 1 and 2 schools also reported that they completed training specific to reading and mathematics instruction. For Cohort 2 schools, this finding marked a substantial increase from midterm—38 percentage points higher for reading and 36 percentage points for mathematics training). However, it is notable that mathematics training was less frequently reported by teachers in Cohort 1 schools, suggesting that this content area was not receiving the same amount of attention as was reading.

Notable changes were observed in supervision and instructional support in Cohort 1 and 2 schools, as reported by teachers. Overall, it appears the Cohort 1 teachers received less support (in the form of support visits, observations, and external inspections) at endline compared to midterm, whereas Cohort 2 teachers received substantially more support compared to midterm. In Cohort 1 schools, nearly half of teachers did not receive visits to provide reading support, two-thirds did not receive

visits to provide mathematics support, and one-third was not observed daily by principals (comparative values in Cohort 2 schools of teachers who did receive such visits were approximately 75%). These findings are in line with the fact that coaches were not supporting Cohort 1 schools after midterm assessment, but also reiterates the reality that the MOE does not provide regular support to schools. Although this is not surprising given the nature and timing of the intervention, it is nevertheless notable how quickly Cohort 1 schools were reverting back to prior behavior patterns. Although this assessment did not track teacher transfer, it is likely that the decline in the application of training and skills is because of teacher transfer.

In terms of instructional resources, most teachers in Cohort 1 and Cohort 2 schools reported having a teacher's guide. In addition, most Cohort 1 (59%) and Cohort 2 (68%) teachers who reported having a lesson plan were willing and able to show a good quality instructional plan. These percentages are a substantial increase over the midterm means of 39 percentage points for Cohort 1 and 57 percentage points in Cohort 2 schools). However, most teachers admitted to *not* following the official curriculum (76% and 73% of Cohort 1 and 2 schools, respectively).

## TRACKING PUPIL PERFORMANCE

“Although everyone made bold claims about the programs’ impact on improving children’s reading and math skills, no one could cite any hard evidence to support the assertion. Moreover, when the assessment team asked about assessing and tracking [pupil] performance, no one could effectively articulate the [pupil] monitoring system or rattle-off data from that system, suggesting that more work needs to be done in this area. For example, teachers need to regularly track each [pupil’s] progress toward the reading and math standards establish by the education system, but few teachers had any records on [pupil] performance.”

—External program evaluation report

Turning to teachers’ actions and beliefs, a similar pattern is evident in **Table 6**, as was observed in the level of support provided to teachers in Cohort 1 and 2 schools. The is that the means for Cohort 1 schools appear to be (slightly) lower than at midterm, whereas those for Cohort 2 schools appear to be (slightly, although in some cases much) higher than at midterm. Teachers were asked whether, in the past five days, they frequently had pupils practice six distinct types of reading activities and whether they had assigned pupils reading tasks at home. First, it is notable that most teachers reportedly did not frequently engage pupils in these literacy activities, which are tied to reading skills and subtasks on EGRA. In most cases, two out of every three teachers did not frequently lead such activities. Second, the frequency of these activities appeared to be declining in Cohort 1 schools, but was increasing slightly in Cohort 2 schools, which is in line with the LTTP design and the fact that if schools are not

supported on a regular basis, then they will revert back to old practices, which is the case with Cohort 1 schools.

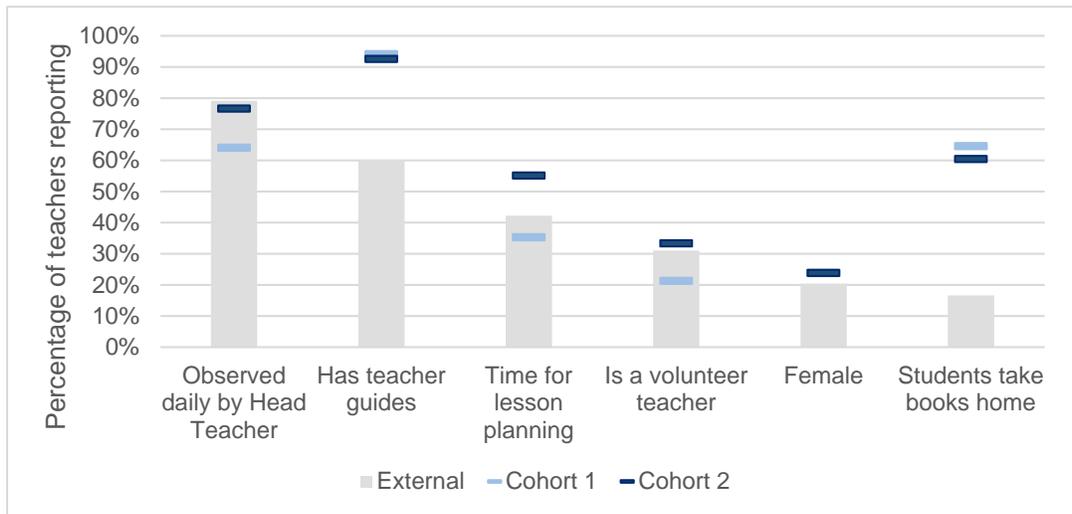
The frequency of various forms of mathematics practice, too, appeared to be on the rise in Cohort 2 schools. However, with the exception of solving problems at the board, the majority of teachers in both Cohort 1 and 2 schools did not engage their pupils in group-based problem solving, individual practice, or oral responses.

On aggregate, many teachers in Cohort 1 and Cohort 2 schools did not frequently lead their pupils in the literacy and mathematics activities of interest. In particular, the proportion of teachers allotting time to pupils for reading in class could be increased.

Teachers were also asked about their beliefs—specifically, whether they believed that pupils should be able to perform specific tasks by the end of the school year. In response to that question, most teachers in Cohort 1 (67%) and Cohort 2 (70%) schools affirmed that pupils should be able to read grade-level stories. In addition, most teachers in Cohort 1 (57%) and Cohort 2 (76%) said that the pupils should perform one-digit addition and subtraction problems. However, the majority of the teachers responding to the question did not believe that their pupils should be able to sound out words, know letter names, tell time, know simple fractions, and perform multiplication and division. Although the latter skills may be advanced for younger grades (recall that grade 1, 2, and 3 teachers were interviewed), knowing letter names and sounding out words should certainly be mastered by children of this age. Awareness of the importance of these skills appears to have increased among Cohort 2 teachers, but, interestingly, seems to have fallen among Cohort 1 teachers.

**Figure 9** displays selected teacher characteristics and classroom resources (as reported by teachers) for teachers in Cohort 1 and Cohort 2 schools, as well as a set of External Cohort schools that serves as a reference marker.

**Figure 9. Teacher characteristics and classroom resources.**



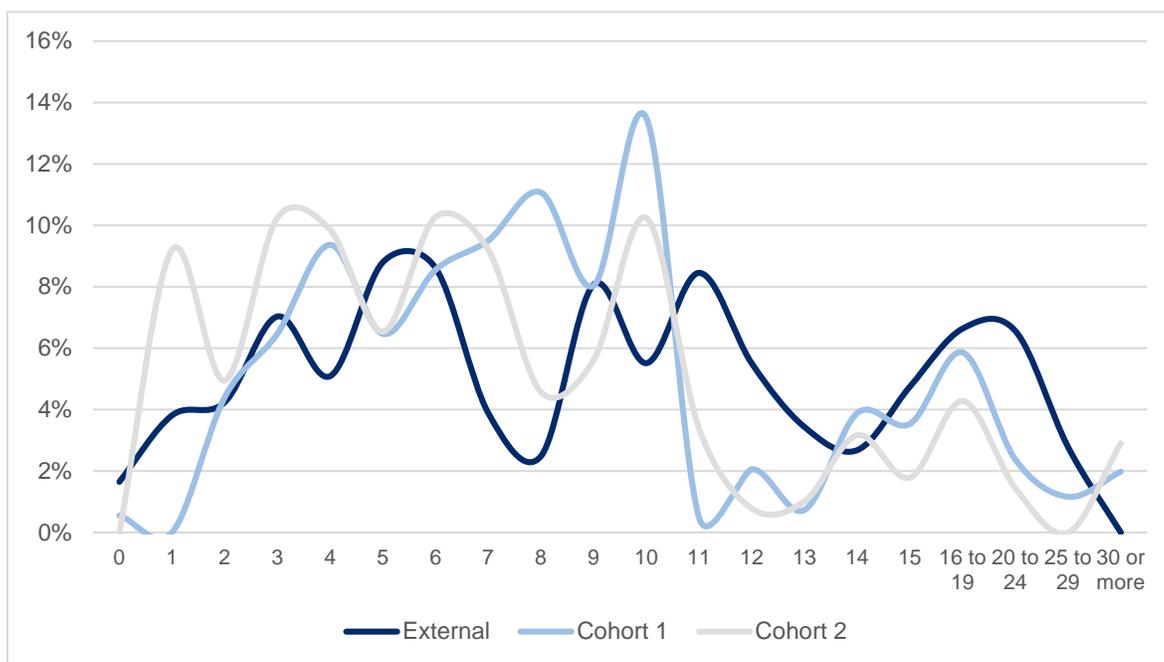
As shown in **Figure 9**, program schools (i.e., Cohorts 1 and 2) differed markedly from the External Cohort schools in terms of accessibility of teacher’s guides. Specifically, 94% of Cohort 1 and 93% of Cohort 2 teachers reported having a guide, as compared to 60% of teachers in External Cohort schools). In addition, approximately two-thirds of program Cohort 1 and 2 teachers reported that pupil reading materials were available for pupils to take home, as compared to only 17% of External Cohort schools). In addition, teachers in Cohort 2 schools were more likely (55%) to report having enough time for lesson planning than their peers in Cohort 1 schools (35%) or External Cohort schools (42%).

These differences are significant, instructionally and pedagogically speaking. The presence of teacher’s guides aids teachers in their attempt to master the material and practice high-quality teaching practices. Time allotted for lesson planning can help ensure that teachers plan well for subsequent lessons and have time to reflect on previous ones. Although Cohort 2 schools reportedly were somewhat better in this regard, the time allotted for lesson planning could be improved in all schools studied here. Pupil access to reading materials is essential for practicing reading outside of school and can augment the reading materials that pupils have in their home with appropriately leveled resources. However, most teachers in program schools reported that they were observed by the head teacher on a daily basis (although the frequency [64%] was somewhat lower in Cohort 1 schools) and relatively few teachers (less than one-third) reported that they were volunteer teachers.

The years of experience reported by teachers (**Figure 10**) appeared to somewhat differ. A significant minority of teachers in Cohort 2 schools (34%) counted less than 5

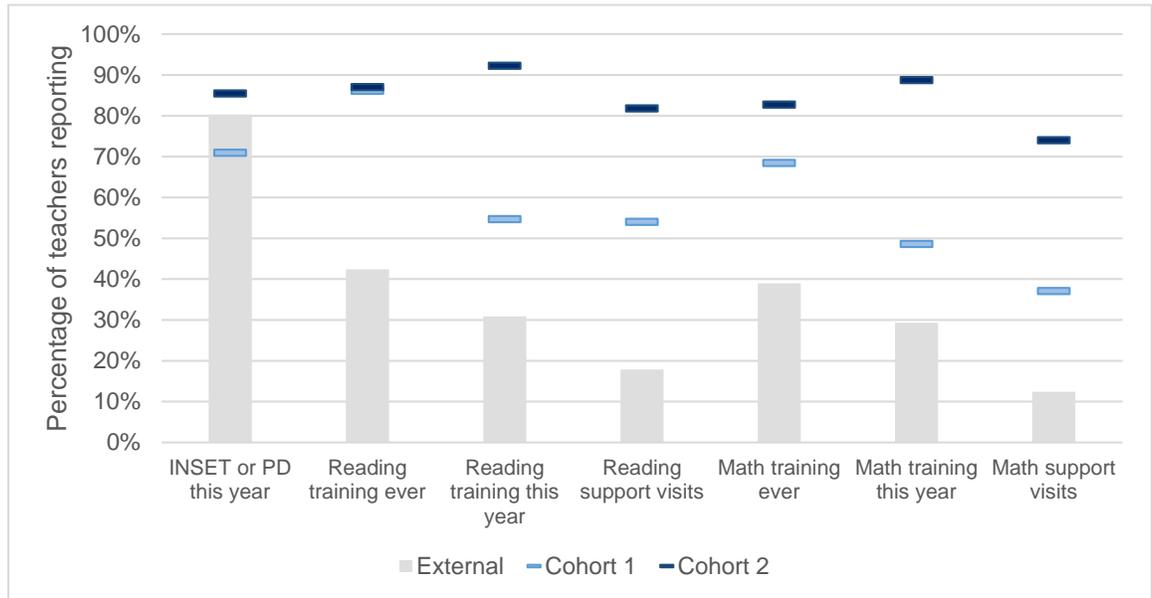
years of experience. However, in Cohort 1 schools, 51% of teachers reported having between 5 and 10 years of experience (compared to 40% in Cohort 2 and 29% in External Cohort schools). External Cohort schools appeared to have a larger proportion of teachers with more than 10 years of experience than program schools.

**Figure 10. Years of experience as reported by the teachers.**



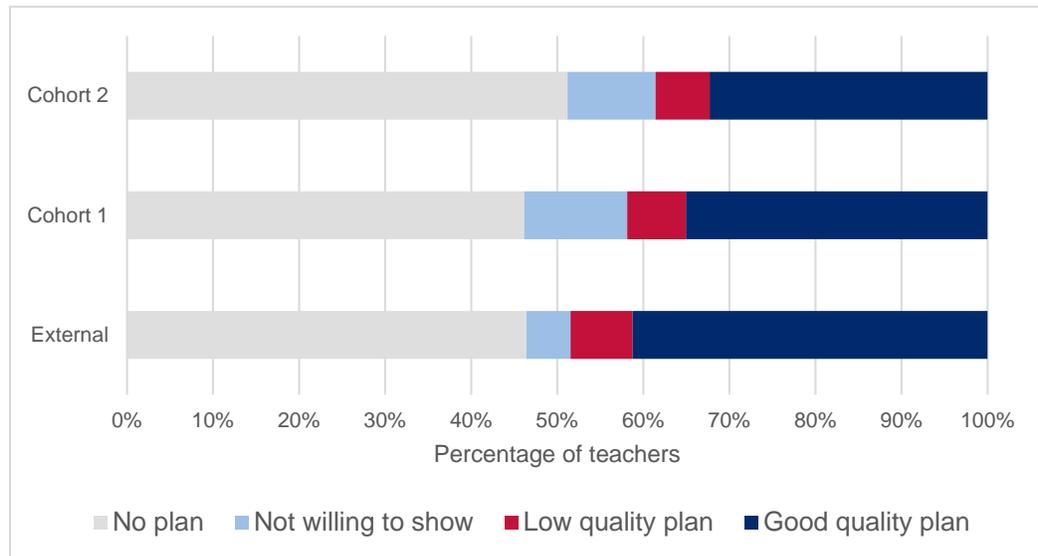
Program schools also differed from each other and from External Cohort schools in terms of training and support provided, as reported by teachers. **Figure 11** shows the proportion of teachers who reported completing specific types of reading or mathematics training and pedagogical support visits. Although there was not a substantial difference between the proportion of teachers who reported receiving general INSET or other PD provided during the past year, **Figure 11** suggests that reading- and mathematics-specific training opportunities and support were much more common in program schools. Interestingly, the provision of content-specific training differed between cohort schools as well. Specifically, teachers in Cohort 2 schools reported having reading and mathematics training and having support visits with greater frequency than their Cohort 1 counterparts. This finding may have been a result of the progression of the intervention program, but it is notable how Cohort 1 schools appear to be reverting to training and support patterns observed in External Cohort schools.

**Figure 11. Teacher training and support in reading and mathematics instruction.**



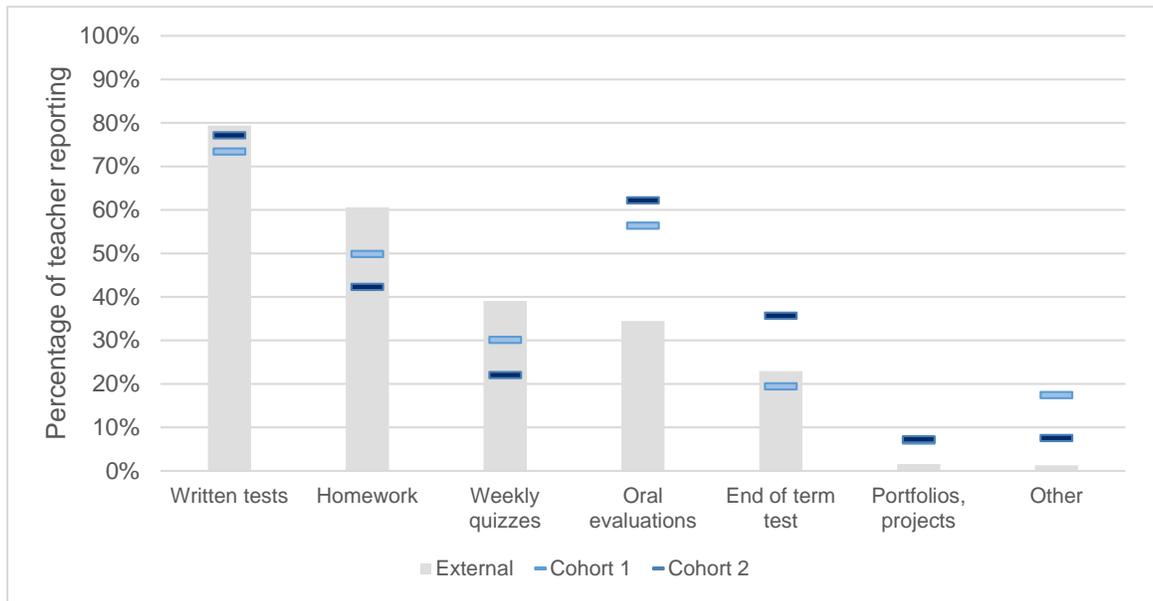
Given that approximately half of the teachers in the program and External Cohort schools reported not having time allocated to lesson planning, it is not surprising that, on the day of the assessment, approximately 60% of these teachers either did not have a lesson plan or were unwilling to show it to assessors (**Figure 12**). Very few (less than 10%) had a low-quality plan, whereas teachers in Cohort 1 (33%), Cohort 2 (35%), and External Cohort (42%) schools had a high-quality plan. Although lesson planning was not an explicit part of the program, it is interesting that a larger proportion of teachers in Cohort 1 and Cohort 2 schools did not have a lesson plan to share with assessors. Lesson planning is an integral part of a professional routine for preparing to teach a particular lesson, clarifying learning objectives and how they will be assessed, and analyzing mistakes pupils are likely to make when encountering the lesson content.

**Figure 12. Teachers' lesson plans on the day of the assessment.**



Teachers also reported how they tend to measure pupil progress (see **Figure 13**) against learning objectives. In External Cohort schools, written tests were the most common vehicle for measuring progress (79%), followed by homework (61%), weekly quizzes (39%), and oral evaluations (34%). In program schools, however, the oral evaluations were used more frequently (62% of Cohort 2 teachers and 56% of Cohort 1 teachers). This finding likely refers to the use of mini-EGRAs conducted by the teachers that were part of the program. If that is indeed the case, then it is worth noting that more than one out of three teachers in program schools did not use the mini-EGRA to evaluate pupils' reading abilities. This finding is in line with the fact that teachers abandoned the use of mini-EGRAs in the later years of the intervention because they deemed them as extra work and time consuming.

**Figure 13. How teachers measure pupil progress.**



Note: Values sum to more than 100% because more than one response was possible.

### 3.5.4 FINDINGS FROM HEAD TEACHER/PRINCIPAL INTERVIEWS

Head teachers were interviewed and responded to questions relating to their academic backgrounds and characteristics, school resources, community participation, and instructional leadership. **Table 21** presents selected responses from the midterm and endline assessments for both Cohort 1 and Cohort 2 schools.

**Table 21. Results from Head Teacher Interviews: Midterm and Endline**

INSTRUMENT ITEM	MIDTERM		ENDLINE		CHANGE	
	COHORT 1	COHORT 2	COHORT 1	COHORT 2	COHORT 1	COHORT 2
Has tried to improve reading: If so, because	95%	93%	99%	100%	4%	7%
Saw other schools doing it	18%	8%	4%	8%	-14%	0%
Independently thought they needed to do it	55%	47%	32%	69%	-23%	22%
The Ministry told them to do it	8%	11%	30%	25%	22%	14%

**Table 21. Results from Head Teacher Interviews: Midterm and Endline**

Teachers received training on how to teach reading	40%	8%	51%	41%	11%	33%
Parents now show more interest in reading	63%	54%	84%	66%	21%	12%
Whole community shows more interest in reading	67%	60%	84%	82%	17%	22%
Pupils show more interest in reading	76%	76%	92%	96%	16%	20%
There is more discussion about improving reading	80%	56%	100%	97%	20%	41%

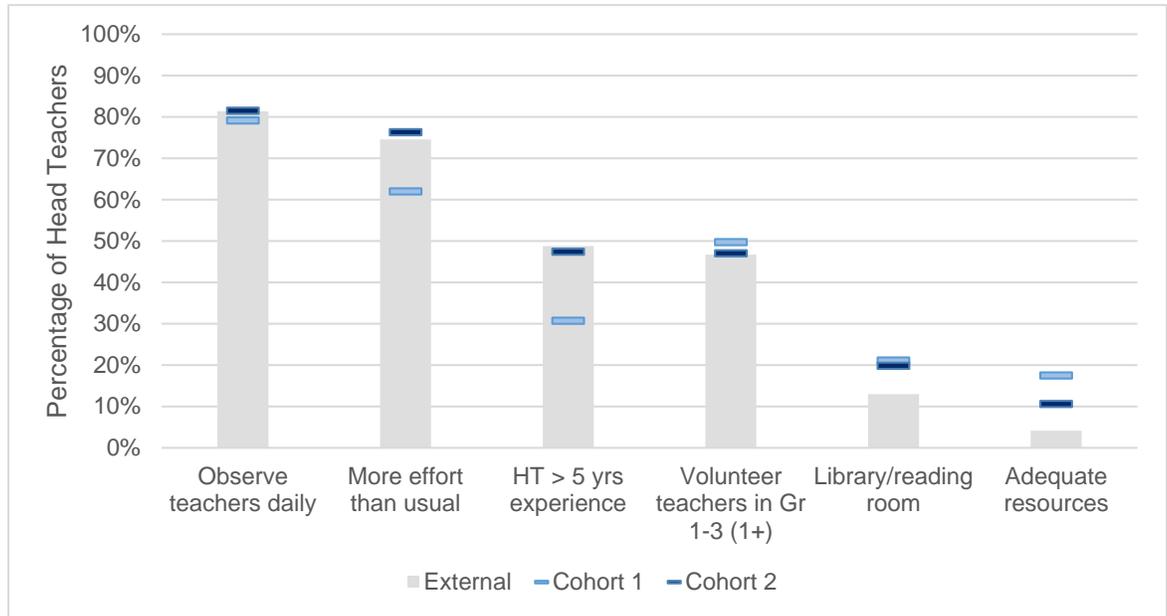
As can be seen in **Table 21**, the vast majority of head teachers reported attempting to improve reading in their schools. With reference to the midterm assessment, head teachers reported that their motivation for doing so was the training that teachers had received (51% in Cohort 1 and 41% in Cohort 2 schools), independent assessments (32% in Cohort 1 and 69% in Cohort 2 schools), and ministerial instructions (30% in Cohort 1 25% in Cohort 2). Unfailingly, the proportion of head teachers who reported that pupils, parents, and communities had shown a greater interest in reading had increased since the midterm assessment. Comparatively, only 60% of head teachers in External Cohort schools reported that parents or communities had expressed more interested in reading (not shown in **Table 21**). In addition, head teachers asserted that more discussions about reading were being held at the school level than in previous years. Further qualitative work must be conducted to determine the quality of this support provided to teachers and in regards to community engagement.

**Figure 14** presents selected characteristics of schools and head teachers at the time of the endline assessment.

As shown in **Figure 14**, most head teachers, including those of External Cohort schools, reported that they observed teachers daily (79%, in Cohort 1, 81% in Cohort 2, and 81% in External Cohort schools) and put forth more effort than usual during this academic year (62% in Cohort 1, 76% in Cohort 2, and 75% in External Cohort schools). Approximately half of the head teachers reported having more than five years of experience (except in Cohort 1, in which only 30% of head teachers reporting having this amount of experience) and that they used volunteer teachers in grades 1 through 3. In terms of school resources, few head teachers (20% or less) reported that the school had a library or reading room, and less than 20% reported that their school had adequate resources. Although the interview did not specify which

resources were lacking in head teachers' estimations, these data do suggest that schools are generally lacking resources.

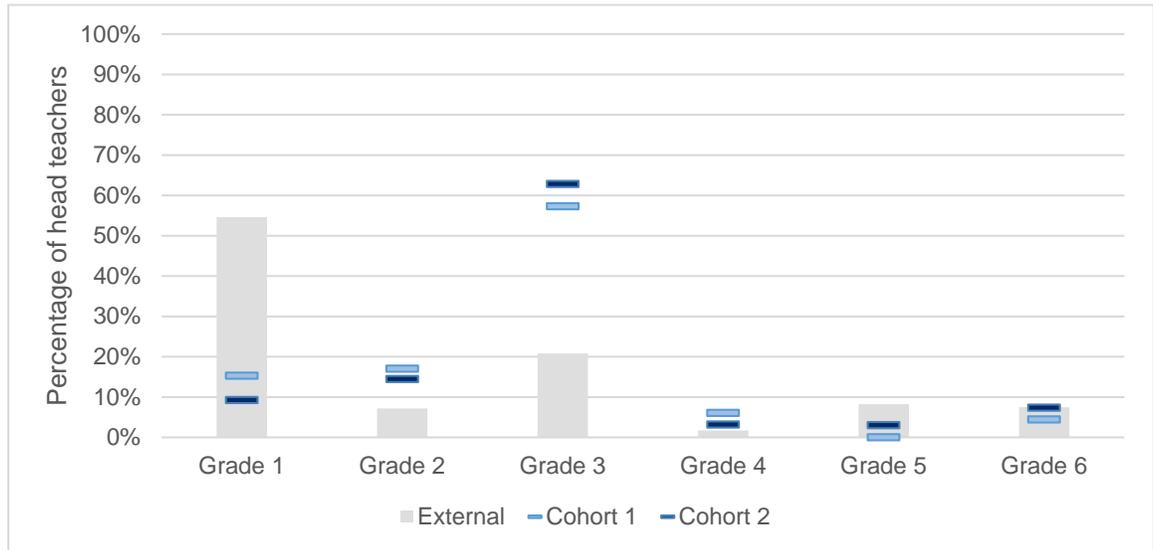
**Figure 14. Characteristics of head teachers and schools.**



It is worth noting that there is a significant discrepancy between the proportion of teachers who reported being observed daily by head teachers (64%) and the proportion of head teachers who reported observing teachers on a daily basis (79%) in Cohort 1 schools. Although this report cannot resolve this discrepancy, it does suggest that classroom observation may occur less frequently than reported by head teachers in Cohort 1 schools.

Head teachers in Cohort 1, Cohort 2, and External Cohort schools were asked at which grade they thought pupils should be able to read fluently (**Figure 15**). Overall, most head teachers in External Cohort schools (55%) asserted that pupils should be able to read fluently by the end of grade 1, and an additional 21% thought that a more developmentally appropriate target for fluent reading was by the end of grade 3. In program schools, most head teachers reported that pupils should be able to read by the end of grade 3 (57% in Cohort 1 and 63% in Cohort 2). As such, the program has appeared to change head teachers' perspectives regarding developmental reading and pupils' progression of skills.

**Figure 15. Head teachers' expectations regarding when all pupils should be able to read fluently.**



### 3.5.5 FACTORS ASSOCIATED WITH PUPIL PERFORMANCE

When combined with the results of the pupil assessments, the SSME questionnaire data provide an opportunity to explore the school-, class-, and pupil-level factors that are associated with learning outcomes. To this end, this section of the report presents the results of several statistical models that are designed to test these linkages. To test whether specific school characteristics, resources, pedagogical practices, and pupil home reading behaviors were associated with reading performance, a linear regression model with ORF was fitted onto the data. To determine a baseline model of pupil-level controls, several pupil-level variables over which schools have little or no influence—such as gender, grade, region, and home language—were entered into the regression analysis to test for statistical significance. Variables that were not found to be significantly associated with reading performance were discarded from the model. After several iterations, a baseline model was fitted with all pupil-level variables found to have a significant association with ORF. These pupil-level variables were: gender, grade, home language, eating breakfast on the day of the assessment, and pupil age (**Table 22**).<sup>6</sup>

<sup>6</sup> Note that region was not statistically significant and was therefore not included in the baseline model.

**Table 22. School Characteristics and Reading Performance (Endline)**

VARIABLE	SUBCATEGORY	BASE MODEL		ITERATION 1		ITERATION 2		ITERATION 3		ITERATION 4	
		COEFF	SE	COEFF	SE	COEFF	SE	COEFF	SE	COEFF	SE
Gender	Male (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Female	-2.12**	0.84	-2.24*	0.88	-1.98*	0.94	-2.16*	0.942	-1.76	0.94
Grade	1 (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2	9.28**	1.40	9.10**	1.43	9.46**	1.57	9.31**	1.55	9.41**	1.51
	3	18.59**	2.03	18.17**	2.07	18.5**	2.25	18.02**	2.17	18.44**	2.21
Home language	English (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Other	4.56**	1.12	3.75**	1.17	4.66**	1.33	4.57**	1.27	4.39**	1.23
Breakfast today	No (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Yes	3.24**	0.95	2.30*	1.02	2.34*	1.11	2.33*	1.07	1.97	1.08
Age		-0.95**	0.26	-0.85**	0.30	-1.03**	0.30	-0.87**	0.28	-1.01**	0.29
Head teacher education	Senior high school (reference)	—	—	0.00	0.00	—	—	—	—	—	—
	Associate	—	—	7.19	4.62	—	—	—	—	—	—
	C Certificate	—	—	-1.03	1.77	—	—	—	—	—	—
	B Certificate	—	—	2.03	2.16	—	—	—	—	—	—
	AA Certificate	—	—	0.47	2.38	—	—	—	—	—	—
	Bachelor's	—	—	10.55**	2.60	—	—	—	—	—	—
	Master's	—	—	-0.51	8.75	—	—	—	—	—	—
% volunteer teachers		—	—	—	—	-0.01	0.02	—	—	—	—
Head teacher experience		—	—	—	—	—	—	-0.27	0.14	—	—
Observation frequency	Biweekly (reference)	—	—	—	—	—	—	—	—	0.00	0.00

**Table 22. School Characteristics and Reading Performance (Endline)**

Weekly	—	—	—	—	—	—	—	2.01	1.82	
Daily	—	—	—	—	—	—	—	6.37**	1.40	
Constant	6.92**	1.61	6.06*	2.38	8.11**	1.92	8.80**	2.13	2.49	2.05
R-squared	0.15		0.18		0.15		0.15		0.15	

COEFF = model coefficient; SE = standard error

\* $p < 0.05$ , \*\* $p < 0.001$

## School-Level Factors Associated with Reading Performance

After fitting the baseline model, we tested whether certain school characteristics were associated with pupil reading performance. Here, we focused on the educational background of head teachers, the proportion of volunteer teachers in grades 1–3 at the school, head teachers' years of experience, and the frequency with which head teachers observed classroom teaching (as reported by head teachers). We entered each of these variables into the model one at a time to determine statistical significance and whether the additional variable added value to the model. **Table 22** presents the results of these series of tests in the form of model coefficients (COEFF) and standard errors (SE). Because the outcome variable is ORF, the coefficients can be taken to signify the difference in the number of words read correctly per minute compared to the variable reference category. For example, in the base model, female pupils appear to read approximately two words per minute slower than boy pupils, after accounting for grade, home language, eating breakfast, and age.

Turning to the school characteristic variables, only head teachers' academic backgrounds and daily classroom observations were found to have statistically significant associations with pupils' reading fluency at the endline. Regarding the head teachers academic backgrounds, pupils in schools where head teachers were holders of a bachelor's degree tended to perform better (reading more than 10 cwpm faster) than pupils in schools where head teachers only had a secondary diploma. Although only a few head teachers in Cohort 1 (3%) and in Cohort 2 (7%) schools held bachelor's degrees, having a higher academic qualification may be correlated with more extensive training, or it may be indicative of other school characteristics conducive to pupil learning. It is unlikely, in any event, that the head teachers' degree itself confers such increases in learning gains compared with schools in which head teachers had finished secondary schooling. With regard to the latter, **Table 22** shows that pupils in schools where head teachers observed classroom instruction on a daily basis (or at least reported doing so) tended to outperform (read 6.4 cwpm faster than) pupils in schools where observations occurred on a biweekly basis. Because these self-reported data mostly aligned with teachers' reports (with the only discrepancy found in Cohort 1 schools), this sort and frequency of instructional support to teachers appears to matter vis-à-vis pupil reading performance in Cohort 1 and Cohort 2 schools.

We tested three more models to determine whether particular school resources tended to be associated with enhanced reading performance at the endline. Specifically, we tested whether sufficient resources (as reported by the head teacher), the presence of a library or reading room, and the pupils' ability to borrow books from schools held a statistically significant linkage with pupil ORF. Results are presented in **Table 23**. The table shows, however, that the only school-level resource that exhibited

a relationship with reading fluency at the endline assessment was the availability of books that pupils could borrow from schools. Recall that this practice was much more common among program Cohort 1 and Cohort 2 schools as compared with the External Cohort schools (**Figure 7**). Pupils in these schools tended to read slightly (approximately 2 cwpm) faster than pupils in schools where books were not available or where borrowing was not allowed.

**Table 23. School Resources and Reading Performance (Endline)**

VARIABLE	SUB-CATEGORY	BASE		ITERATION 1		ITERATION 2		ITERATION 3	
		COEFF	SE	COEFF	SE	COEFF	SE	COEFF	SE
Gender	Male (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Female	-2.12**	0.84	-2.04*	0.94	-2.11*	0.90	-2.12**	0.84
Grade	1 (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2	9.28**	1.40	9.58**	1.50	9.35**	1.49	9.31**	1.39
	3	18.59**	2.03	18.54**	2.19	18.29**	2.17	18.55**	2.02
Home language	English (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Other	4.56**	1.12	4.65**	1.24	4.48**	1.22	4.55**	1.12
Breakfast today	No (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Yes	3.24**	0.95	2.35*	1.05*	2.17*	1.05	2.93**	0.95
Age		-0.95**	0.26	-0.97**	0.29	-0.90**	0.30	-0.95**	0.26
Sufficient resources	No (reference)	—	—	0.00	0.00	—	—	—	—
	Yes	—	—	-1.99	1.78	—	—	—	—
Library or reading room	No (reference)	—	—	—	—	0.00	0.00	—	—
	Yes	—	—	—	—	3.73	1.65	—	—
Pupils can borrow books	No (reference)	—	—	—	—	—	—	0.00	0.00
	Yes	—	—	—	—	—	—	1.97*	1.01
Constant		6.92**	1.61	7.80**	1.82	6.16**	1.96	5.62**	1.75
R-squared		0.15		0.15		0.15		0.15	

COEFF = model coefficient; SE = standard error

\* $p < 0.05$ , \*\* $p < 0.001$

### Classroom-Level Factors Associated with Reading Performance

Although data from teacher interviews were not collected in such a way as to associate them with pupil outcomes, the pupils were also asked about their teachers' pedagogical habits. As such, we were able to include this reported instructional behavior in models designed to test whether specific pedagogical activities in the classroom were associated with pupil reading outcomes. During interviews, the pupils reported whether their teachers instructed them to read aloud in class, practice letter sounds, or read silently, or whether the teacher read aloud to pupils. As shown in **Table 24**, the only classroom practice associated with pupil ORF at the endline assessment was practicing letter sounds. Specifically, pupils in classrooms where teachers had them practice letter sounds tended to read more than 5 cwpm faster than their peers in classrooms that did not engage in such practice. Although this finding is not meant to suggest that other practices, such as reading aloud or silently, are unimportant, it does perhaps indicate that pupils benefit (in terms of reading ability) from teachers who lead them in more explicit phonetic practice.

**Table 24. Pupil-Reported Teacher Practices and Reading Performance (Endline)**

VARIABLE	SUB-CATEGORY	BASE		ITERATION 1		ITERATION 2		ITERATION 3		ITERATION 4	
		COEFF	SE	COEFF	SE	COEFF	SE	COEFF	SE	COEFF	SE
Gender	Male (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Female	-2.12**	0.84	-2.18**	0.84	-2.00*	0.83	-2.16**	0.84	-2.10**	0.84
Grade	1 (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2	9.28**	1.40	9.30**	1.40	9.00**	1.39	9.19**	1.39	9.26**	1.40
	3	18.59**	2.03	18.61**	2.03	18.24**	2.01	18.50**	2.02	18.59**	2.02
Home language	English (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Other	4.56**	1.12	4.52**	1.12	4.58**	1.10	4.44**	1.12	4.47**	1.12
Breakfast today	No (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Yes	3.24**	0.95	2.97**	0.98	2.78**	0.95	2.93**	0.98	2.97**	0.96
Age		-0.95**	0.26	-0.96**	0.26	-0.93**	0.26	-0.94**	0.26	-0.93**	0.26
Pupils read aloud	No (reference)	—	—	0.00	0.00	—	—	—	—	—	—
	Yes	—	—	2.34	1.24	—	—	—	—	—	—
Practices letter sounds	No (reference)	—	—	—	—	0.00	0.00	—	—	—	—
	Yes	—	—	—	—	5.27**	1.03	—	—	—	—
Silent reading	No (reference)	—	—	—	—	—	—	0.00	0.00	—	—
	Yes	—	—	—	—	—	—	1.92	1.03	—	—
Teacher reads aloud	No (reference)	—	—	—	—	—	—	—	—	0.00	0.00
	Yes	—	—	—	—	—	—	—	—	3.57	0.89
Constant		6.92**	1.61	5.05**	1.82	2.55	1.69	5.52**	1.79	3.73*	1.72
R-squared		0.15		0.15		0.16		0.15		0.15	

COEFF = model coefficient; SE = standard error

\* $p < 0.05$ , \*\* $p < 0.001$

### Pupil-Level Factors Associated with Reading Performance

Lastly, we tested several pupil-level variables that were thought to impact reading ability. During interviews, pupils were asked questions about the availability of books and their reading practices at home (whether others read to them and whether they read aloud to others). We added these variables one at a time to the base model. The results are presented in **Table 25**.

Notably (although perhaps unsurprisingly), all pupil-level variables tested in the models showed a statistically significant association with pupil reading ability at the endline assessment. Pupils who reported that they had reading books at home tended to perform better (3.5 cwpm) on the oral reading subtask than pupils who did not. Of course, the mere presence of books in a home does not beget a person's reading ability, but it is a proxy for the accessibility of resources, a variety of reading material, and the opportunity to practice—all of which are important for reading development. Being read to by others was also positively associated with reading performance: pupils who reported being read to in their homes read, on average, approximately 4 cwpm faster than pupils who were not read to at home. The strongest association, however, was with pupils who practiced reading aloud to others in their homes. These pupils tended to read approximately 6.8 cwpm faster than pupils who did not practice reading at home.

Although the finding that pupils who practice reading outside of school tend to read faster and more accurately than pupils who do not borders on the banal, it is nevertheless important. Recall that only half of interviewed pupils reported practicing reading to others at home and that this proportion had decreased from the midterm assessment. Encouraging pupils to read with someone outside of school and providing them with resources to do so could be a school-level priority. In addition, allocating time within the school day for pupils to practice reading aloud could also help to overcome the lack of access to resources or opportunities to practice for some children.

**Table 25. Home Reading Practices and Reading Performance (Endline)**

VARIABLE	SUB-CATEGORY	BASE		ITERATION 1		ITERATION 2		ITERATION 3	
		COEFF	SE	COEFF	SE	COEFF	SE	COEFF	SE
Gender	Male (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Female	-2.12**	0.84	-2.09**	0.83	-1.98*	0.82	-1.96*	0.82
Grade	1 (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	2	9.28**	1.40	8.89**	1.38	9.00**	1.36	8.31**	1.36
	3	18.59**	2.03	18.25**	2.02	18.45**	2.01	17.58**	1.95

**Table 25. Home Reading Practices and Reading Performance (Endline)**

VARIABLE	SUB-CATEGORY	BASE		ITERATION 1		ITERATION 2		ITERATION 3	
		COEFF	SE	COEFF	SE	COEFF	SE	COEFF	SE
Home language	English (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Other	4.56**	1.12	4.34**	1.11	4.03**	1.10	3.51**	1.10
Breakfast today	No (reference)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Yes	3.24**	0.95	3.10**	0.95	2.69**	0.92	2.84**	0.91
Age		-0.95**	0.26	-0.95**	0.25	-0.92**	0.25	-0.91**	0.25
Books at Home	No (reference)	—	—	0.00	0.00	—	—	—	—
	Yes	—	—	3.52**	0.87	—	—	—	—
Others read to pupil	No (reference)	—	—	—	—	0.00	0.00	—	—
	Yes	—	—	—	—	4.13**	0.84	—	—
Pupil reads to others	No (reference)	—	—	—	—	—	—	0.00	0.00
	Yes	—	—	—	—	—	—	6.82**	0.95
Constant		6.92**	1.61	4.93**	1.65	5.30**	1.67	4.38**	1.62
R-squared		0.15		0.16		0.16		0.18	

COEFF = model coefficient; SE = standard error  
 \* $p < 0.05$ , \*\* $p < 0.001$

### Limitations

Although the results presented in this section of the report describe many school-, classroom-, and pupil-level factors associated with pupil reading performance during the endline assessment, it is important to note several limitations to this analysis.

First and foremost, the association between scores registered and variables observed only at the endline ignores similar data obtained at the midterm. In particular, the largest predictor of pupil performance—prior pupil performance—is not included in the models. In other words, the analysis attempts to explain the variation observed regarding pupil reading fluency at one point in time, but it does not account for variation observed at prior points in time. In addition, although the results described here are relevant for school leaders and policy makers, they do not explain the growth observed in pupil reading or mathematics scores observed in different cohorts of pupils and schools from midterm to endline. Growth, rather than raw achievement levels, is what the intervention attempted to achieve.

Second, data from the teacher interview could not be linked empirically with the pupil-level data. Because of the nature of the data collection, the teachers who were interviewed were not necessarily the teachers of pupils who were assessed at endline. Including the teacher data in the models here could have been accomplished at the school level (by aggregating teacher variables and creating variables of “school-

level” human resources). However, doing so would have tended to obscure the analysis because the interviewed teachers did not necessarily have direct contact with pupils.

Third, the same pupils were not assessed at midterm and endline; thus, any growth that is observed by the endline assessment could only be aggregated at the school level. Assessing the same pupils at different points in time is difficult logistically and costly for many reasons; therefore, this approach was not taken in Liberia. Despite these limitations, the findings presented here provide insight regarding many factors that are associated with pupil reading performance. To the extent that enhanced performance is a priority in Liberia, policy makers and school leaders would do well to consider these factors as important elements in a holistic approach toward improving reading ability.

### 3.5.6 SUMMARY OF SSME FINDINGS

This entire section of the report has presented data from interviews with head teachers, teachers, and pupils to supplement the results from the EGRA and EGMA assessments previously described. The main purpose of these data is to describe schools, classrooms, and pupils in the intervention regions and to provide a snapshot of pupils within their learning environments. To this end, the section also discussed findings from head teacher, teacher, and pupil interviews from the midterm and endline assessments to describe how those learning environments have changed over time.

With regard to pupils, several background characteristics appear to have changed markedly between the midterm and endline assessments, suggesting that the pupil populations in these schools have changed. The proportion of pupils who spoke English at home and whose parents speak English increased substantially. However, the percentages of pupils who had books at home, were read to by others, or practiced reading at home fell in Cohort 1 and Cohort 2 schools since the midterm, which is concerning (although children in both cohorts still out-read their peers in External Cohort schools). Lastly, two positive school-level changes were reported by children in Cohort 2 schools: significantly more pupils reported taking books home from school and practicing letter sounds with their teachers at endline as compared to midterm. These trends may alleviate the gap in reading resources and teacher actions observed during the midterm assessment of these variables between Cohort 1 and 2 schools.

From the teacher interviews, we noted changes in supervision and instructional support in Cohort 1 and Cohort 2 schools. Overall, it appears that the Cohort 1 teachers received less support at endline compared to midterm (in the form of support

visits, observations, and external inspections), whereas Cohort 2 teachers received substantially more support compared to midterm. When asked whether, in the past five days, the teachers frequently had pupils engage in specific reading activities, most of the teachers reported that they did not frequently engage their pupils during these literacy activities. The frequency of these activities appeared to be declining in Cohort 1 schools, but was increasing slightly in Cohort 2 schools. The frequency of various forms of mathematics practice, too, appeared to be on the rise in Cohort 2 schools. These findings were expected, given the staged nature of the intervention. Still, it is notable that the longevity of the practices undertaken during the intervention does not appear to last long after the intervention has finished; Cohort 1 schools and teachers had already begun to revert to previous practices.

The vast majority of head teachers reported that they are attempting to improve reading in their schools. The teachers reported that their motivation for doing so was because of the training they had received, independent assessments, and ministerial instructions. Unfailingly, the proportion of head teachers who reported that pupils, parents, and communities had shown a greater interest in reading had increased since the midterm assessment. In addition, head teachers asserted that more discussions about reading were being held at the school level as compared to previous years.

The regression analyses revealed several school-, classroom-, and pupil-level factors associated with pupil reading performance as measured by the oral reading subtask for the endline EGRA. At the school level, the academic backgrounds of head teachers and observing classroom instruction on a daily basis were positively associated with reading outcomes. At the classroom level (based on reports from the pupils regarding the behavior of their teachers), practicing letter sounds and pupils' ability to borrow books from school to read at home (two phenomena that were more common in Cohort 1 and Cohort 2 schools as compared with External Cohort schools) were both positively associated with ORF. Perhaps the most consistent, if unsurprising, finding was observed at the pupil level: access to books, being read to, and reading to others were all significantly associated with reading fluency. Such findings are perhaps predictable; however, not all schools, teachers, and pupils were privy to the resources or engaged in the activities that are predictably associated with pupil reading outcomes.

## 4 CHALLENGES

The LTTP has made an important step toward improving the quality of instruction and support provided to 1,073 schools (792 Cohort 1 schools and 326 Cohort 2 schools). The program design, nevertheless, overestimated the MOE's capacity—from the central to the school level—to take on a systemic change. This change would require significant financial and human resources, content-based planning, and management improvements. In addition to the Ebola crisis, the program also ran into a series of operational and technical challenges that caused serious delays. This section of the report discusses the challenges that the program had to deal with and recommendations for future programming.

### LONG-TERM ROLLOUT AND SUSTAINABILITY

“Reading and math programs are making progress: The team is impressed by the work by the EGRA and EGMA consultants on the LTTP II. The assessment team recommends reinforcing the training for the teachers already trained and a go-slow approach to rolling-out of the second component. While this reinforcement is taking place, the LTTP needs to work from within the Center of Curriculum Development to develop a longer term roll-out strategy for training teachers in reading and to institutionalize the reading and math program with the Center of Curriculum Development.”

—*External evaluation report*

#### 4.1 IMPACT OF EBOLA

The Ebola emergency and the subsequent closing of all primary schools in Liberia for seven months during the 2014/2015 academic year had a significant impact on the performance of Cohort 2 schools (Save the Children, 2015). Most of the activities that were planned for implementation just before and/or at the beginning of the new academic year had to be canceled. One activity that was impacted included the distribution of a large consignment of new reading and mathematics books to all Cohort 1 and Cohort 2 schools. Another activity that was impacted was the 10-day refresher training in reading and mathematics for relevant LTTP staff (Reading and Mathematics Specialists and Capacity-Building Officers) and MOE representatives and multiple simultaneous face-to-face trainings in LTTP implementation counties at the cluster level for District Education Officers and all Cohort 2 school teachers and principals.

The closing of schools exacerbated the challenges teachers faced in delivering the lesson plans by the end of the 2014/2015 academic year (in June 2015). As shown in **Table 26**, the majority of Cohort 2 schools had reached, on average, only Week 6 or Week 7 in the reading teaching and learning materials by the end of the school year, meaning that the teaching of some critical skills was not achieved.

**Table 26. Cohort 2 Schools' Progress on Lesson Plans at End of 2014/2015 Academic Year**

COUNTY	NUMBER OF COHORT 2 SCHOOLS	LESSON STATUS (AVERAGE)
Bong	70	Week 7–10
Lofa	75	Week 5–6
Margibi	37	Week 6–8
Montserrado	38	Week 7–8
Nimba	106	Week 8–10
<b>Total/Average</b>	<b>326</b>	<b>Week 6–8</b>

## 4.2 SYSTEMIC CHALLENGES

**Severe shortage of skills.** The MOE suffers from very severe capacity constraints, which is a cause of great concern and frustration to many members of the MOE. A recent MOE human resources exercise, undertaken in conjunction with the Civil Service Authority, resulted in many vacancies being declared, but the Ebola situation has meant that the posts have not yet been filled. The gap in capacity had serious implications for LTTP work with the MOE. Overall, the few people who have the skills and knowledge to handle the activities already have too much work or are chasing other work; therefore, this leads to a shortage of time to reflect on policy, develop it, and implement it appropriately. Teachers, principals, and other education staff need continued assistance because the system itself will be unable to carry on with the activities now that LTTP support has ended.

**Teacher transfers.** Transfers implemented by MOE's District Education Officers of LTTP-trained teachers and principals to nonprogram schools resulted in instructional gaps in Cohort 1 and Cohort 2 schools; that is, the pupils in these schools were left without effective reading and mathematics instruction. These transfers took place, despite a Memorandum of Agreement between LTTP II and target schools to *not* transfer teachers, resource teachers, or principals to non-LTTP II reading and mathematics schools. Frequently, because of these transfers, the schools were left

with fewer teachers because the transferred teachers were not replaced. This practice of teacher transfers forced the remaining teachers to combine classes that were not on the same level, making the teaching much harder.

**School-based support.** The lack of periodic monitoring and support visits by District Education Officers to schools made it difficult to ensure that the training and the instruction would continue in schools. The lack of periodic monitoring and support visits was a result of human resource and financial constraints with which the MOE had been grappling. In conditions such as that of Liberia, where public financing of education is among the lowest in Africa, it is inconceivable to expect sustainability over a short period of time. LTTP introduced the position of the coach, an individual who supports schools in person on a regular basis. Despite the MOE's promises to continue this role, it is still deemed unsustainable, given that the position does not exist in the MOE career ladder. Yet experience showed, and the teachers confirmed, that the coach was the key ingredient to implementing the training they received.

## MOE CAPACITY

“All parties over-estimated the capacity of the MOE: The Liberian MOE is characterized by frequent changes in leadership, weak educational planning and management, and lack of financial and human resources throughout. The Ministry is striving to redress the debilitating effects of past conflicts, while laying the foundation for a nationwide education delivery system. Consequently, it is not in a position to say ‘no’ to offers of development assistance by any donor, even when it lacks the framework and capacity to carry out the duties expected of a true development partner. The LTTP is based on the principles of joint engagement, but currently the MOE lacks the capacity to work as a full partner with the LTTP. The LTTP needs to adjust its operating strategy to account for the poor capacity within the MOE.”

—*External evaluation report*

**Time on task.** Apart from the many interruptions to the academic year because of frequent exam periods (every six weeks, children are tested, which takes one week to prepare for exams and another week to administer them), teachers' time on task was affected by holidays, market days, teachers' second jobs, times when teachers left the classrooms to collect their salaries from larger towns, and poor time management within the instructional time. Added to these interruptions was the challenge of teachers' abandoning classes and/or staging “go-slow” actions in demand of pay, and many teachers avoiding coming to work because of the ongoing MOE payroll verification process. Another perennial challenge faced was the absence of children from formal schools because of their participation in Sande and Poro (“bush”) schools. It is unlikely that these systemic challenges will be fully resolved until the root causes are addressed (e.g., low teacher salaries, accountability measures).

**Logistics.** The book distribution to schools was delayed because of procurement and distribution challenges, which subsequently had an impact on pupils' reading scores. The procurement in the first year of the program was delayed because the reading and mathematics books needed to be developed. In later years, the availability of books was delayed by a lengthy procurement processes, heavy rains, the Ebola crisis, and a lack of the logistical capability to swiftly distribute books once they arrived in Liberia.

## 5 RECOMMENDATIONS

**MOE capacity building.** The MOE's capacity to improve the quality of education in Liberia has been characterized as underfunded and inadequate. The LTTP II external evaluators suggested that future programming should focus on improving some of the systemic challenges that the MOE is facing first, work in fewer schools, and introduce a whole-school approach, which would then make the MOE more capable of receiving and sustaining the support in the long run. Unless these systematic challenges are addressed, it is not recommended to take LTTP II methodology to scale.

**Over-age pupils.** More research must be conducted to identify the reasons why most pupils in the early grades of schooling are over-age. Once these reasons are known, then steps can be undertaken to address this issue.

**Teacher transfers.** The MOE and District Education Officers need to refrain from using teacher transfers as a punitive tool for poor performance. Instead, schools that are underperforming should be identified and supported.

**School-based support.** The schools (principals, teachers, and PTAs) continue to require significant support. It is recommended that the MOE and donors focus on rapid support to improve the system's capacity immediately while, perhaps, not losing the sight of service delivery, given the shortage of skills and financing.

## OPTIMIZING PROGRAM CAPACITY

“Although the early grade reading and mathematics program appears to have made a strong start and is widely heralded by teachers and administrators, the LTTP staff implementing the program [and] most teachers and school administrators suggest that [it] is spread too thinly. Nearly everyone interviewed suggested that the training of early grade (1–3) teachers and principals was good and that the materials provided were terrific. They [said] that more training was needed for the teachers, principals, and coaches and that more reading and mathematics materials should be provided. Many [said] that a ‘whole school’ approach should be adopted for training teachers and providing teaching and learning, given [teacher] turnover and the presence of [many] over-age children [without] reading and math skills in all primary grades. The lack of MOE counterpart funds and MOE personnel who are [not entirely] committed to the program, along with the other factors cited above, suggest that the program may not be sustainable in the medium-term. USAID/Liberia and the LTTP may want to re-examine the programs’ output targets and use the savings to reinforce the gains already made.”

—*External evaluation report*

**Whole-school approach.** Even though it was not a direct focus of LTTP II, some coaches succeeded in reviving and strengthening the PTA. Without strong school leadership and engagement of parents, however, the continuity of any such program in schools will be in question. Apart from the accountability side of community and parental engagement, future programs should focus on ensuring that children are safe in and around the school, and as they travel to and from school.

**Time on task.** The MOE needs to enact measures to improve time on task. These measures may include the following: ensuring that teachers show up in classrooms, that they teach for the full class period, that the teaching aligns with the lesson plan, and that schools remain open all five days each week. Other efforts may include piloting mobile banking for paying teacher salaries and electronic data collection.

**Pupil performance tracking.** More attention needs to be given to tracking pupil performance and using data for decision making and accountability. In all four counties (i.e., Bong, Lofa, Montserrado, and Nimba), the only systematic assessment of reading and mathematics during implementation of this program was that which was carried out by LTTP II, through the EGRA, EGMA, and periodic mini-EGRAs (i.e., smaller scale assessments).

**Logistics.** Teacher’s manuals and books for pupils must be delivered on time. Given that the LTTP and the MOE developed the reading and mathematics curricula for grades 1 through 3, future programs should continue to use these books. Moreover, the capacity of the MOE to procure and distribute books must be developed.

**Focus on fluency and comprehension.** Improving the ORF and decoding skills of pupils is possible in a short period of time, and such improvements are highly correlated with reading comprehension. We know that children can understand a higher percentage of what they hear than what they read. Explicit instruction and modeling are necessary to match pupils' listening comprehension with their comprehension after they read written texts.

**Use reading benchmarks for decision making.** The wealth of data obtained since EGRA was first introduced in Liberia in 2008 provided enough evidence for the MOE, LTTP, and other nongovernmental organization and donor stakeholders to determine what rates of fluency, comprehension, and word skills are necessary at each level. These benchmarks should be used in financial and educational planning and prioritizing.

**Continue with the focus.** Liberian teachers have proven to be receptive to new pedagogical techniques and strategies. The RTTIs responded well to a renewed focus on the reading and mathematics curriculum. It is with targeted efforts that teachers can improve how well children read, quite quickly. We recommend that the focus on reading and mathematics in early grades continues. MOE and USAID invested significant resources to develop reading and mathematics books; therefore, we are recommending that the MOE assumes full ownership of these materials.

**Scripted lessons.** Liberian teachers lack lesson planning skills and content knowledge. The reading and mathematics curriculum for grades 1, 2, and 3 that was developed in collaboration with the MOE provided teachers with scripted lessons and with abbreviated outlines of a lesson for those teachers who were better equipped and could plan on their own. Many teachers improved their content knowledge by implementing these lessons. We recommend that the MOE continue to use this curriculum. We also recommend that various PD institutions use the LTTP–developed training-of-trainers manuals.

**Pilot instruction in mother tongue.** We recommend that the MOE explore introducing the use of mother tongue as a language of instruction in early grades. Research indicates that if children learn to read in the language they learn first orally, then these skills will transfer to English or another second language much faster.

**Decoding skills and kindergarten.** These skills are not taught in schools to the extent required, and yet they are crucial gateways to the rapid acceleration of learning outcomes and improved reading performance. Some performance improvements under LTTP II were noticeable, but not at the level required for pupils as a matter of course to improve their fluency in reading and comprehension. The emphasis on decoding skills—and reading overall—should be introduced in kindergarten so that rising grade 1 pupils are better prepared.

**Reading at home.** Pupils who reported that they had reading books at home tended to perform better (3.5 cwpm) on the oral reading subtask than pupils who did not. Regardless of parental literacy, pupils should be required to read at home out loud to their parents or siblings.

**Mathematics.** In mathematics, LTTP II generated ample evidence regarding the importance of key competencies that can be considered as building blocks for future, more complex tasks that will be introduced in later grades. It is crucial that children develop deep and flexible mathematics knowledge at an early age. Research has shown that early mathematics outcomes are predictive of children's later reading and mathematics outcomes. The attention paid to mathematics must equal that of reading, and it cannot be put aside. Pupils in the early grades need a consistent mathematics program that will teach them the basic competencies. We recommend that any future program include a comprehensive mathematics program that will allow pupils to develop this deep and flexible mathematics knowledge.

# REFERENCES

- Aron, A., Aron, E. N., & Coups, E. (2013). *Statistics for psychology (6th ed.)*. Boston, Massachusetts, USA: Pearson.
- Cooper-Trent, Jennifer. (2015). Teach the world to read: Free reading program for developing communities, charities, and countries; and Fantastic Phonics program [Websites]. Retrieved from <http://www.teachtheworldtoread.com> and <http://www.early-reading.com>
- Crouch, L., & Korda, M. (2008). *EGRA Liberia: baseline assessment of reading levels and associated factors*. Report prepared as part of a process of collaboration between USAID and the World Bank. Research Triangle Park, North Carolina: RTI International. Retrieved from <https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=158>
- Davidson, M., Korda, M., and Collins, Ollie White. (2011). Teachers' use of EGRA for continuous assessment: The case of EGRA Plus: Liberia. In A. Gove & A. Wetterberg (Eds.), *The Early Grade Reading Assessment: Applications and interventions to improve basic literacy* (pp. 113–137). Research Triangle Park, North Carolina: RTI Press. Retrieved from <http://www.rti.org/pubs/bk-0007-1109-wetterberg.pdf>
- Heneveld, W., & Craig, H. (1996). *Schools count: World Bank project designs and the quality of African primary education*. Washington, DC: World Bank. Retrieved from <https://www.openknowledge.worldbank.org/handle/10986/9974>
- Institute of Education Sciences, U.S. Department of Education. (2014). *What Works Clearinghouse: Procedures and standards handbook (version 3.0)*. Retrieved from [http://ies.ed.gov/ncee/wwc/pdf/reference\\_resources/wwc\\_procedures\\_v3\\_0\\_standards\\_handbook.pdf](http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf)
- International Monetary Fund (IMF). (2008). *Poverty reduction strategy: Republic of Liberia*. Washington, DC: IMF. Retrieved from <http://www.imf.org/external/pubs/ft/scr/2008/cr08219.pdf>

- Piper, B., & Korda, M. (2011). *[Early Grade Reading Assessment] EGRA Plus: Liberia—Program evaluation report*. Final report prepared for USAID under the Education Data for Decision Making (EdData II) project, Task Order No. EHC-E-06-04-00004-00 (RTI Task 6). Research Triangle Park, North Carolina: RTI International. Retrieved from [http://pdf.usaid.gov/pdf\\_docs/pdacr618.pdf](http://pdf.usaid.gov/pdf_docs/pdacr618.pdf)
- RTI International. (n.d.). *Early grade reading and mathematics methodology: From pilot to scale*. Prepared for FHI 360 and USAID/Liberia under the Liberia Teacher Training Program (LTTP), Phase II of the EQUIP2 Leaders Award under contract with FHI 360, Inc. Research Triangle Park, North Carolina: RTI International.
- RTI International. (2011a). *EGRA Plus: Liberia—Overview of a sustainable reading intervention*. Report prepared for USAID under the Education Data for Decision Making (EdData II) project, Task Order No. EHC-E-06-04-00004-00 (RTI Task 6). Research Triangle Park, North Carolina: RTI International. Retrieved from [http://pdf.usaid.gov/pdf\\_docs/pnaea383.pdf](http://pdf.usaid.gov/pdf_docs/pnaea383.pdf)
- RTI International. (2011b). *Baseline assessment report*. Prepared for FHI 360 and USAID/Liberia under the Liberia Teacher Training Program (LTTP), Phase II of the EQUIP2 Leaders Award under contract with FHI 360. Research Triangle Park, North Carolina: RTI International.
- Save the Children. (2015). *Liberia Education Cluster: Assessment of the effect of Ebola on education in Liberia*. February. Retrieved from [http://educationcluster.net/?get=002241|2015/02/Liberia\\_Education\\_Cluster\\_Ebola\\_Assessment\\_Report\\_FINAL.pdf](http://educationcluster.net/?get=002241|2015/02/Liberia_Education_Cluster_Ebola_Assessment_Report_FINAL.pdf)
- United Nations Development Programme (UNDP). 2013. *Human development reports: Human Development Index (HDI) [Data files]*. Retrieved from <http://hdr.undp.org/en/content/human-development-index-hdi>

This page intentionally left blank.

# ANNEX A: TECHNICAL DETAILS ABOUT SAMPLE DESIGN AND WEIGHTING PROCEDURES

## SAMPLING DESIGN FOR THE ENDLINE ASSESSMENT

School sampling was probability proportional to size (PPS) sampling with the three cohorts as clusters.

Two-stage sampling was used as follows:

- Stage 1—PPS sampling was used to select schools for each cohort. The probability of a school being selected was the number of grade pupils in the school divided by the number of grade pupils in the cohort.
- Stage 2—Systematic sampling was used to select pupils. The pupils for each grade were separately lined up, and then every  $n^{\text{th}}$  pupil (as appropriate for the number of pupils in the school that day) was selected.

## WEIGHTS

Weights for the two stages were calculated by using the inverse of probability of selection.

Stage 1 weight:

$$\frac{\text{Number of pupils in the selected cohort}}{\text{Number of pupils in school} \times \text{number of schools selected in the cohort}}$$

The weights for Stage 2 were obtained from pupil counts by gender and school.

Stage 2 weight:

$$\frac{\text{Number of pupils in the school by grade}}{\text{Number of pupils selected in the school for the sample by grade}}$$

The final weight by pupil level was the product of the weights for Stages 1 and 2.

## REPLACEMENTS

An additional list of replacement schools was generated—a representation of approximately 20% of the original sample. If a school was inaccessible or unavailable, then a randomized school from the same cohort was used as a replacement. It is important to note that logistical problems during the rainy season caused many pupils to miss school; therefore, a number of schools had to be replaced because no pupils were present in those schools on the date of the assessment.

# ANNEX B: TECHNICAL DETAILS ABOUT SUBTASK EQUATING

Equating is a statistical procedure used to convert scores from multiple forms of a test to the same common measurement scale. This conversion process adjusts for any differences in difficulty that exist between forms so that a score on one form can be equated to its equivalent value on another form. As a result, equating makes it possible to estimate the score that a person taking one test form would have received if he or she had taken a different test form (Holland & Dorans, 2006).

Similar to any procedure based on statistical estimates, equating is typically performed by using samples from a population of individuals, and the results are susceptible to some combination of random sampling error and bias. As sample sizes increase and become more representative of the population, standard error is expected to decrease. Conversely, as sample sizes decrease and potentially become less representative of the population, statistical assumptions are introduced into the equating process in hopes of reducing the amount of estimation and thereby limiting the expected increase in standard error.

The endline assessment used the same pupil Early Grade Reading Assessment (EGRA) instruments as were used during the baseline assessment in 2011. As a consequence of this, no equating was required for the endline pupil scores.

It is important to note that the midterm scores were equated because a different instrument was used for that assessment. **Table B-1** presents the mean scores.

**Table B-1. Mean scores for reading fluency and comprehension, baseline and midterm**

SUBTASKS	BASELINE PASSAGE	MIDTERM PASSAGE	RATIO MULTIPLIER APPLIED TO MIDTERM MEANS (BASELINE MEAN/MIDTERM MEAN)
Oral reading fluency score (correct words per minute)	49.73	42.24	1.18
Reading comprehension score (number correct out of 5)	2.84	2.09	1.36

## REFERENCE

Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 187–220). Westport, Connecticut, USA: Praeger.

# ANNEX C: EGRA, EGMA, AND SSME INSTRUMENTS FOR THE LIBERIA ENDLINE ASSESSMENT

1. Student Instrument
2. Student Reading Stimulus Sheets
3. Student Math Stimulus Sheets
4. Teacher Instrument
5. Principal Instrument
6. School Inventory



## Student Instrument

LTTP2: Reading and Mathematics Assessment  
Endline Assessment: June, 2015



### General Instructions

It is important to establish a playful and relaxed rapport with the children to be assessed, via some simple initial conversation among topics of interest to the child. The child should perceive the following assessment as a game to be enjoyed rather than a severe situation. After you have finished, thank the child for his/her time and effort.

### Verbal Consent

Read the text in the box clearly to the child:

**My name is \_\_\_\_\_ . I work with the Ministry of Education in Liberia.**

- **We are trying to understand how children learn to read. You were picked by chance, like in a raffle or lottery.**
- **We would like your help in this. But you do not have to take part if you do not want to.**
- **We are going to play a reading game. I am going to ask you to read letters, words and a short story out loud. We are also going to play some counting games and some number games.**
- **Using this stopwatch, I will see how long it takes you to read.**
- **This is NOT a test and it will not affect your grade at school.**
- **I will also ask you questions about your family, like what language your family uses at home and some of the things your family has.**
- **I will NOT write down your name so no one will know these are your answers.**
- **Once again, you do not have to participate if you do not wish to. Once we begin, if you would rather not answer a question, that's all right.**
- **Can we get started?**

**Check box if verbal consent is obtained:**       **YES**

*(If verbal consent is not obtained, thank the child and move on to the next child, using this same form)*

A. Date of assessment :		H. Unique student code :	
B. Assessor name/code :		I . Student's grade level :	<input type="radio"/> 1 = 1 <sup>st</sup> grade <input type="radio"/> 2 = 2nd grade <input type="radio"/> 3 = 3rd grade
C. NAME and location of school :		J. Class section:	
D. Unique School code :		K. Student's month and year of birth :	Month : _____ Year : _____ Age: _____
E. School shift :	<input type="radio"/> 1 = Full day <input type="radio"/> 2 = Morning <input type="radio"/> 3 = Afternoon	L. Longitudinal selection If YES, write the assigned code number for this student.	<input type="radio"/> 1 = YES <input type="radio"/> 0 = No  Code: _____
F. School Type:	<input type="radio"/> 1 = cohort 1 <input type="radio"/> 2 = cohort 2 <input type="radio"/> 3 = external control <input type="radio"/> 4 = longitudinal	M. Student's gender	<input type="radio"/> 1 = girl    0 = boy
G. Teacher name <b><i>(important!)</i></b>			

## Task 1. Orientation to Print

Show the child the paragraph segment on the last page of the student assessment (Section 6).

Read the instructions in the gray boxes below, recording the child's response before moving to the next instruction.

**I don't want you to read this now. On this page, where would you begin to read?  
Show me with your finger.**

*[Child puts finger on the top row, left-most word]*     Correct     Incorrect     No Response

**Now show me in which direction you would read next.**

*[Child moves finger from left to right]*     Correct     Incorrect     No Response

**When you get to the end of the line, where would you read next?**

*[Child moves finger to left-most word of second  
line]*     Correct     Incorrect     No Response

## Section 2. Letter Name Knowledge

Show the child the sheet of letters on the first page of the student assessment. Say,

Here is a page full of letters of the alphabet. Please tell me the NAMES of as many letters as you can--not the SOUNDS of the letters, but the names.

1. For example, the name of this letter [point to O] is "OH".

Now you try: tell me the name of this letter [point to V]:

[If correct:] Good, the name of this letter is "VEE."

[ If incorrect:] The name of this letter is "VEE."

2. Now try another one: tell me the name of this letter [point to L]:

[If correct:] Good, the name of this letter is "ELL."

[If incorrect:] The name of this letter is "ELL."

Do you understand what you are supposed to do? When I say "begin," name the letters as best as you can. I will keep quiet and listen to you, unless you need help. Ready? Begin.



Set the timer on 1 minute. Start the timer when the child reads the first letter. Follow along with your pen and clearly mark any incorrect letters with a slash (/). Count self-corrections as correct. Stay quiet, except when providing answers as follows: if the child hesitates for 3 seconds, provide the name of the letter, point to the next letter and say "Please go on." Mark the letter you provide to the child as incorrect.

WHEN THE TIMER REACHES 0, SAY, "stop." Mark the final letter read with a bracket ( ] ).

**Early stop rule:** If the child does not give a single correct response on the first line, say "Thank you!", draw a line through the letters in the first row, discontinue this exercise, check the box at the bottom, and go on to the next exercise.

L i h R S y E O n T	10
i e T D A t a d e w	20
h O e m U r L G R U	30
g R B E i f m t s r	40
S T C N p A F c a E	50
y s Q A M C O t n P	60
e A e s O F h u A t	70
R G H b S i g m i L	80
L i N O e o E r p X	90
N A c D d l O j e n	100

Time left on stopwatch if student completes in LESS than 60 seconds: \_\_\_\_\_

Exercise was discontinued as child had no correct answers in the first line.

### Task 3. Phonemic Awareness

This is **NOT** a timed exercise and **THERE IS NO STUDENT SHEET**. Read aloud each set of words **once** and have the student say which word begins with a different sound. Read these instructions to the child:

**This is listening exercise. I'm going to say THREE words. ONE of them begins with a different sound, and you tell me which word BEGINS WITH A DIFFERENT SOUND**

1. For example:

“lost”, “map”, “like”. Which word begins with a different sound?

[If correct:] **Very good, “map” begins with a different sound.**

[If incorrect:] “lost”, “map”, “like”. “map” begins with a different sound than “lost” and “like.”

2. Now try another one: “train”, “trip”, “stop”. Which word begins with a different sound?

[If correct:] **Very good, “stop” begins with a different sound.**

[If incorrect:] “train”, “trip”, “stop”. “stop” begins with a different sound than “train” and “trip.”

**Do you understand what you are supposed to do?**

Pronounce each set of words **once slowly** (about 1 word per second). If the child does not respond after 3 seconds mark it no response and move on.

**Early stop rule:** If the child gets the **first 5 sets** of answers **incorrect or no response**, draw the line through each of the 5 first rows, discontinue this exercise, check the box at the bottom of this page and go on to the next exercise.

Which word begins with a different sound? [repeat each set ONCE]										
1	boy	ball	cat	[cat]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
2	man	can	mad	[can]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
3	pan	late	pin	[late]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
4	back	ten	tin	[back]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
5	fish	fat	cat	[cat]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
6	boat	bit	coat	[coat]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
7	day	bag	dot	[bag]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
8	can	girl	cold	[girl]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
9	run	race	sand	[sand]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response
10	leg	make	lay	[make]	<input type="checkbox"/>	Correct	<input type="checkbox"/>	Incorrect	<input type="checkbox"/>	No Response

Exercise was discontinued as child had no correct answers in **the first five sets** of words.

## Task 4. Familiar Word Identification

Show the child the sheet of words on the second page of the student assessment. Say,

**Here are some words. I would like you to read me as many words as you can (do not spell the words, but read them).**

**For example, this word is: "CAT".**

**1. Now you try:** [point to the word "mat" and say ]**please read this word:**

[If correct]: **Good, this word is "mat."**

[If incorrect]: **This word is "mat."**

**2. Now try another one:** [point to the word "top"] **please read this word :**

[If correct]: **Good, this word is "top."**

[If incorrect]: **This word is "top."**

**Do you understand what are you supposed to do? When I say "begin," read the words as best as you can. I will keep quiet and listen to you, unless you need help. Ready? Begin.**



**Start the timer when the child reads the first word.** Follow along with your pencil and clearly mark any incorrect words with a slash (/). Count self-corrections as correct. **Stay quiet**, except when providing answers as follows: if the child hesitates for 3 seconds, read the word, point to the next word and say "**Please go on.**" Mark the word you read to the child as incorrect.

**WHEN THE TIMER REACHES 0, SAY, "stop." Mark the final word read with a bracket (⌋).**

**Early stop rule:** *If the child gives no correct answers on the first line, say, "Thank you!", discontinue this exercise, draw the line through the words in the first row, check the box at the bottom of the page, and go on to the next exercise.*

but	time	in	the	also	5
make	no	its	said	were	10
came	very	do	after	long	15
water	as	all	for	even	20
her	was	three	been	more	25
that	must	can	around	it	30
another	words	back	called	work	35
could	an	him	on	see	40
than	get	not	where	what	45
you	if	their	through	when	50

Time on stopwatch if student completes in LESS than 60 seconds: \_\_\_\_\_

Exercise was discontinued as child had no correct answers in the first line.

## Task 5. Simple unfamiliar nonword decoding

Show the child the sheet of nonwords on the third page on the student form. Say,

**Here are some made-up words. I would like you to read me as many made-up words as you can (do not spell the words, but read them).**

**For example, this made-up word is: “ut”.**

**1. Now you try:** [point to the next word: “dif” and say] **please read this word**

[If correct]: **“Very good: dif”**

[If incorrect]: **This made-up word is “dif.”**

**2. Now try another one:** [point to the next word: mab and say] **please read this word.**

[If correct]: **“Very good: mab”**

[If incorrect]: **This made-up word is “mab.”**

**Do you understand what you are supposed to do? When I say “begin,” read the words as best as you can. I will keep quiet and listen to you, unless you need help. Ready? Begin.**

 **Start the timer when the child reads the first word.** Follow along with your pencil and clearly mark any incorrect words with a slash (/). Count self-corrections as correct. **Stay quiet**, except when providing answers as follows: if the child hesitates for 3 seconds, provide the word, point to the next word and say **“Please go on.”** Mark the word you provide to the child as incorrect.

**WHEN THE TIMER REACHES 0, SAY, “Stop.” Mark the final word read with a bracket ( ] ).**

**Early stop rule:** *If the child gives no correct answers on the first line, say “Thank you!”, discontinue this exercise, draw the line through the words in the first row, check the box at the bottom of the page, and go on to the next exercise.*

loz	ep	yat	zam	tob	5
zom	ras	mon	jaf	duz	10
tam	af	ked	ig	el	15
tig	pek	dop	zac	ik	20
uf	ral	ep	bab	vif	25
lut	sig	zop	zar	jaf	30
ruz	huf	wab	ak	jep	35
wub	dod	ik	vus	nux	40
pek	zel	bef	wab	hiz	45
wof	ib	dek	zek	vok	50

Time left on stopwatch if student completes in LESS than 60 seconds: \_\_\_\_\_

Exercise was discontinued as child had no correct answers in the first line.

## Task 6. Passage reading and Comprehension.

Show the child the story on the last page of the student form. Say,

**Here is a short story. I want you to read this aloud. When you finish, I will ask you some questions about what you have read. Do you understand what are you supposed to do? When I say “begin,” read the story as best as you can. I will keep quiet and listen to you, unless you need help. Ready? Begin.**



**Set the timer on 1 minute. Start the timer when the child reads the first word.** Follow along with your pencil and clearly mark any incorrect words with a slash (/). Count self-corrections as correct. **Stay quiet**, except when providing answers as follows: if the child hesitates for 3 seconds, provide the word, point to the next word and say “**Please go on.**”

Mark the word you provide to the child as incorrect. **WHEN THE TIMER REACHES 0, SAY,, “stop.”** Mark the final word read with a bracket (⌋). If the child gets the entire first line incorrect, discontinue this exercise – both reading and comprehension questions -, check the box below and go on to the next exercise.

**STOP THE CHILD AT 0 SECONDS AND MARK WITH A BRACKET (⌋).**

Take **the text away** from the child after they read it. Read instructions to the child. Then read each question slowly and clearly. After you read each question, give the child at most 15 seconds to answer each question. Mark the answers to the questions as correct or incorrect.

**Now I am going to ask you a few questions about the story you just read. Try to answer the questions as best as you can.**

Kemah lives near the big river.	6	<b>Where does Kemah live?</b> [near the big river] <input type="checkbox"/> Correct <input type="checkbox"/> Incorrect <input type="checkbox"/> No Response
There is a big tree by the river where Kemah lives. Kemah likes to sit in the tree.	24	<b>Where does Kemah like to sit when she goes to the river?</b> [in the tree, in the big tree near the river] <input type="checkbox"/> Correct <input type="checkbox"/> Incorrect <input type="checkbox"/> No Response
Every day after school, she stops by the tree and looks for a place to sit. She climbs the tree and sits on a branch.	49	<b>What does Kemah do after she climbs the tree?</b> [she sits on a branch, she finds a place to sit in the tree] <input type="checkbox"/> Correct <input type="checkbox"/> Incorrect <input type="checkbox"/> No Response
She looks at the fish in the river.	57	<b>What does Kemah do when she sits in the tree?</b> [she looks at the fish in the river, looks at fish] <input type="checkbox"/> Correct <input type="checkbox"/> Incorrect <input type="checkbox"/> No Response
Kemah is happy.	60	<b>Why is Kemah happy when she sits in the tree?</b> [ she likes to look at fish in the river, she likes the tree] <input type="checkbox"/> Correct <input type="checkbox"/> Incorrect <input type="checkbox"/> No Response

Time left on stopwatch: \_\_\_\_\_

Test Discontinued because child read NO words on the first line:

## Task 7. Listening Comprehension

This is NOT a timed exercise and **THERE IS NO STUDENT SHEET**. The administrator reads aloud the following passage **ONLY ONE TIME**, slowly (about 1 word per second). Say,

**I am going to read you a short story aloud ONCE and then ask you some questions. Please listen carefully and answer the questions as best as you can. Do you understand what are you supposed to do?**

**Musu goes to the Bong Town School every day on a motor bike. One day, Musu could not get a motor bike to take her to school because it was raining and they were all busy. Musu did not want to get wet. Then, one old man said, “You can have my son’s raincoat.” Musu was happy. She did not have to be wet at school.**

1. How does Musu usually get to school?

[on a motorbike] Correct Incorrect No Response

2. Why did the old man give Musu a coat?

[Because it was raining] Correct Incorrect No response

3. Why was Musu happy at school that day?

[because she did not have to be wet at school] Correct Incorrect No Response

*Say: Now... I have some mathematics tasks that I want you to do for me. Please listen carefully and do the best you can. Some tasks are harder than others, so don't worry if you're not sure about all of them. Just give it your best try. OK?*

Task 1: Number Identification					Sheet T1	60 seconds																																										
<p>👤 Here are some numbers. I want you to point to each number and tell me what the number is. I am going to use this stopwatch and will tell you when to begin and when to stop.</p> <p>- [point to first number] Start here and go across (sweep finger across first line). Are you ready? . . . Start.</p> <p>- What number is this ?</p>					<p>• Discontinue test if the child answers <b>0 correct in the first line going ACROSS.</b></p>																																											
<p>✂ ( / ) Incorrect or no response ( ) After the last number read</p> <table border="1"> <thead> <tr> <th colspan="5"></th> <th>Tot. Cum.</th> </tr> </thead> <tbody> <tr> <td>4</td> <td>10</td> <td>28</td> <td>58</td> <td>807</td> <td>(5)</td> </tr> <tr> <td>94</td> <td>368</td> <td>30</td> <td>106</td> <td>17</td> <td>(10)</td> </tr> <tr> <td>9</td> <td>39</td> <td>14</td> <td>711</td> <td>83</td> <td>(15)</td> </tr> <tr> <td>423</td> <td>34</td> <td>72</td> <td>245</td> <td>77</td> <td>(20)</td> </tr> <tr> <td>187</td> <td>52</td> <td>22</td> <td>19</td> <td>33</td> <td>(25)</td> </tr> <tr> <td>646</td> <td>12</td> <td>64</td> <td>49</td> <td>301</td> <td>(30)</td> </tr> </tbody> </table>										Tot. Cum.	4	10	28	58	807	(5)	94	368	30	106	17	(10)	9	39	14	711	83	(15)	423	34	72	245	77	(20)	187	52	22	19	33	(25)	646	12	64	49	301	(30)	<p>Tell the child to go on</p> <p>• If a child stops on a number for <u>5 SECONDS.</u></p>	
					Tot. Cum.																																											
4	10	28	58	807	(5)																																											
94	368	30	106	17	(10)																																											
9	39	14	711	83	(15)																																											
423	34	72	245	77	(20)																																											
187	52	22	19	33	(25)																																											
646	12	64	49	301	(30)																																											
<p>✂ Time left (seconds):</p>					<p>Test discontinued: <input type="checkbox"/></p>																																											
<p>✂ Total Number correct:</p>																																																

Task 2: Quantity Discrimination - Example		Sheet T2	NOT TIMED
<p><u>P1:</u></p> <p>👤 Look at these numbers. Tell me which number is bigger.</p> <p style="text-align: center;">10    4</p> <p>✓ 👤 That's correct, 10 is bigger. Let's do another one.</p> <p>✗ 👤 The bigger number is 10. [Point to 10] This is 10. [Point to 4] This is 4. 10 is bigger than 4. Let's do another one.</p>			
<p><u>P2:</u></p> <p>👤 Look at these numbers. Tell me which number is bigger.</p> <p style="text-align: center;">8    12</p> <p>✓ 👤 That's right, 12 is bigger. Let's continue.</p> <p>✗ 👤 The bigger number is 12. [Point to 8] This number is 8. [Point to 12] This is 12. 12 is bigger than 8. Let's continue.</p>			

Task 2: Quantity Discrimination - Exercise	Sheet T2	NOT TIMED																																								
<p>(say) <b>Look at these numbers. Tell me which number is bigger.</b> [repeat for each item]</p>		<p>Discontinue test if the child gets <b>4 successive errors</b></p> <p>Tell child to go on,  <ul style="list-style-type: none"> <li>• If the child doesn't respond after <b>5 SECONDS</b>.</li> </ul> </p>																																								
<p>(✓) on the "1" for correct response  (✓) on the "0" for Incorrect or no response  Circle the "0" for SELF-CORRECT.</p> <table border="1" style="width: 100%; text-align: center;"> <tbody> <tr> <td>6</td><td><u>8</u></td><td>1</td><td>0</td> <td>229</td><td><u>238</u></td><td>1</td><td>0</td> </tr> <tr> <td>54</td><td><u>63</u></td><td>1</td><td>0</td> <td>675</td><td><u>684</u></td><td>1</td><td>0</td> </tr> <tr> <td>279</td><td><u>381</u></td><td>1</td><td>0</td> <td>16</td><td><u>25</u></td><td>1</td><td>0</td> </tr> <tr> <td>79</td><td><u>80</u></td><td>1</td><td>0</td> <td>82</td><td><u>91</u></td><td>1</td><td>0</td> </tr> <tr> <td>44</td><td><u>53</u></td><td>1</td><td>0</td> <td>523</td><td><u>532</u></td><td>1</td><td>0</td> </tr> </tbody> </table>			6	<u>8</u>	1	0	229	<u>238</u>	1	0	54	<u>63</u>	1	0	675	<u>684</u>	1	0	279	<u>381</u>	1	0	16	<u>25</u>	1	0	79	<u>80</u>	1	0	82	<u>91</u>	1	0	44	<u>53</u>	1	0	523	<u>532</u>	1	0
6	<u>8</u>		1	0	229	<u>238</u>	1	0																																		
54	<u>63</u>	1	0	675	<u>684</u>	1	0																																			
279	<u>381</u>	1	0	16	<u>25</u>	1	0																																			
79	<u>80</u>	1	0	82	<u>91</u>	1	0																																			
44	<u>53</u>	1	0	523	<u>532</u>	1	0																																			
<p>4 successive errors so Task Discontinued: <input type="checkbox"/></p>		<p>Total Correct: <input type="text"/></p>																																								

Task 3: Missing Number - Example	Sheet T3	NOT TIMED
<p>(say) <b>Here are some numbers. 1, 2, 3, what number goes here (point to the dash)?</b>  <b>1, 2, 3, ____</b>  (if student is correct, say) <b>That's correct, 4. Let's do another one.</b>  (if student is incorrect, say) <b>The number four goes here. Say the numbers with me.</b>  [Point to each number] . . . <b>1, 2, 3, 4. 4 goes here. Let's do another one.</b></p>		
<p><u>P2:</u>  (say) <b>Here are some numbers. 17, [point to dash], 19, what number goes here?</b> [point to dash again]  <b>17, ____, 19</b>  (if student is correct, say) <b>That's right, 18. Let's do some more.</b>  (if student is incorrect, say) <b>The number 18 goes here. Say the numbers with me.</b> [Point to each number] . . . <b>17, 18, 19.</b> [Point to dash] <b>18 goes here. Let's do some more.</b></p>		

Task 3: Missing Number - Exercise	Sheet T3	NOT TIMED																																																												
<p>(say) <b>Here are some more numbers.</b> [Point to the dash] . . . <b>What number goes here?</b> [Repeat for each item]</p>		<p>Discontinue test if the child gets <b>4 successive errors</b></p> <p>Tell child to go on,  <ul style="list-style-type: none"> <li>• If the child doesn't respond after <b>5 SECONDS</b>.</li> </ul> </p>																																																												
<p>(✓) on the "1" for correct response  (✓) on the "0" for Incorrect or no response  Circle the "0" for SELF-CORRECT.</p> <table border="1" style="width: 100%; text-align: center;"> <tbody> <tr> <td>2</td><td><u>4</u></td><td>6</td><td></td><td>1</td><td>0</td> <td>100</td><td>200</td><td>300</td><td><u>400</u></td><td>1</td><td>0</td> </tr> <tr> <td>245</td><td>250</td><td>255</td><td><u>260</u></td><td>1</td><td>0</td> <td>30</td><td>35</td><td><u>40</u></td><td>45</td><td>1</td><td>0</td> </tr> <tr> <td><u>30</u></td><td>40</td><td>50</td><td>60</td><td>1</td><td>0</td> <td>18</td><td><u>20</u></td><td>22</td><td>24</td><td>1</td><td>0</td> </tr> <tr> <td><u>89</u></td><td>90</td><td>91</td><td></td><td>1</td><td>0</td> <td>348</td><td>349</td><td><u>350</u></td><td></td><td>1</td><td>0</td> </tr> <tr> <td>305</td><td>310</td><td><u>315</u></td><td>320</td><td>1</td><td>0</td> <td>500</td><td>400</td><td><u>300</u></td><td></td><td>1</td><td>0</td> </tr> </tbody> </table>			2	<u>4</u>	6		1	0	100	200	300	<u>400</u>	1	0	245	250	255	<u>260</u>	1	0	30	35	<u>40</u>	45	1	0	<u>30</u>	40	50	60	1	0	18	<u>20</u>	22	24	1	0	<u>89</u>	90	91		1	0	348	349	<u>350</u>		1	0	305	310	<u>315</u>	320	1	0	500	400	<u>300</u>		1	0
2	<u>4</u>		6		1	0	100	200	300	<u>400</u>	1	0																																																		
245	250	255	<u>260</u>	1	0	30	35	<u>40</u>	45	1	0																																																			
<u>30</u>	40	50	60	1	0	18	<u>20</u>	22	24	1	0																																																			
<u>89</u>	90	91		1	0	348	349	<u>350</u>		1	0																																																			
305	310	<u>315</u>	320	1	0	500	400	<u>300</u>		1	0																																																			
<p>4 successive errors so Task Discontinued: <input type="checkbox"/></p>		<p>Total Correct: <input type="text"/></p>																																																												

Task Discontinued:	Total Correct:	
--------------------	----------------	--

<b>Task 4A: Addition</b>		Sheet T4	<b>TIMED: 60</b> seconds																																												
<p>(say) <i>Here are some addition (plus) problems. I am now going to use this stopwatch. Tell me the answers to as many as you can. If you don't know an answer, move to the next problem. Are you ready? . . .</i></p> <p><i>- Start here [point to first problem], go across, (sweep finger across first row) and tell me the answer to each problem.</i></p>		<p>Discontinue test if child misses <b>4 successive items.</b></p> <p>Tell child to go on,  <ul style="list-style-type: none"> <li>• If a child stops on a number for <b><u>5 SECONDS.</u></b></li> </ul> </p>																																													
<p>✎ ( / ) Incorrect or no response ( ) After last problem attempted</p> <table border="1"> <thead> <tr> <th colspan="3"></th> <th>Tot Cum.</th> </tr> </thead> <tbody> <tr> <td>4 + 2 = (6)</td> <td>8 + 2 = (10)</td> <td>8 + 6 = (14)</td> <td>(3)</td> </tr> <tr> <td>16 + 4 = (20)</td> <td>7 + 1 = (8)</td> <td>5 + 4 = (9)</td> <td>(6)</td> </tr> <tr> <td>10 + 3 = (13)</td> <td>10 + 10 = (20)</td> <td>2 + 2 = (4)</td> <td>(9)</td> </tr> <tr> <td>5 + 7 = (12)</td> <td>6 + 6 = (12)</td> <td>3 + 4 = (7)</td> <td>(12)</td> </tr> <tr> <td>6 + 2 = (8)</td> <td>5 + 6 = (11)</td> <td>15 + 5 = (20)</td> <td>(15)</td> </tr> <tr> <td>4 + 5 = (9)</td> <td>7 + 2 = (9)</td> <td>3 + 9 = (12)</td> <td>(18)</td> </tr> <tr> <td>13 + 3 = (16)</td> <td>1 + 5 = (6)</td> <td>5 + 5 = (10)</td> <td>(21)</td> </tr> <tr> <td>2 + 11 = (13)</td> <td>3 + 2 = (5)</td> <td>6 + 4 = (10)</td> <td>(24)</td> </tr> <tr> <td>6 + 10 = (16)</td> <td>10 + 5 = (15)</td> <td>5 + 3 = (8)</td> <td>(27)</td> </tr> <tr> <td>7 + 3 = (10)</td> <td>4 + 7 = (11)</td> <td>11 + 9 = (20)</td> <td>(30)</td> </tr> </tbody> </table>					Tot Cum.	4 + 2 = (6)	8 + 2 = (10)	8 + 6 = (14)	(3)	16 + 4 = (20)	7 + 1 = (8)	5 + 4 = (9)	(6)	10 + 3 = (13)	10 + 10 = (20)	2 + 2 = (4)	(9)	5 + 7 = (12)	6 + 6 = (12)	3 + 4 = (7)	(12)	6 + 2 = (8)	5 + 6 = (11)	15 + 5 = (20)	(15)	4 + 5 = (9)	7 + 2 = (9)	3 + 9 = (12)	(18)	13 + 3 = (16)	1 + 5 = (6)	5 + 5 = (10)	(21)	2 + 11 = (13)	3 + 2 = (5)	6 + 4 = (10)	(24)	6 + 10 = (16)	10 + 5 = (15)	5 + 3 = (8)	(27)	7 + 3 = (10)	4 + 7 = (11)	11 + 9 = (20)	(30)		
			Tot Cum.																																												
4 + 2 = (6)	8 + 2 = (10)	8 + 6 = (14)	(3)																																												
16 + 4 = (20)	7 + 1 = (8)	5 + 4 = (9)	(6)																																												
10 + 3 = (13)	10 + 10 = (20)	2 + 2 = (4)	(9)																																												
5 + 7 = (12)	6 + 6 = (12)	3 + 4 = (7)	(12)																																												
6 + 2 = (8)	5 + 6 = (11)	15 + 5 = (20)	(15)																																												
4 + 5 = (9)	7 + 2 = (9)	3 + 9 = (12)	(18)																																												
13 + 3 = (16)	1 + 5 = (6)	5 + 5 = (10)	(21)																																												
2 + 11 = (13)	3 + 2 = (5)	6 + 4 = (10)	(24)																																												
6 + 10 = (16)	10 + 5 = (15)	5 + 3 = (8)	(27)																																												
7 + 3 = (10)	4 + 7 = (11)	11 + 9 = (20)	(30)																																												
Time left (seconds):		Total number correct:																																													
Task Discontinued:																																															

**Task 4B: Subtraction**

Sheet T4

**TIMED: 60**  
seconds

(say) Now we are going to do some subtraction problems. Do as many as you can correctly. If you don't know an answer, move to the next problem. Are you ready? . . . Start here [point to first problem], go across (sweep finger across line) and tell me the answer to each problem.

Discontinue test if student gets **4 successive errors**.

Tell child to go on,  
• If a child stops on a number for **5 SECONDS**.

✂ ( / ) Incorrect or no response			Tot Cum.
( ) After the last problem attempted			
6 - 2 = (4)	10 - 2 = (8)	14 - 6 = (8)	(3)
20 - 4 = (16)	8 - 1 = (7)	9 - 4 = (5)	(6)
13 - 3 = (10)	20 - 10 = (10)	4 - 2 = (2)	(9)
12 - 7 = (5)	12 - 6 = (6)	7 - 4 = (3)	(12)
8 - 2 = (6)	11 - 6 = (5)	20 - 5 = (15)	(15)
9 - 5 = (4)	9 - 2 = (7)	12 - 9 = (3)	(18)
16 - 3 = (13)	6 - 5 = (1)	10 - 5 = (5)	(21)
13 - 11 = (2)	5 - 2 = (3)	10 - 4 = (6)	(24)
16 - 10 = (6)	15 - 5 = (10)	8 - 3 = (5)	(27)
10 - 3 = (7)	11 - 7 = (4)	20 - 9 = (11)	(30)

✂ Time left (seconds):	<input type="text"/>	Total number correct:	<input type="text"/>
Task Discontinued	<input type="text"/>		<input type="text"/>

Task 5: Word Problems	NO student sheet	NOT TIMED
✎ ✦ Counters		
<p>(say) <i>I have some problems that I am going to ask you to solve for me. Here are some things to help you count. You can use them or your fingers if you need them, but you don't have to use them. Listen very carefully to each problem. If you need, I will repeat problem for you. Okay, let's get started.</i></p> <p><u>Example:</u>  (say) <b>Foya has 2 mangoes. Helen has 3 mangoes. How many mangoes do they have altogether?</b>  (if answer is correct, say) <b>That's right. They have 5 mangoes altogether. Let's do some more.</b>  (if answer is incorrect, say) <b>They have 5 mangoes altogether. Watch me.</b>  <i>[Use the counters and read the problem, demonstrating 2 counters for Foya and 3 counters for Helen]</i> <b>When we count add them altogether, we get 5 mangoes. Let's do some more.</b></p>	<p>Discontinue task If the child gets <b>4 successive errors</b></p> <p>Tell child to go on, If she stops on a number for <u>5 SECONDS</u>.  (without beginning to use the counters, count on her fingers, etc.)</p> <p>Tell child to go on, if, after 60 seconds, she has not given a response <b>even if</b> she is working on the problem with counters and/or fingers.</p>	
<p>(say) <i>Now you will work out more questions that I will read to you. Remember some of these questions may be hard even for older children, so do your best. Remember, you can use these or your fingers to help you answer the questions [point to the counters]. Okay, let's get started.</i></p> <p>✎ For each problem:  (✓) 1 = Correct.  (✓) 0 = Incorrect or no response.</p>		
<p><u>Problem 1</u>  (say) <b>Momo has 2 mangoes. His father gives him 5 more. How many does he have now ?</b>  Correct answer: 7 <input type="text" value="1"/> <input type="text" value="0"/></p>		
<p><u>Problem 2</u>  (say) <b>Fayah has 6 oranges. He eats 3 oranges. How many oranges does he have now?</b>  Correct answer: 3 <input type="text" value="1"/> <input type="text" value="0"/></p>		
<p><u>Problem 3</u>  (say) <b>Yaya has 8 pencils. Abdul has 4 pencils. How many more pencils does Yaya have than Abdul?</b>  Correct answer: 4 <input type="text" value="1"/> <input type="text" value="0"/></p>		

Problem 4

(say) ***There are 8 children walking to school. 6 are boys and the rest are girls. How many girls are walking to school?***

Correct answer: 2

Problem 5

(say) ***I have 7 bananas. How many more bananas do I need if I want to give one to each of my 12 friends?***

Correct answer: 5

 Total number correct:


## Student Context Interview

Say to the child: *Thank you very much. Now, I am going to ask you some questions about your family and about reading.*

S1	What language/dialect does your family speak most often at home?	English = 1 Others = 2 [ <b>Specify main one</b> ] _____ Don't know = 88 Refuse/No Answer = 99	
S2	What language/dialect do your parents read or write in?	Cannot read and write = 0 English = 1 Others = 2 [ <b>Specify main one</b> ] _____ Don't know = 88 Refuse/No Answer = 99	
S3	Do you have any reading books at home? (If no, skip to S5.)	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S4	If answer to S3 is Yes, in what language/dialects?	English = 1 Other = 2 [ <b>Specify main one</b> ] _____	
S5	Does anyone read aloud to you at home? (If No, skip to S7.)	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S6	If answer to S5 is Yes, in what language/dialects do they read to you?	English = 1 Other = 2 [ <b>Specify main one</b> ] _____	
S7	Do you practice reading aloud to someone at home? (If No, skip to S9.)	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S8	If answer to S7 is Yes, in what language(s) do you read?	English = 1 Other = 2 [ <b>Specify main one</b> ] _____ Don't know = 88      Refuse/No answer = 99	
S9	Have you ever repeated a grade?  If yes, which ones? ( <b>CIRCLE</b> the grades repeated.)	No = 0 Yes, Grade 1 = 1 Yes, Grade 2 = 2 Yes, Grade 3 = 3 Don't know = 88 Refuse/No answer = 99	
S10	Does your current teacher ever practice letter sounds with you? [Give student example of /k/ and /m/].	Never = 0 Often = 1 Always = 2 Refuse/No answer = 99	
S11	Does your teacher ever read aloud to you?	Never = 0 Often = 1 Always = 2 Refuse/No answer = 99	
S12	Did you eat lunch at break time at school yesterday [or last school day]?	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S13	Did you miss any school days last week?	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S15	The last time you did not do well on a test or assignment in school, did your parent(s) find out?  If yes, what did s/he do? ( <b>CIRCLE</b> all that apply. Do not read the choices to the child.)	No/Never learned about it = 0 Learned but did nothing = 1 Helped/encouraged me to do better = 2 Punished me physically = 3 Criticized me verbally = 4 Discussed with the teacher = 5 Argued with the teacher = 6	

		Other (Don't specify) = 7 Don't know = 88 Refuse/No answer = 99	
S16	The last time you did well on a test or assignment in school, did your parent(s) find out? (If No, skip to S18.)	No/Never learned about it = 0 Yes = 1 Don't know = 88 Refuse/No answer = 99	
S17	If the answer to S16 is Yes, what did your parent(s) do?	Learned but did nothing = 0 Congratulated or encouraged me = 1 Other = 3 _____	
S18	Do you have a library at your school?	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S19	Do you watch television at home?	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S20	Do you listen to radio at home?	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S21	Do you have electricity/current at home?	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S22	Did you eat before coming to school today?	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S23	Do you have books at school that you can take home to read?	No = 0                      Yes = 1 Don't know = 88      Refuse/No answer = 99	
S24	Does your teacher make you practice silent reading in class?	Never = 0 Often = 1 Always = 2 Refuse/No answer = 99	
S25	Does your teacher make you practice reading out loud in class?	Never = 0 Often = 1 Always = 2 Refuse/No answer = 99	
S26	Does your teacher assign reading for you to do at home?	Never = 0 Often = 1 Always = 2 Refuse/No answer = 99	
S27	Does your teacher ever make you re-tell a story during class?	Never = 0 Often = 1 Always = 2 Refuse/No answer = 99	

Thank the student by shaking his/her hand!

Example : O V L

---

L	i	h	R	S	y	E	O	n	T
i	e	T	D	A	t	a	d	e	w
h	O	e	m	U	r	L	G	R	U
g	R	B	E	i	f	m	t	s	r
S	T	C	N	p	A	F	c	a	E
y	s	Q	A	M	C	O	t	n	P
e	A	e	s	O	F	h	u	A	t
R	G	H	b	S	i	g	m	i	L
L	i	N	O	e	o	E	r	p	X
N	A	c	D	d	L	O	j	e	n

Example : cat mat top

---

but	time	in	the	also
make	no	its	said	were
came	very	do	after	long
water	as	all	for	even
her	was	three	been	more
that	must	can	around	it
another	words	back	called	work
could	an	him	on	see
than	get	not	where	what
you	if	their	through	when

*Example :*

---

ut	dif	mab		
loz	ep	yat	zam	tob
zom	ras	mon	jaf	duz
tam	af	ked	ig	el
tig	pek	dop	zac	ik
uf	ral	ep	bab	vif
lut	sig	zop	zar	jaf
ruz	huf	wab	ak	jep
wub	dod	ik	vus	nux
pek	zel	bef	wab	hiz
wof	ib	dek	zek	vok

Kemah lives near the big river. There is a big tree by the river where  
Kemah lives. Kemah likes to sit in the tree. Every day after school, she  
stops by the tree and looks for a place to sit. She climbs the tree and sits  
on a branch. She looks at the fish in the river. Kemah is happy.

EGRA (Early Grade Reading Assessment)  
Student Sheet  
Endline

EGMA (Early Grade Mathematics Assessment)

***Student Sheets***

T1.

4	10	28	58	807
94	368	30	106	17
9	39	14	711	83
423	34	72	245	77
187	52	22	19	33
646	12	64	49	301

10

4

8

12

6

8

63

54

381

279

79

80

44

53

238

229

675

684

25

16

82

91

523

532

T3.

1, 2, 3, \_\_\_\_\_

17, \_\_\_\_\_, 19

T3.

2, \_\_\_\_\_, 6

245, 250, 255, \_\_\_\_\_

\_\_\_\_\_40, 50, 60

\_\_\_\_\_, 90, 91

305, 310, \_\_\_\_\_, 320

100, 200, 300, \_\_\_\_\_

30, 35, \_\_\_\_\_, 45

18, \_\_\_\_\_, 22, 24

348, 349, \_\_\_\_\_

500, 400, \_\_\_\_\_

T4.

1	$4 + 2 = \square$	2	$8 + 2 = \square$	3	$8 + 6 = \square$
4	$16 + 4 = \square$	5	$7 + 1 = \square$	6	$5 + 4 = \square$
7	$10 + 3 = \square$	8	$10 + 10 = \square$	9	$2 + 2 = \square$
10	$5 + 7 = \square$	11	$6 + 6 = \square$	12	$3 + 4 = \square$
13	$6 + 2 = \square$	14	$5 + 6 = \square$	15	$15 + 5 = \square$
16	$4 + 5 = \square$	17	$7 + 2 = \square$	18	$3 + 9 = \square$
19	$13 + 3 = \square$	20	$1 + 5 = \square$	21	$3 + 9 = \square$
22	$2 + 11 = \square$	23	$3 + 2 = \square$	24	$6 + 4 = \square$
25	$6 + 10 =$	26	$10 + 5 = \square$	27	$5 + 3 = \square$
28	$7 + 3 =$	29	$4 + 7 = \square$	30	$11 + 9 = \square$

T4.

1	$6 - 2 = \square$	2	$10 - 2 = \square$	3	$14 - 6 = \square$
4	$20 - 4 = \square$	5	$8 - 1 = \square$	6	$9 - 4 = \square$
7	$13 - 3 = \square$	8	$20 - 10 = \square$	9	$4 - 2 = \square$
10	$12 - 7 = \square$	11	$12 - 6 = \square$	12	$7 - 4 = \square$
13	$8 - 2 = \square$	14	$11 - 6 = \square$	15	$20 - 5 = \square$
16	$9 - 5 = \square$	17	$9 - 2 = \square$	18	$12 - 9 = \square$
19	$16 - 3 = \square$	20	$6 - 5 = \square$	21	$10 - 5 = \square$
22	$13 - 11 = \square$	23	$5 - 2 = \square$	24	$10 - 4 = \square$
25	$16 - 10 = \square$	26	$15 - 5 = \square$	27	$8 - 3 = \square$
28	$10 - 3 = \square$	29	$11 - 7 = \square$	30	$20 - 9 = \square$

## TEACHER Instrument

Liberia Teacher Training Program 2  
Baseline Assessment, April-May, 2011

### Verbal Consent

My name is \_\_\_\_\_. I work with the Ministry of Education in Liberia.

- We are trying to understand how children learn to read and do math. Your school was selected through the process of statistical sampling. We would like your help in this. But you do not have to take part if you do not want to.
- Your name will not be mentioned anywhere in reports based on this survey. The survey data and the results will be published in the form of collective tables. The information acquired through this instrument will be shared with the Ministry of Education in the hopes of identifying areas where additional support may be needed. Your name may be used to track the training provided to you and whether the training is helping you.
- The name of your school will be recorded so that we can correctly link school, principal, teacher, and student data. Your school's name may be used to track your training and the help given to your school. The information acquired through this instrument will be shared with the Ministry of Education in the hopes of identifying areas where additional support may be needed.
- If you agree, I would ask you some questions regarding your normal activities at school, including your interactions with school staff, Ministry office staff, students, and parents.
- Then, with your assistance, I would randomly select 10 students in Grade 1, 10 students in Grade 2 and 10 students in Grade 3 to assess their reading and math skills. I would also ask these students about some of their normal school activities, school assets, language use, and reading practices at home, as well as home asset ownership. Selected students need only participate if they wish. I will spend 15-20 minutes interviewing each child. My interview with you will take 10 minutes.
- Are you willing to participate? Once again, you do not have to participate if you do not wish to. Once we begin, if you would rather not answer a question, that's all right.

Can we get started?

Check box if verbal consent is obtained :  YES  NO

*(If verbal consent is not obtained, thank the teacher and select the next one)*

## Section 1. Teacher Interview

SC1	School Name	
SC2	School Code	
SC3	School Type	Cohort 1= 1 ; Cohort 2=2; External=3
T1	Interviewer Name	
T2	Interviewer Code	
T3	Starting Time of Interview	____:____AM
T4	Ending Time of Interview	____:____PM
T5	Interview Date (dd/mm/yyyy)	Day:____Month:____Year:____
T6	Interview Status	Refused =1 Partially Completed = 2 Complete = 3
T7	[Is the teacher male or female?]	Male=0, Female=1
T8	Teacher name (make sure the same name is used in the questionnaires of the children who are taught by this teacher, so the questionnaires can be linked)	
T9	What type of teaching certificate do you have?	C Certificate = 0 B Certificate = 1 AA Certificate = 2 Other =3 Don't know = 88 Refuse/No Answer = 99
T10	What is your highest level of education?	Elementary=0 Junior High School =1 Senior High School =2 Associate = 3 Certificate C = 4 Certificate B = 5 Certificate AA = 6 Bachelors Degree = 7 Masters degree or other = 8 Other (Do not specify) = 9 Don't Know = 88 Refuse/No Answer = 99
T11	How many years of teaching experience do you have (Enter number)	
T12	Have you attended any in-service training or professional development sessions such as workshops in the last year?	No = 0                      Yes = 1 Don't Know = 88              Refuse/No Answer = 99
T13	Have you <i>ever</i> received training on how to teach reading?	No = 0                      Yes = 1 Don't Know = 88              Refuse/No Answer = 99
T14	If the answer to T13 is Yes, did you receive training for <i>this school year</i> ?	No = 0                      Yes = 1 Don't Know = 88              Refuse/No Answer = 99
T15	If the answer to T14 is Yes, about how many hours? (Enter the hours)	
T16	Have you <i>ever</i> received training on how to teach math?	No = 0                      Yes = 1 Don't Know = 88              Refuse/No Answer = 99
T17	If the answer to T16 is Yes, did you receive training for <i>this school year</i> ?	No = 0                      Yes = 1 Don't Know = 88              Refuse/No Answer = 99

T18	If the answer to T17 is Yes, about how many hours? (Enter the hours)	
T19	Have you received support visits at school this past year on how to teach reading?	No = 0 Don't Know = 88 Yes = 1 Refuse/No Answer = 99
T20	What grade or grades do you teach in this school year? [ <b>CIRCLE</b> all that apply]	Grade 1 Grade 2 Grade 3 Grade 4 Grade 5 Grade 6
T21	Have you been teaching the same class since the beginning of the school year?	No = 0 Don't Know = 88 Yes = 1 Refuse/No Answer = 99
T22	Do you keep an attendance record of students?	No = 0 Don't Know = 88 Yes = 1 Refuse/No Answer = 99
T23	How often do you develop lesson plans?	Daily = 1 Weekly = 2 Bi-weekly = 3 Monthly = 4 Don't Know = 88 Refuse/No Answer = 99
T24	Where do you develop lesson plans?	Home = 0 School = 1 Don't Know = 88 Refuse/No Answer = 99
T25	Do you have a written lesson plan for today?	No = 0 Don't Know = 88 Yes = 1 Refuse/No Answer = 99
T26	Could I please see it?	Was not able or willing to show it = 0 Could show it but it did not look very good = 1 Was able to show it and it looked pretty good = 3
T27	Do you have any scheduled time during the school day for lesson planning?	No = 0 Don't Know = 88 Yes = 1 Refuse/No Answer = 99
<b><i>I'm going to ask you about different activities you might do with your students. Think about the last 5 school days and tell me how frequently the following activities took place</i></b>		
T28.1	The whole class repeated letters or words that you said first, when teaching reading skills	Never = 0 Sometimes = 1 Frequently = 2 Everyday = 3
T28.2	Students sounded unfamiliar words they are learning	Never = 0 Sometimes = 1 Frequently = 2 Everyday = 3
T28.3	Students read aloud to teacher or another student	Never = 0 Sometimes = 1 Frequently = 2 Everyday = 3
T28.4	Students learned the meaning of new words	Never = 0 Sometimes = 1 Frequently = 2 Everyday = 3
T28.5	Students had to retell a story that they read during the week	Never = 0 Sometimes = 1 Frequently = 2 Everyday = 3

T28.6	Students were assigned reading to do on their own in school time	Never = 0 Sometimes = 1 Frequently = 2 Everyday = 3
T28.7	Students were assigned reading to do at home	Never = 0 Sometimes = 1 Frequently = 2 Everyday = 3
<b><i>Now we are going back to some questions about your overall work.</i></b>		
T29	Do you use the official reading curriculum in your classroom lessons?	Never = 0 Sometimes = 1 Frequently = 2 Everyday = 3
T30	Do you have teacher guides?	No = 0 Don't Know = 88 Yes = 1 Refuse/No Answer = 99
T31	If answer to T30 is Yes, "How useful do you find them?"	Not very useful = 0 Moderately useful = 1 Very Useful = 2 Don't Know = 88 Refuse/No Answer = 99
T32	How frequently does your principal observe (your) classes when they were in session?	Never = 0 Once a year = 1 Once every 2-3 months = 2 Once every month = 3 Once every two weeks = 4 Once every week = 5 Daily = 6 Don't Know = 88 Refuse/No Answer = 99
T33	Within the last year, did you receive an external inspection or support visit from Education Officers?	No = 0 Don't Know = 88 Yes = 1 Refuse/No Answer = 99
T34	How do you measure your students' progress? [Do not read options and <b>CIRCLE</b> those mentioned.]	Written tests = 1 Oral evaluations = 2 Their portfolios and other projects = 3 Their homework = 4 End of term evaluations = 5 Other (Do not specify) = 6
T35	What reading skills should your children have at the end of school year? [Do not read options. Just <b>CIRCLE</b> the number for those mentioned.]	Read grade level stories = 1 Sound out words they don't know = 2 Understand stories that they read = 3 Know letter names = 4 Other (Do not specify) = 5
T36	What math skills should your children have at the end of the school year? [do not read options. Just <b>CIRCLE</b> the number for those mentioned]	Do 1 digit addition and subtraction = 1 Know simple fractions = 2 Do multiplication and division Know how to tell time Know how to make change
T37	Parents have shown more interest in their children's reading and math in the last year.	Strongly disagree = 1 Disagree = 2 Neither agree nor disagree = 3 Agree = 4 Strongly Agree = 5 Refuse/No Answer = 99

T38	The whole school community has shown more interest in their children's reading and math in the last year.	Strongly disagree = 1 Disagree = 2 Neither agree nor disagree = 3 Agree = 4 Strongly Agree = 5 Refuse/No Answer = 99
T39	Students have shown more interest in their own reading and math skills in the last year.	Strongly disagree = 1 Disagree = 2 Neither agree nor disagree = 3 Agree = 4 Strongly Agree = 5 Refuse/No Answer = 99
T40	There have been more discussions at school about how to teach reading and math, in the past year.	Strongly disagree = 1 Disagree = 2 Neither agree nor disagree = 3 Agree = 4 Strongly Agree = 5 Refuse/No Answer = 99
T41	Do students take books from school to read at home?	Never = 0 Almost never = 1 Sometimes = 2 Almost always = 3 Always = 5 Refuse/No Answer = 99
T42	Do the teachers at this school work together as teams to solve problems related to teaching?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
T43	Have you made any special efforts to improve reading and/or math at your school this past year?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
T44	If the answer to T43 is Yes, ask "Why did you do that?" Do not read the answers, just <b>CIRCLE</b> the ones that apply.	We saw other schools doing it = 1 We thought it might be important because the children were tested = 2 The Ministry told us to do it = 3 An NGO told us to do it = 4 Teachers got teacher training that showed them how to teach reading better = 5 Other = 6 (Do not specify)
T45	If the answer to T43 is Yes, ask "How much more effort did you make that in a normal year?"	A little = 1 Some = 2 A good bit = 3 A lot = 4 Refuse / No Answer = 99
T48	Are you a volunteer teacher?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
T49	If Yes to T48, ask: 'Will you continue teaching in September 2011?'	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99

**THANK YOU VERY MUCH**

General Impression \_\_\_\_\_

Made possible with assistance from the American people

U.S. Agency for International Development

[www.usaid.gov](http://www.usaid.gov)

**PRINCIPAL Instrument**  
Liberia Teacher Training Program 2  
Baseline Assessment, April-May, 2011

Verbal Consent

My name is \_\_\_\_\_. I work with the Ministry of Education in Liberia.

- We are trying to understand how children learn to read and to do math. Your school was selected through the process of statistical sampling. We would like your help in this. But you do not have to take part if you do not want to.
- Your name might be used to decide how to improve services to your school, but it will not be mentioned in any reports written for this project. The information acquired through this instrument will be shared with the Ministry of Education in the hopes of identifying areas where additional support may be needed.
- The name of your school will be recorded so that we can correctly link school, principal, teacher, and student data. The information acquired through this instrument will be shared with the Ministry of Education in the hopes of identifying areas where additional support may be needed.
- If you agree, I would ask you some questions regarding your normal activities at school, including your interactions with your staff, Ministry office staff, students, and parents.
- Then, with your assistance, I would randomly select 10 students in Grade 1, 10 students in Grade 2, and 10 students in Grade 3 to assess their reading and math skills. I would also ask these students about some of their normal school activities, school assets, language use, and reading practices at home, as well as home asset ownership. Selected students need only participate if they wish. The entire interview process for all 30 students will last around 5 hours (30 minutes per each child). My interview with you will take 10 minutes. Finally, I would interview Grade 1, Grade 2 and Grade 3 teachers. I will spend 10 minutes interviewing each teacher.
- Are you willing to participate? Once again, you do not have to participate if you do not wish to. Once we begin, if you would rather not answer a question, that's all right.

Can we get started?

Check box if verbal consent is obtained :       YES       NO

*(If verbal consent is not obtained, thank the principal and terminate the exercise in this school)*

## Section 1. Classroom Observation

Before you start selecting students from the classrooms, conduct the following three activities. We need to get the data for as many classrooms as you end up visiting. If there are more than 3 teachers for grades 1, 2 or 3, just ignore the 4<sup>th</sup>. If there are only 2, leave the third column empty. If there is only 1, leave the second two columns empty. Note the first table is for Grade 1, the second is for Grade 2, and the third for Grade 3.

NOTE: you must make sure that you collect this data for the time of this assessment. Each cell that is empty must be filled out with the appropriate number. For example, in C02, for each TEACHER, you write his/or her name where indicated by cell titled 'Name of Teachers' and where the line is provided. Underneath and to the right of 'Grade 2: Male' enter the number of students who are in Grade 2 and who are boys and are taught by this teach. Repeat the same for all other teachers.

**April/May 2011**

C01			Name of teacher:	Name of teacher:	Name of teacher:
C02	Obtain a school register from the start of academic year, record enrollment for both female and male students	Grade 1: Male			
		Grade 1: Female			
C03			Name of teacher:	Name of teacher:	Name of teacher:
C04	Obtain a school register from the start of academic year, record enrollment for both female and male students	Grade 2: Male			
		Grade 2: Female			
C05			Name of teacher:	Name of teacher:	Name of teacher:
C06	Obtain a school register from the start of academic year, record enrollment for both female and male students	Grade 3: Male			
		Grade 3: Female			

## Section 2. Principal Interview

	School Name	
	School Code	
	Principal Name	
	School Type	Cohort 1= 1 ; Cohort 2=2; External=3
	Interviewer Name	
	Starting Time of Interview	____:____AM/PM
	Ending Time of Interview	____:____PM/PM
	Interview Date (dd/mm/yyyy)	Day: ____Month: ____Year: ____
HT5	Interview Status	Refused = 1 Partially Completed = 2 Complete = 3
HT6	What is your position at this school? (circle all that apply)	Principal = 1 Vice Principal = 2 Teacher = 3
HT7	[Is the principal male or female?]	Male = 0 Female=1
HT8	How many years have you been a principal? Enter years:	
HT9	What is your highest level of education?	Elementary=0 Junior High School =1 Senior High School =2 Associate = 3 Certificate C = 4 Certificate B = 5 Certificate AA = 6 Bachelors Degree = 7 Masters degree or other = 8 Other (Do not specify) = 9 Don't Know = 88 Refuse/No Answer = 99
HT10a & HT10b	What grades are taught at this school this year? (Enter the first grade taught and the last grade taught.)	First grade taught Last grade taught
HT11	Which grade(s) does your best teacher teach? (Enter grade):	
HT12	At what grade do you expect all of your students to read fluently? (Enter the grade:)	
HT13	Do you go into classrooms to observe your teachers teaching?	No = 0 Yes = 1 Don't Know = 88 Refuse/No Answer = 99

HT14	If yes to HT13, ask how often:	Never = 0 Once a year = 1 Once every 2-3 months = 2 Once every month = 3 Once every two weeks = 4 Once every week = 5 Daily = 6 Don't Know = 88 Refuse/No Answer = 99
HT15	Within the last year, did you receive an external inspection or support visit from Education Officers?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
HT16	Do you have a feeding program at school?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
HT17	Does your school observe student dress code?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
HT18	Do you have sufficient resource materials/textbooks?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
HT19	Do you have a library or reading room?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
HT20	Do you hold regular PTA meetings?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
HT21	If answer to HT20 was Yes, how many parents come to the PTA meetings	Few = 1 Some = 2 Most = 3 All = 4
HT22	In one year, how often do you send teachers for in-service teacher training, on average?	Less than 5 = 1 5 to 10 times = 2 More than 10 times = 3
HT23	Do you have lockable book storage at your school?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
HT24	Do you keep records of teachers' attendance?	No = 0                      Yes = 1 Don't Know = 88      Refuse/No Answer = 99
HT25	If answer to HT24 is Yes, say "Could I see the attendance records please?"	Was not able or willing to show = 1 Was able to show but badly kept (out of date, incomplete) = 2 Was able to show and in good shape = 3
HT26	Parents have shown more interest in their children's reading and math in the last year.	Strongly disagree = 1 Disagree = 2 Neither agree nor disagree = 3 Agree = 4 Strongly Agree = 5 Refuse or No Answer = 99

HT27	The whole school community has shown more interest in their children’s reading and math in the last year.	Strongly disagree = 1 Disagree = 2 Neither agree nor agree = 3 Agree = 4 Strongly Agree = 5 Refuse/No Answer = 99
HT28	Students have shown more interest in their own reading and math in the last year.	Strongly disagree = 1 Disagree = 2 Neither agree nor agree = 3 Agree = 4 Strongly Agree = 5 Refuse/No Answer = 99
HT29	There have been more discussions at school about how to teach reading and math, in the past year.	Strongly disagree = 1 Disagree = 2 Neither agree nor agree = 3 Agree = 4 Strongly Agree = 5 Refuse/No Answer = 99
HT30	Have you made any special efforts to improve reading and/or math at your school this past year?	No = 0                      Yes = 1 Don't Know = 88    Refuse/No Answer = 99
HT31	If the answer to HT31 is Yes, ask, “Why did you do that?” Do not read the answers, just <b>CIRCLE</b> the ones that apply.	We saw other schools doing it = 1 We thought it might be important because the children were tested = 2 The Ministry told us to do it = 3 An NGO told us to do it = 4 Teachers got teacher training that showed them how to teach reading better = 5 Other = 6 (Do not specify)
HT32	If the answer to HT31 is Yes, ask “How much more effort did you make that in a normal year?”	A little = 1 Some = 2 A good bit = 3 A lot = 4 Refuse/No Answer = 99
HT33	Does your school implement the Accelerated Learning Program (ALP) or Non Formal Education Program (NFE)?	No = 0                      Yes = 1 Don't Know = 88    Refuse/No Answer = 99
HT34	Are you currently teaching in this school?	No = 0                      Yes = 1 Don't Know = 88    Refuse/No Answer = 99
HT35	If answer to HT34 is Yes, “Ask: ‘What grades are you teaching? [CIRCLE all that apply]”	Grade 1 Grade 2 Grade 3 Grade 4 Grade 5 Grade 6

HT36	Are any of your teachers (other than you) also principals or vice-principals?	No = 0 Don't Know = 88	Yes = 1 Refuse/No Answer = 99
HT37	If answer to HT36 is Yes, ask 'How many?' [write the number]		
HT38	How many volunteer teachers did your school have at the beginning of 2010/2011 academic year? [write the number]		
HT39	How many volunteer teachers did you have at the beginning of 2010/2011 academic year? [write the number]		
HT40	How many volunteer teachers does your school have now? [write the number]		
HT41	How many Grade 1, Grade 2, and Grade 3 teachers does your school have? [write the number]		
HT42	How many Grade 1, Grade 2 and Grade 3 teachers are currently volunteer teachers? [write the number]		

**THANK YOU VERY MUCH**

General Impression \_\_\_\_\_

**Made possible with assistance from the American people**

U.S. Agency for International Development

[www.usaid.gov](http://www.usaid.gov)



## School Data Collection Summary Sheet –April 2011

Name of Fieldwork Team:	<input type="text"/>	Name of school:	<input type="text"/>
Date:	<input type="text"/>	School code:	<input type="text"/>
Location:	<input type="text"/>	Address of school:	<input type="text"/>
		Telephone number:	<input type="text"/>

### SUMMARY OF PUPIL SAMPLING, AND PUPIL & TEACHER INFORMATION COLLECTED

Grade/Name of class section	Total grade enrollment	N of absentees on date of visit	N of selected pupils who decline to participate	N of pupil assessments collected	N of Completed Teacher questionnaire collected?	N of Completed principal questionnaire collected? (record only once on Grade 2)
<b>GRADE 1</b>						
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
<i>Total Grade 1:</i>						
<b>GRADE 2</b>						
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
<i>Total Grade 2:</i>						
<b>GRADE 3</b>						
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
<i>Total Grade 3:</i>						

Signature of Team Representative: \_\_\_\_\_ Date: \_\_\_\_\_



**School Data Collection Summary Sheet –April 2011**

Name of Fieldwork Team:	<input type="text"/>	Name of school:	<input type="text"/>
Date:	<input type="text"/>	School code:	<input type="text"/>
Location:	<input type="text"/>	Address of school:	<input type="text"/>
		Telephone number:	<input type="text"/>

**SUMMARY OF PUPIL SAMPLING, AND PUPIL & TEACHER INFORMATION COLLECTED**

Grade/Name of class section	Total grade enrollment	N of absentees on date of visit	N of selected pupils who decline to participate	N of pupil assessments collected	N of Completed Teacher questionnaire collected?	N of Completed principal questionnaire collected? (record only once on Grade 2)
<b>GRADE 1</b>						
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
<b>Total Grade 1:</b>						
<b>GRADE 2</b>						
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
<b>Total Grade 2:</b>						
<b>GRADE 3</b>						
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
Section _____	M: F:	M: F:	M: F:			
<b>Total Grade 3:</b>						

Signature of Team Representative: \_\_\_\_\_ Date: \_\_\_\_\_

# ANNEX D: GENDER ANALYSIS OF SCORES, COHORT 2

An examination of reading scores by gender showed that boys and girls overall followed different patterns across assessments. **Table D-1** below shows the change in reading scores for pupils in Cohort 2 at the midterm and endline assessments. The boys scored consistently higher in most grades for the mathematics subtasks, but not the reading subtasks. This is explained in part by examination of the baseline scores, where the girls also consistently scored lower than the boys. However, of concern is that the girls did not make up the gap on the boys. **Figure D-1** demonstrates this observation; the gap by gender (boys minus girls) was 0.1 correct words per minute for the baseline assessment and  $-1.2$  for the endline. For correct number identification problems per minute (**Figure D-2**), the gap by gender (boys minus girls) was 4.3 and 3.8 between baseline and endline, still a statistically significant gender gap.

**Table D-1. Change in Boys' and Girls' Reading Scores, Cohort 2, by Subtask and Grade**

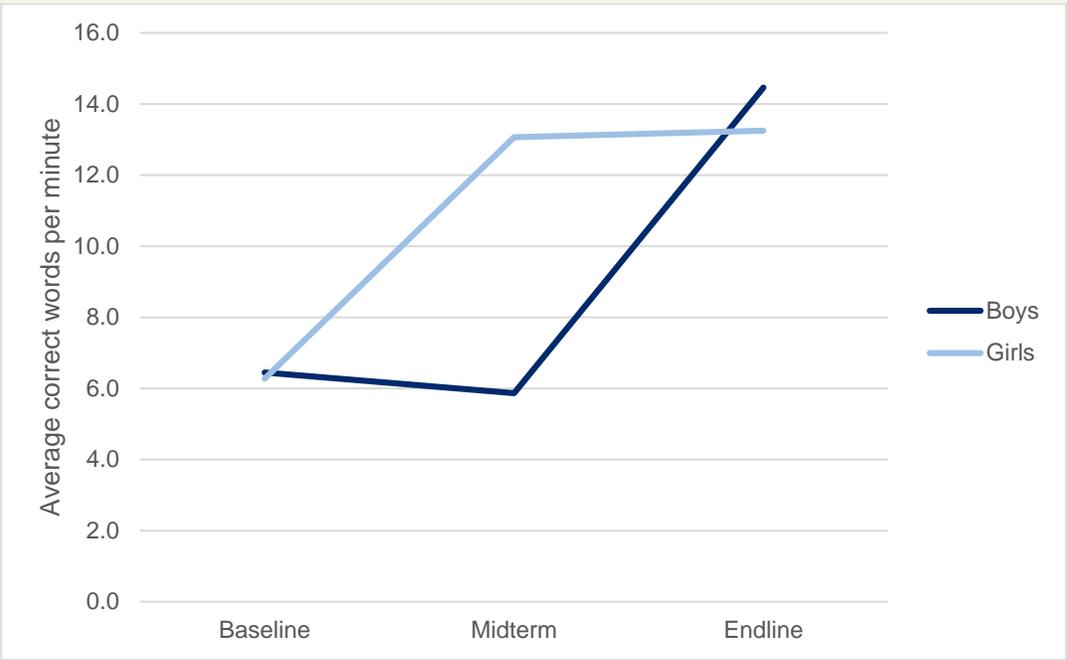
SUBTASK	GRADE	GENDER	BASELINE	MIDTERM	ENDLINE	ENDLINE DIFFERENCE (BOYS-GIRLS)	T-SCORE OF DIFFERENCE IN MEANS OF ENDLINE
Oral reading fluency (correct words per minute)	1	Boys	2.4	7.2	5.6	0.4	0.5
		Girls	12.9	5.3	5.2		
	2	Boys	6.4	5.9	14.5	1.2*	0.6
		Girls	6.3	13.1	13.2		
	3	Boys	18.2	21.0	30.0	8.0	2.0
		Girls	20.0	27.0	22.0		
Correct letters per minute	1	Boys	54.8	52.1	62.8	2.7	1.0
		Girls	59.4	55.3	60.1		
	2	Boys	65.6	47.1	81.2	1.5	0.6
		Girls	62.8	63.9	79.7		
	3	Boys	65.4	75.7	93.7	3.1*	1.1
		Girls	64.9	81.7	90.6		
Correct familiar words per minute	1	Boys	4.0	10.1	5.5	1.5	2.6
		Girls	6.7	7.9	4.0		
	2	Boys	8.0	8.5	11.1	0.4*	0.3
		Girls	4.7	17.3	10.6		
	3	Boys	11.3	25.1	20.4	6.1	2.5
		Girls	11.3	31.3	14.3		
Correct invented words per minute	1	Boys	0.4	0.7	0.7	-0.2	0.7
		Girls	0.5	0.7	0.9		
	2	Boys	0.8	0.7	3.8	0.8*	0.9
		Girls	1.2	1.8	3.0		
	3	Boys	1.7	5.1	6.9	3.3	2.5
		Girls	1.0	5.0	3.6		
Listening comprehension (percentage correct)	1	Boys	41%	46%	44%	2%	0.5
		Girls	32%	48%	42%		
	2	Boys	39%	48%	51%	-5%	1.1
		Girls	27%	55%	56%		
	3	Boys	51%	66%	64%	6%	1.9
		Girls	41%	79%	58%		
Reading comprehension score (percentage correct)	1	Boys	3%	13%	2%	-2%	2.0
		Girls	16%	7%	4%		
	2	Boys	10%	6%	10%	-1%	0.6
		Girls	5%	17%	11%		
	3	Boys	23%	27%	25%	5%	1.2
		Girls	15%	22%	20%		
	1	Boys	28%	29%	35%	1%	0.4

**Table D-1. Change in Boys' and Girls' Reading Scores, Cohort 2, by Subtask and Grade**

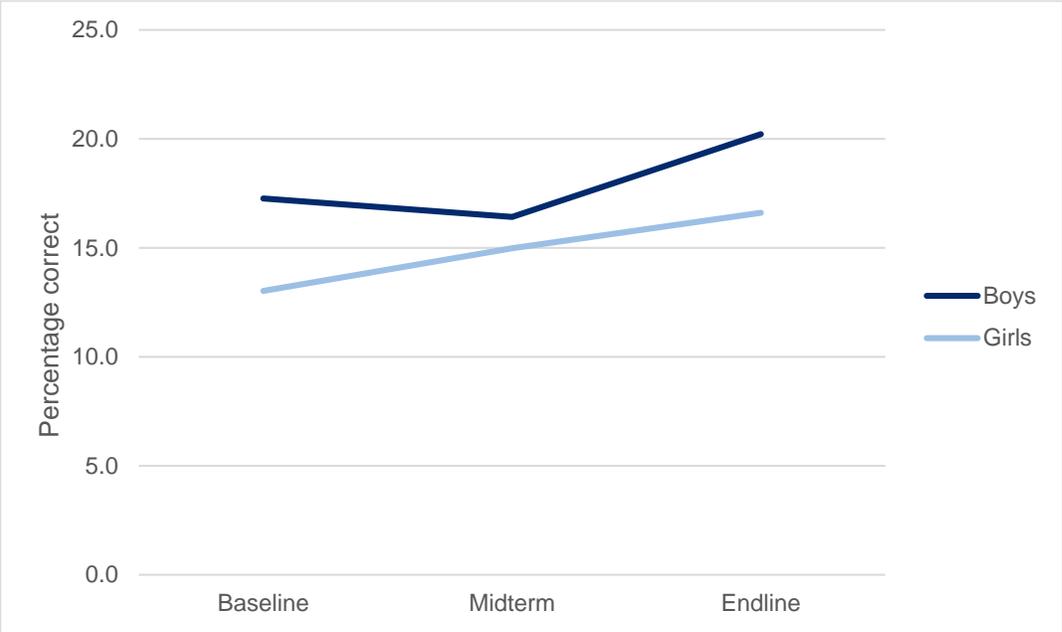
SUBTASK	GRADE	GENDER	BASELINE	MIDTERM	ENDLINE	ENDLINE DIFFERENCE (BOYS-GIRLS)	T-SCORE OF DIFFERENCE IN MEANS OF ENDLINE
Initial sound score (percentage correct)	2	Girls	9%	26%	34%	-1%	0.2
		Boys	36%	30%	42%		
	3	Girls	25%	25%	43%	1%**	0.3
		Boys	47%	44%	49%		
		Girls	38%	48%	48%		
		Boys	47%	44%	49%		
Correct addition problems per minute	1	Boys	7.4	7.4	5.5	1.8**	3.7
		Girls	8.0	5.0	3.7		
	2	Boys	13.5	8.6	9.7	2.1**	3.5
		Girls	8.6	6.5	7.6		
	3	Boys	12.3	12.0	11.6	2.7*	3.8
		Girls	10.5	11.0	8.9		
Correct subtraction problems per minute	1	Boys	5.7	4.8	3.3	1.0*	2.7
		Girls	6.7	3.8	2.2		
	2	Boys	8.5	5.2	6.7	1.3*	2.6
		Girls	7.3	3.9	5.5		
	3	Boys	10.9	8.4	8.2	1.9**	2.4
		Girls	8.4	7.4	6.3		
Correct number identification problems per minute	1	Boys	12.3	14.7	13.8	3.1*	3.6
		Girls	14.0	11.0	10.7		
	2	Boys	17.3	16.4	20.2	3.6**	2.5
		Girls	13.0	15.0	16.6		
	3	Boys	17.6	22.6	25.8	5.2*	4.6
		Girls	17.1	22.1	20.6		
Quantity comparison score (percentage correct)	1	Boys	24%	49%	50%	7%**	2.1
		Girls	46%	36%	44%		
	2	Boys	40%	42%	70%	13%**	3.8
		Girls	27%	46%	57%		
	3	Boys	54%	71%	76%	10%	3.9
		Girls	48%	72%	66%		
Word problems (percentage correct)	1	Boys	35%	43%	31%	5%*	1.7
		Girls	26%	37%	26%		
	2	Boys	53%	52%	47%	7%*	2.9
		Girls	36%	45%	40%		
	3	Boys	65%	63%	54%	8%**	2.9
		Girls	56%	66%	46%		

\*p < 0.05,  
\*\*p < 0.001

**Figure D-1. Gap in oral reading fluency scores, grade 2 in Cohort 2, by gender (correct words per minute).**



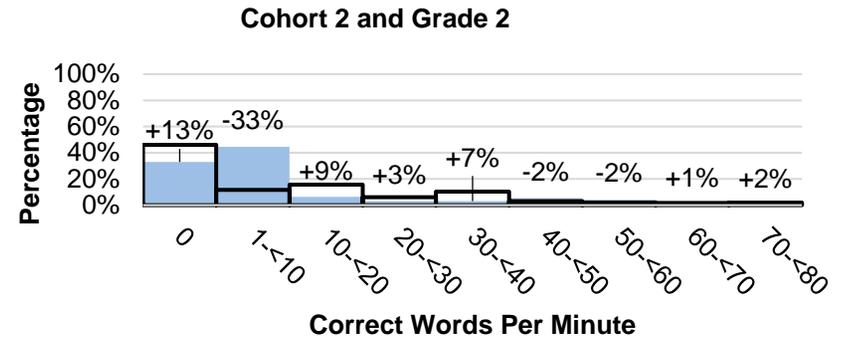
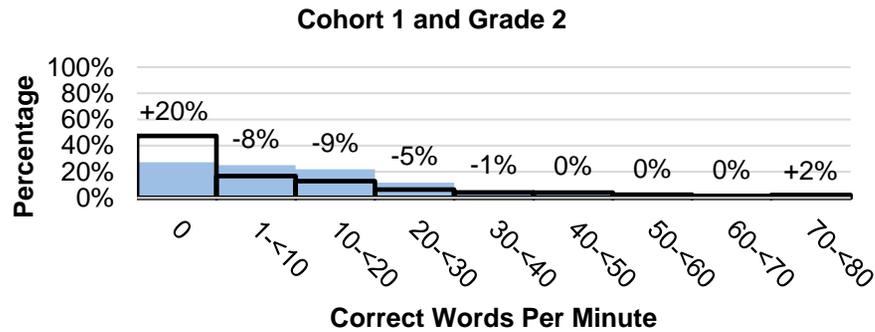
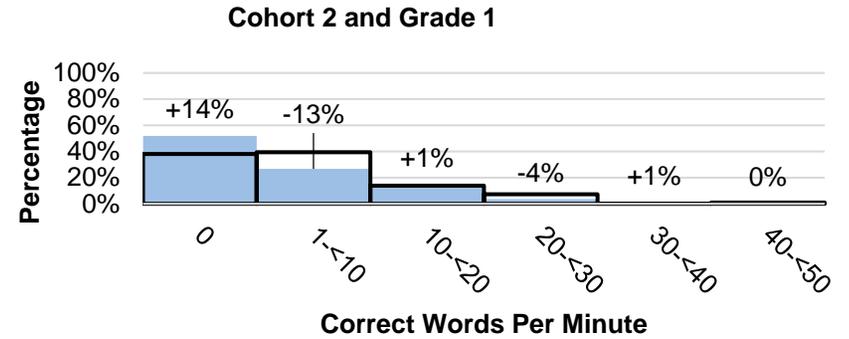
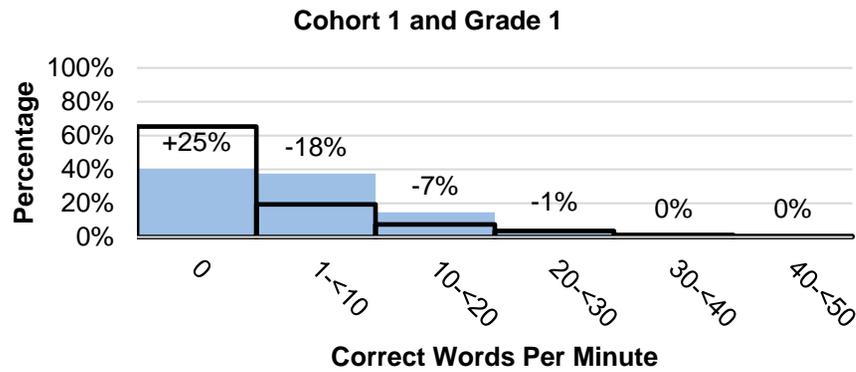
**Figure D-2. Gap in scores for number identification problems per minute, grade 2 in Cohort 2, by gender.**



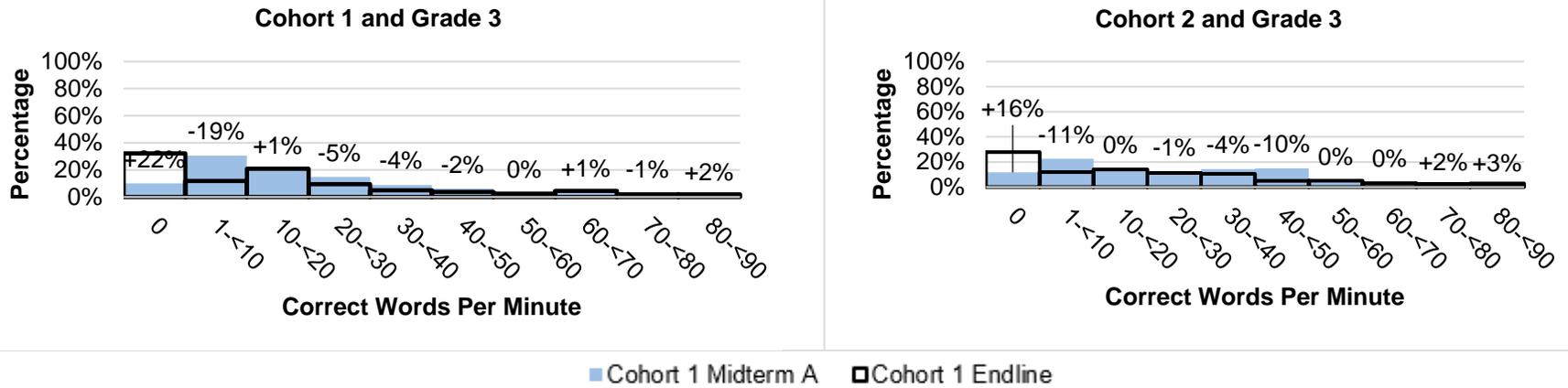
This page intentionally left blank.

# ANNEX E: OTHER STATISTICAL ANALYSES—DETAIL

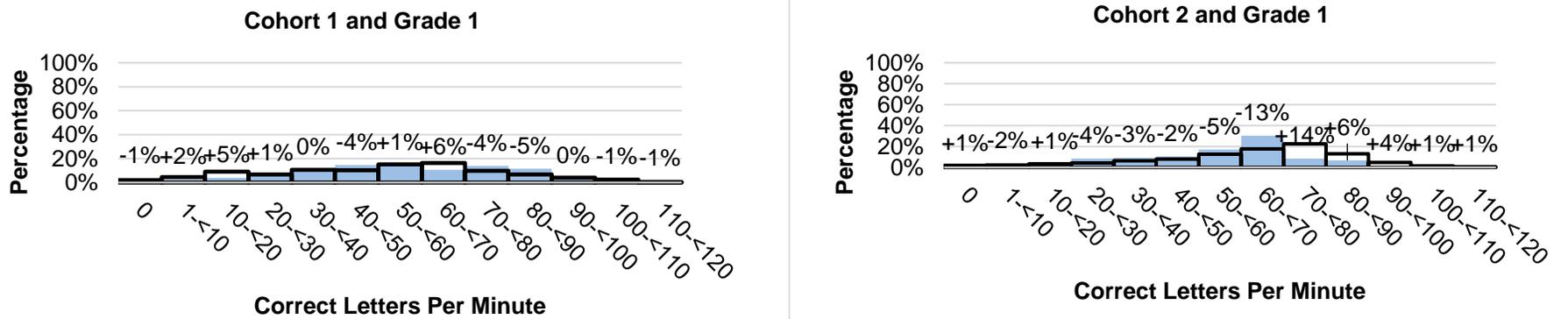
**Figure E-1. Change in distribution for oral reading fluency.**



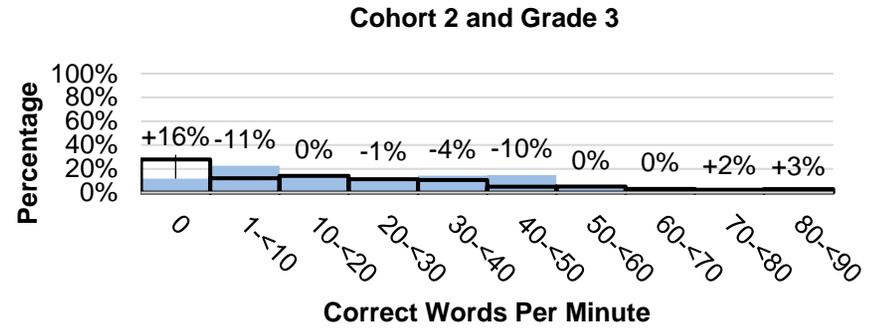
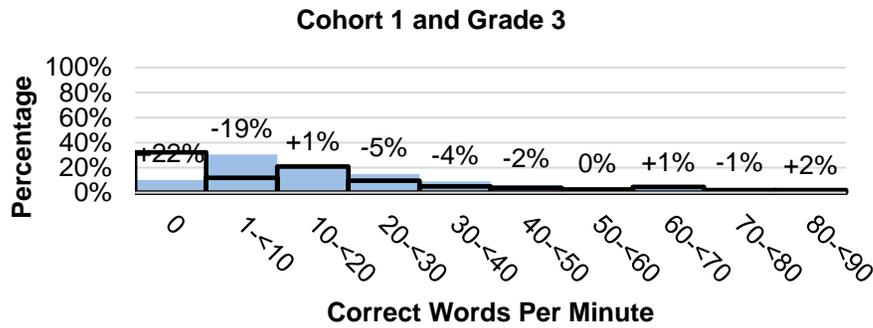
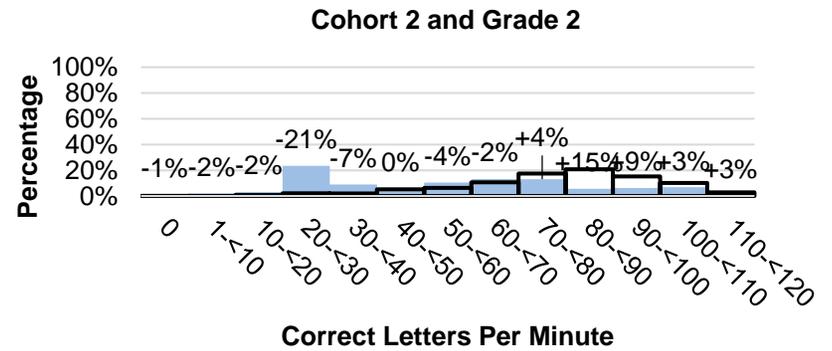
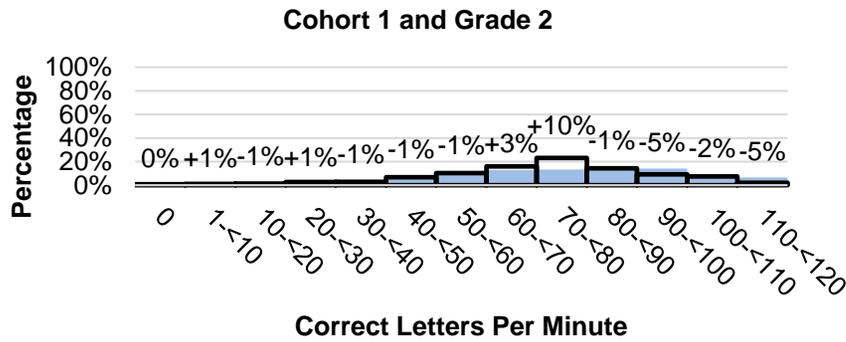
**Figure E-1. Change in distribution for oral reading fluency.**



**Figure E-2. Change in distribution for correct letters per minute.**

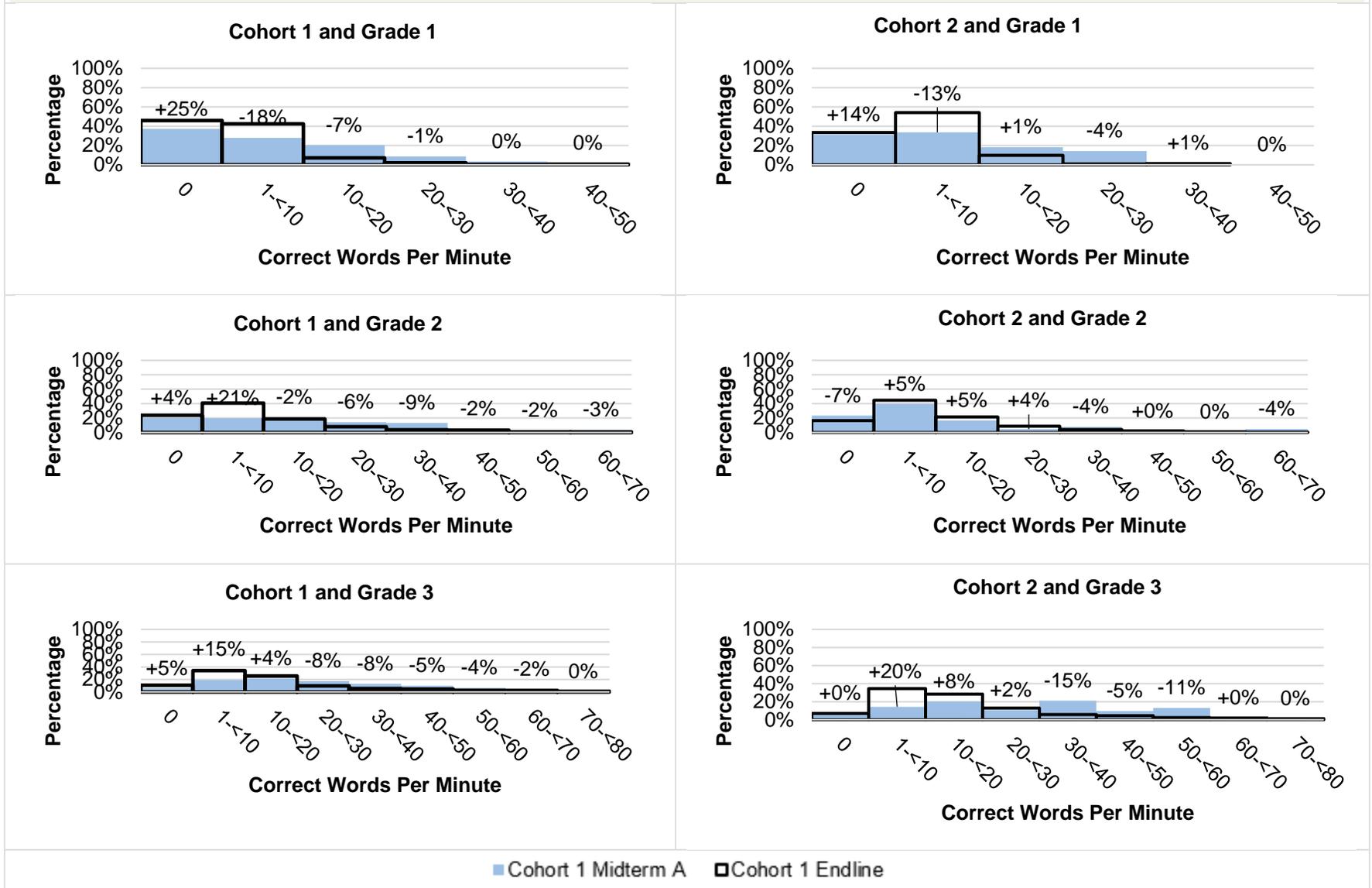


**Figure E-2. Change in distribution for correct letters per minute.**

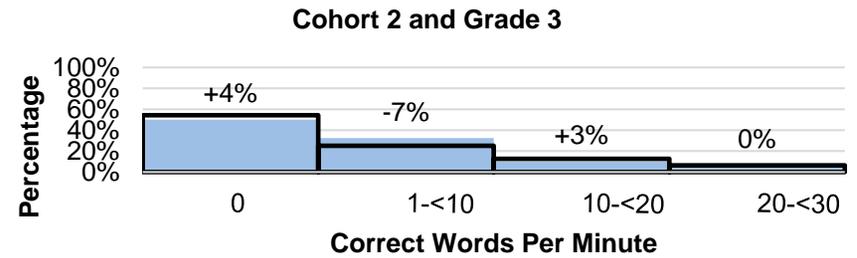
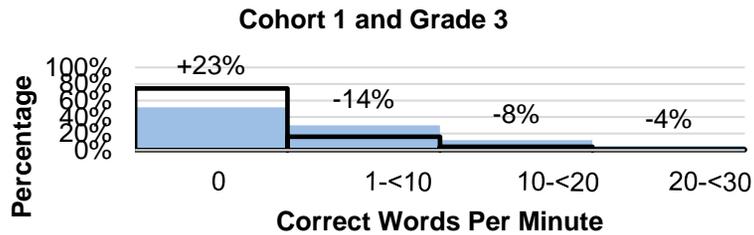
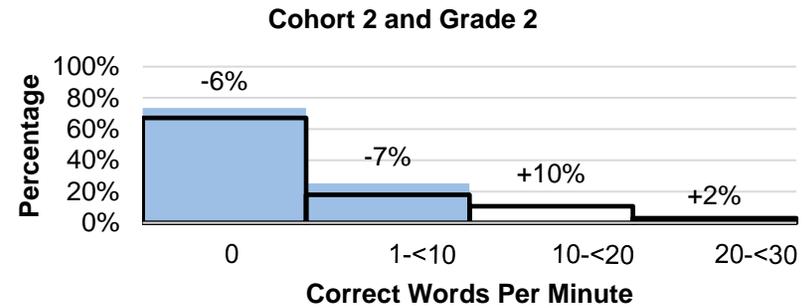
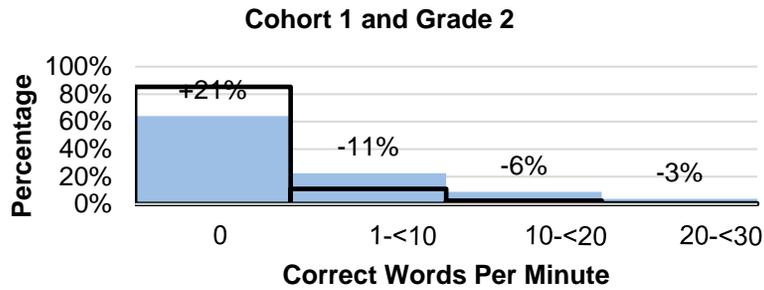
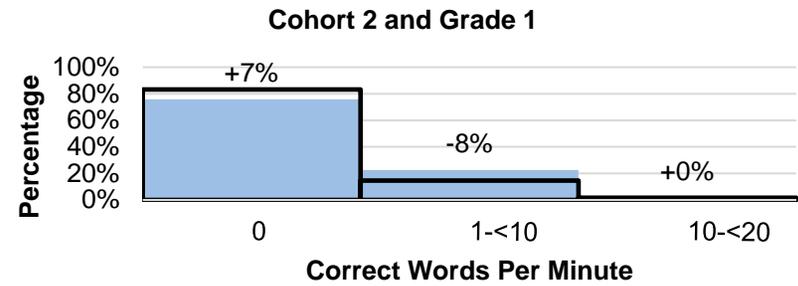
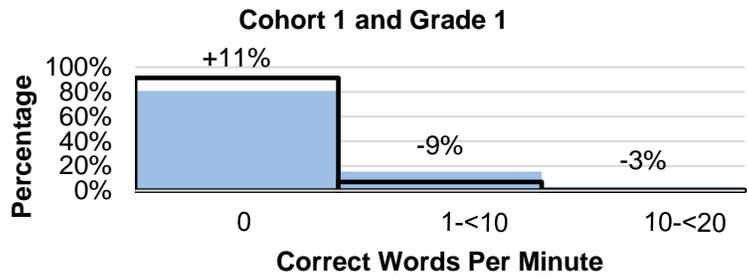


■ Cohort 1 Midterm A   ■ Cohort 1 Endline

**Figure E-3. Change in distribution for correct words per minute.**

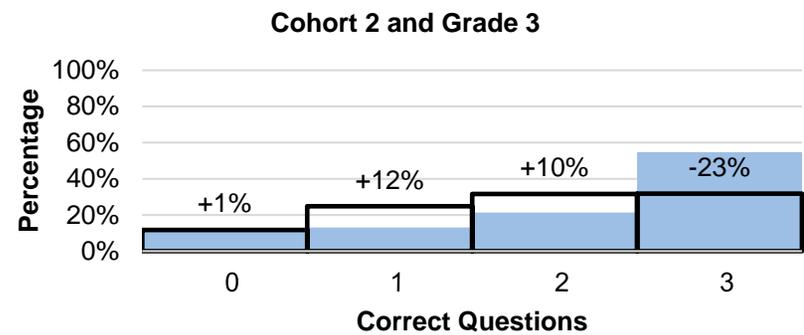
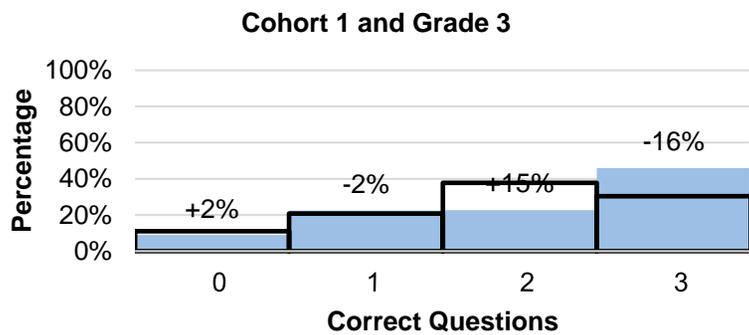
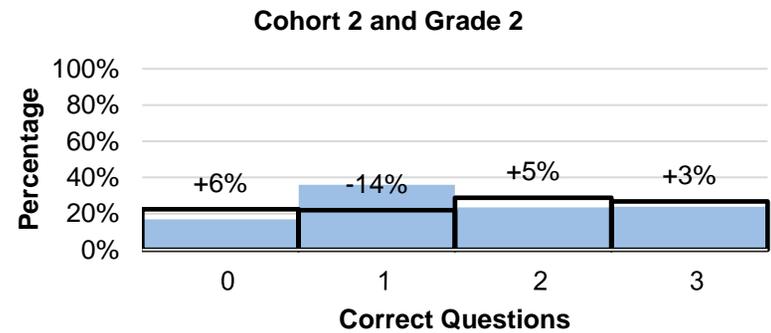
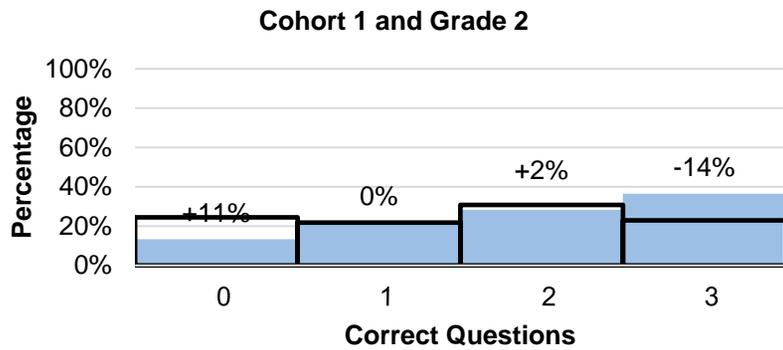
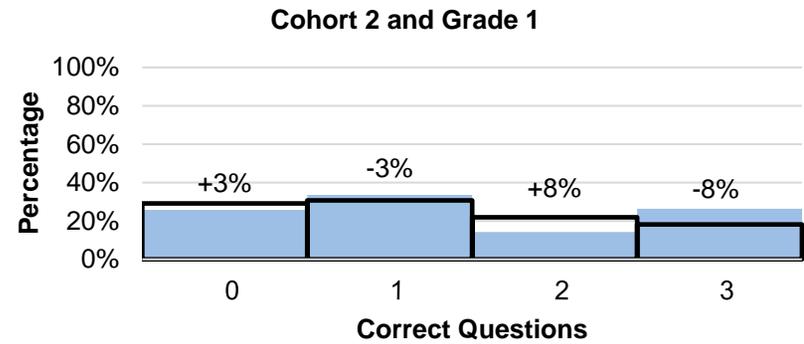
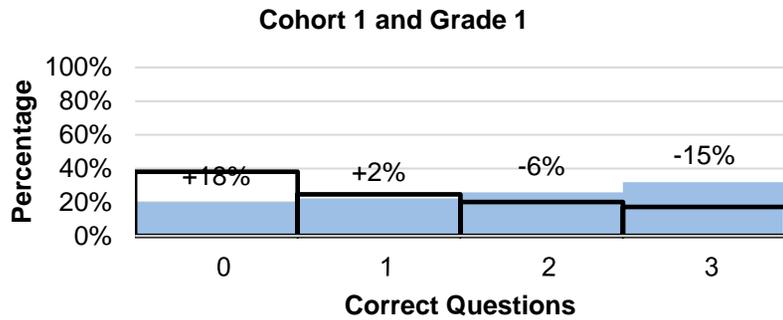


**Figure E-4. Change in distribution for correct invented words per minute.**



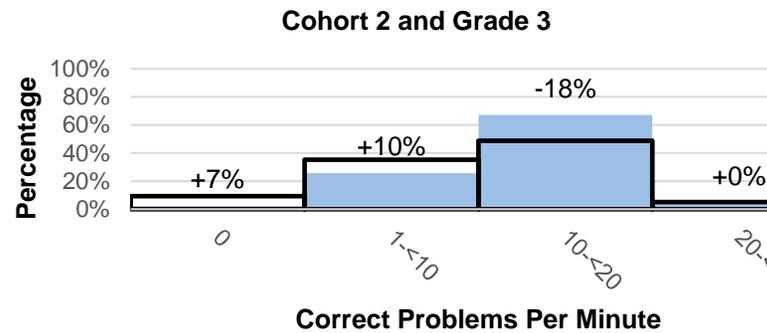
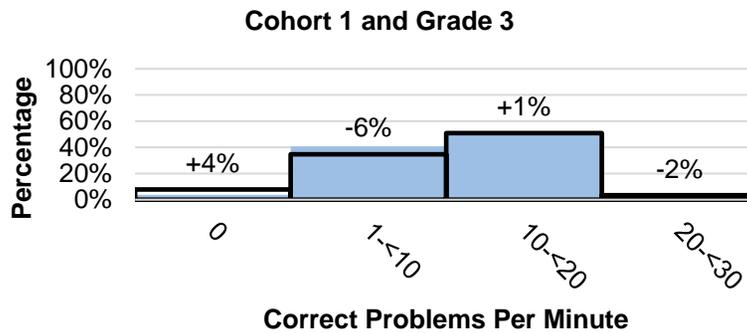
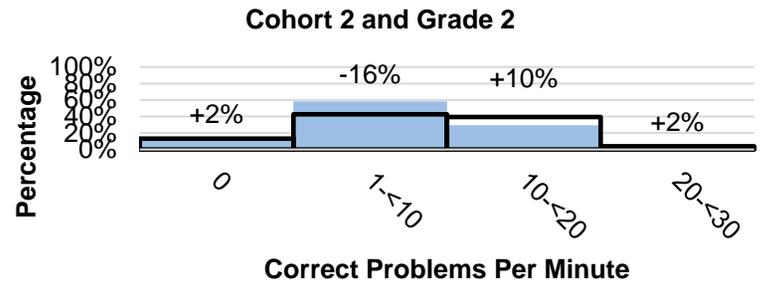
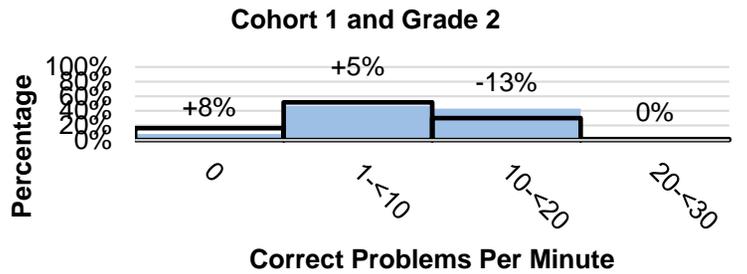
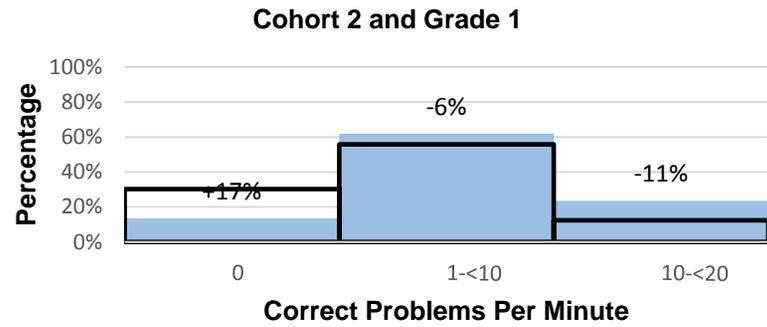
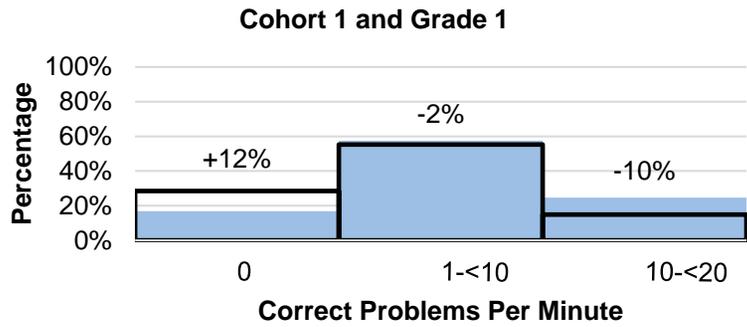
■ Cohort 1 Midterm A   ■ Cohort 1 Endline

**Figure E-5. Change in distribution for listening comprehension score.**



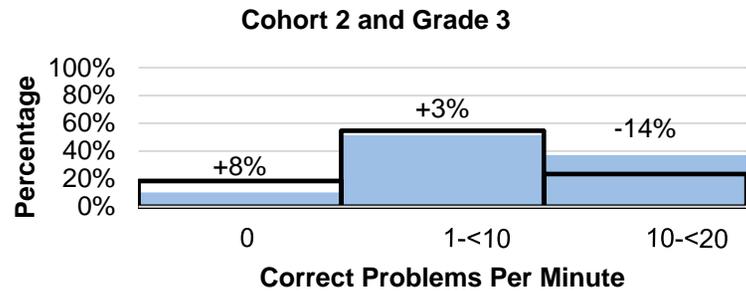
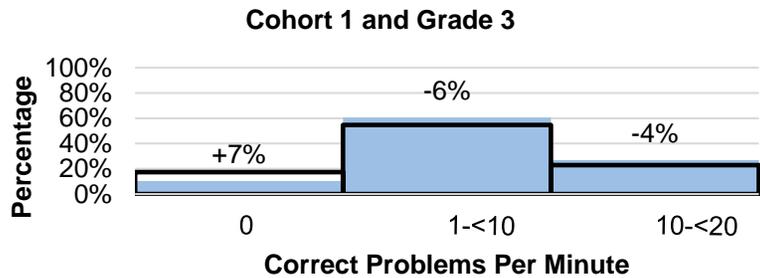
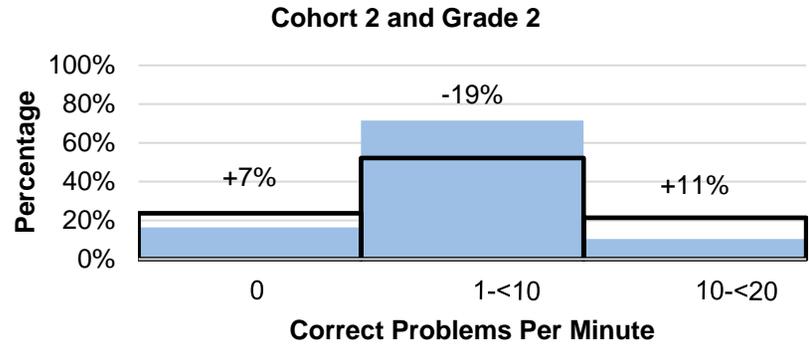
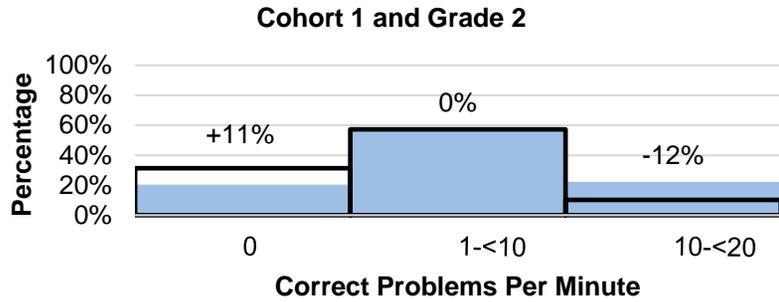
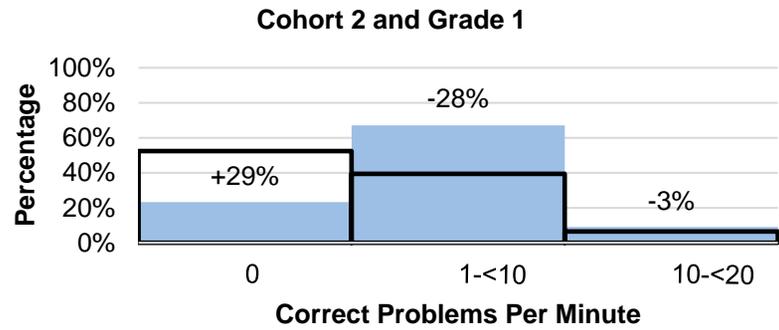
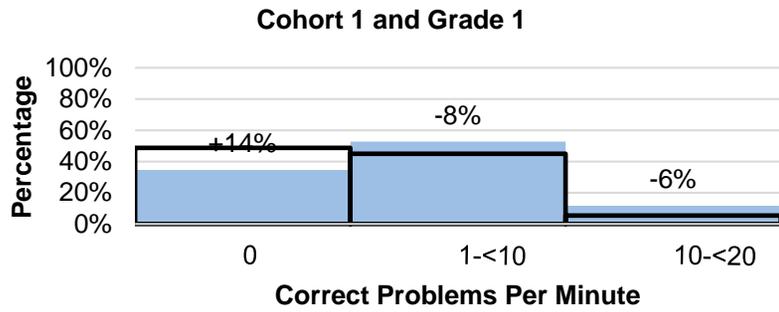
■ Cohort 1 Midterm A    □ Cohort 1 Endline

**Figure E-6. Change in distribution for correct addition problems per minute.**



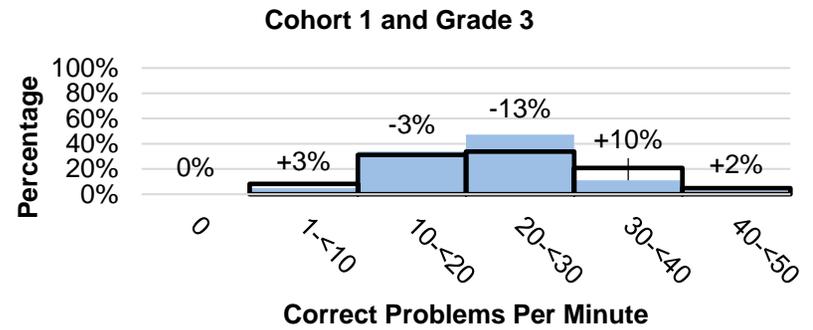
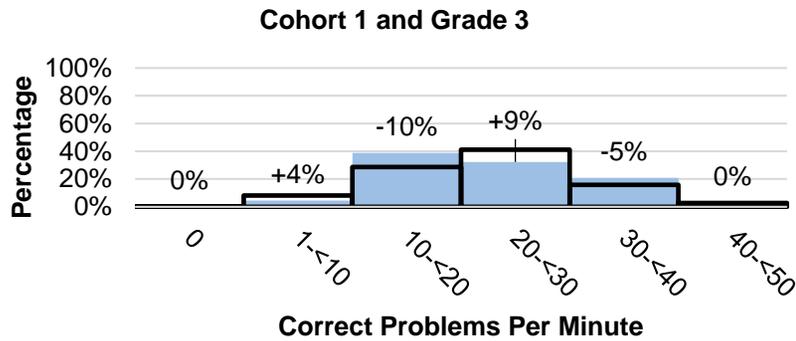
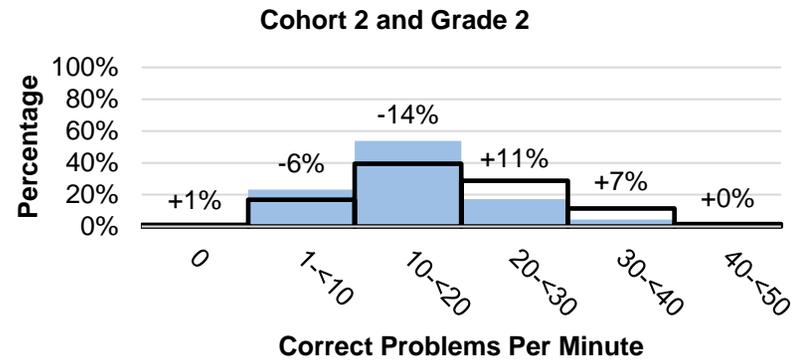
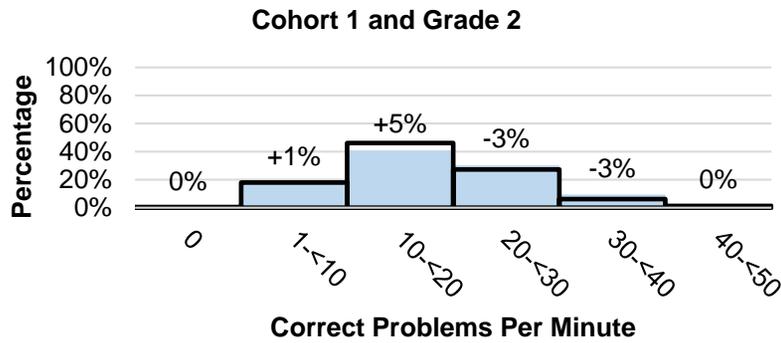
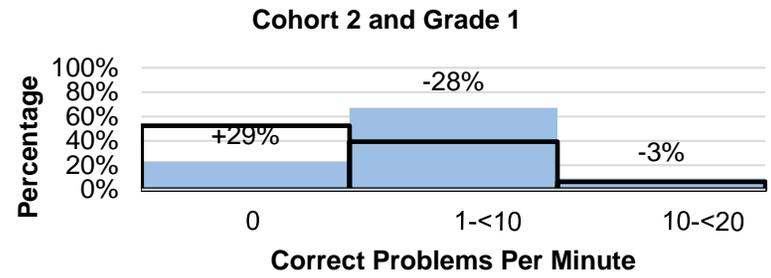
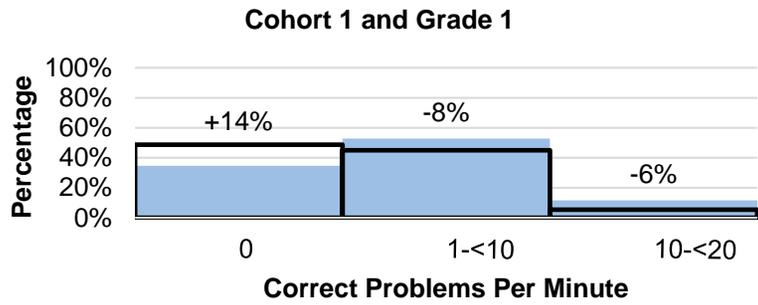
■ Cohort 1 Midterm A    □ Cohort 1 Endline

**Figure E-7. Change in distribution for correct subtraction problems per minute.**



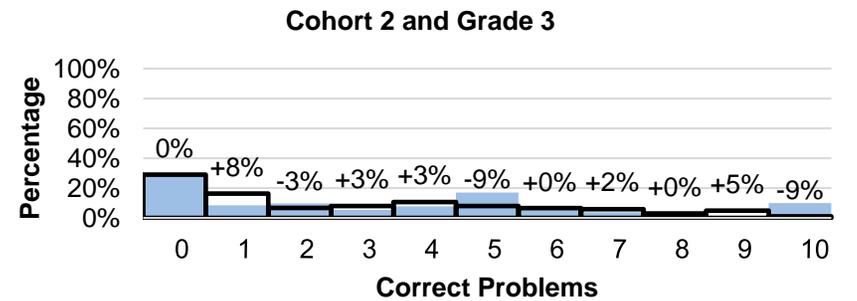
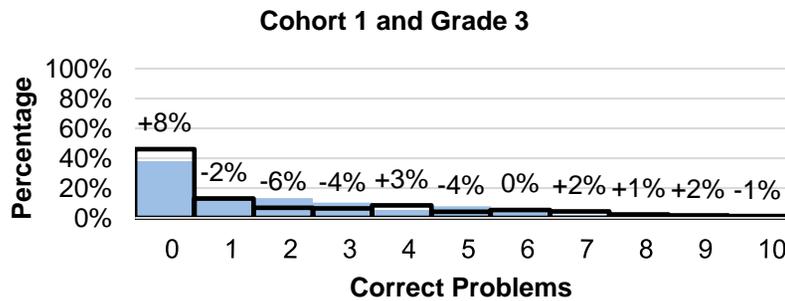
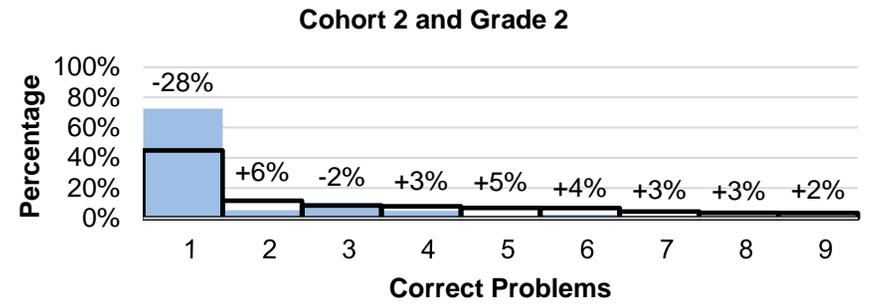
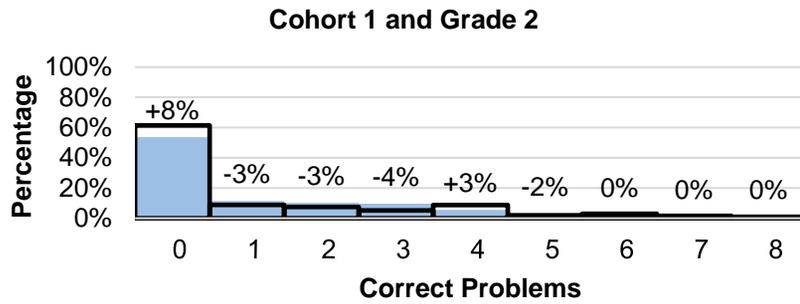
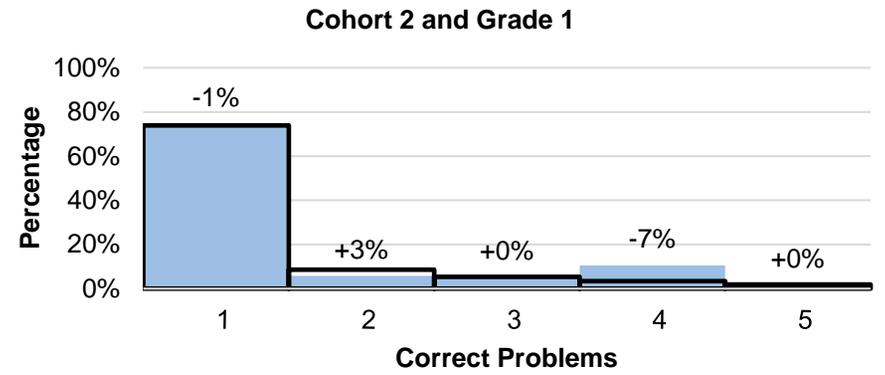
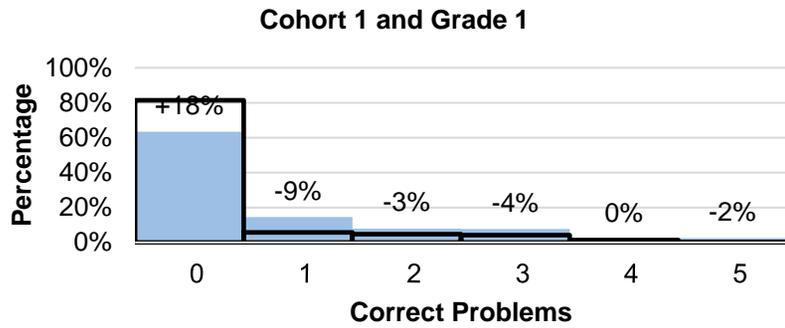
■ Cohort 1 Midterm A    □ Cohort 1 Endline

**Figure E-8. Change in distribution for correct number identification problems per minute.**



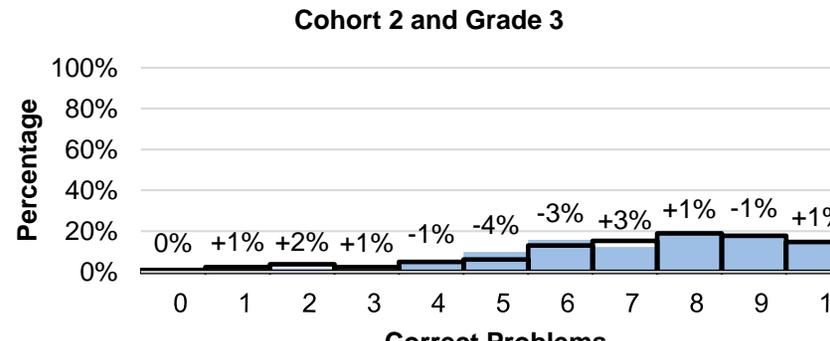
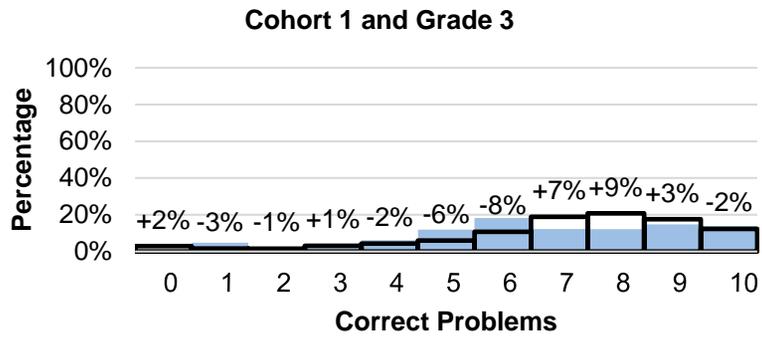
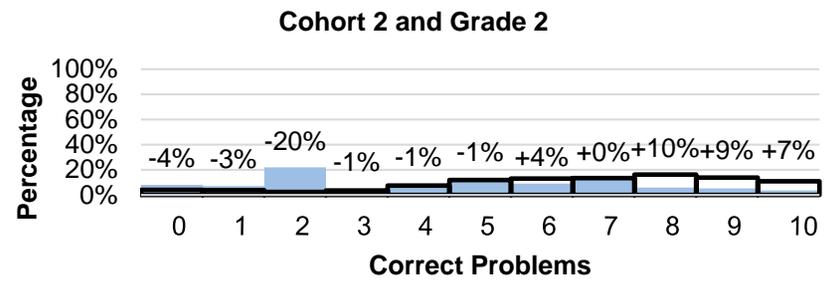
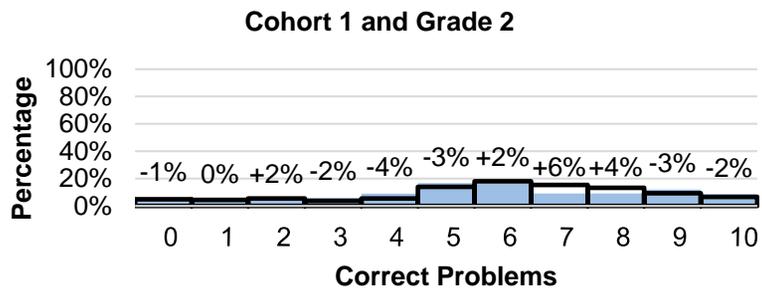
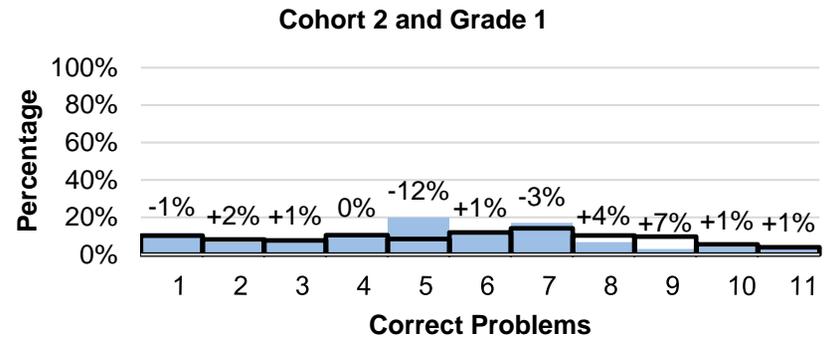
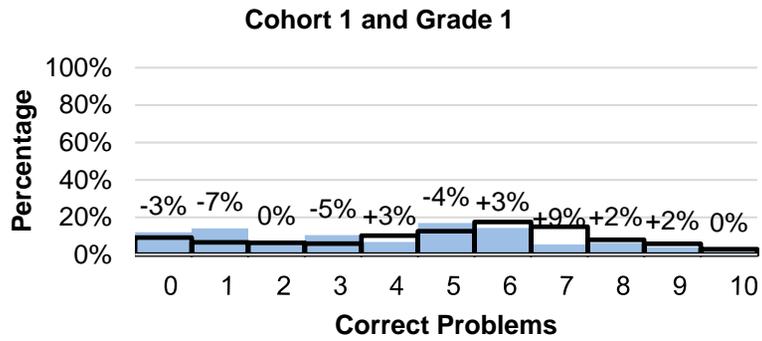
■ Cohort 1 Midterm A    □ Cohort 1 Endline

**Figure E-9. Change in distribution for missing number score.**



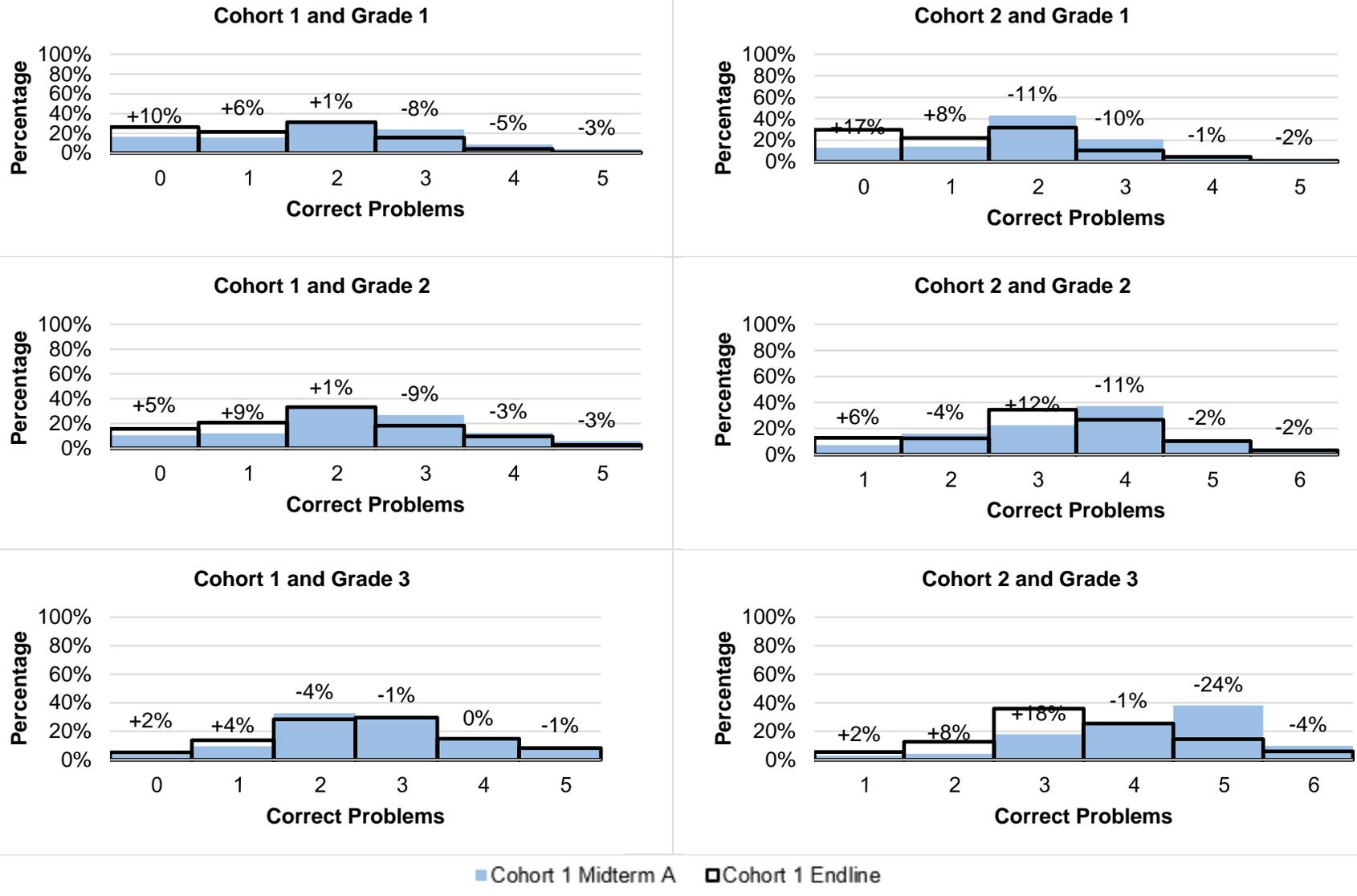
■ Cohort 1 Midterm A   ■ Cohort 1 Endline

**Figure E-10. Change in distribution for quantitative comparison score.**



■ Cohort 1 Midterm A   ■ Cohort 1 Endline

**Figure E-11. Change in distribution for word problem score.**



This page intentionally left blank.

# ANNEX F: TECHNICAL DETAILS ON INSTRUMENT RELIABILITY AND VALIDITY TESTING

The research team conducted an assessment of internal consistency to evaluate reliability. Internal consistency is an appropriate and standard classical evaluation approach for cross-sectional data, and when combined with item-level evaluative psychometric methods, provides insight regarding item and/or subtask functioning. Internal consistency (Cronbach, 1951) is the average correlation of all possible half-scale divisions and is frequently provided in published assessment psychometrics. The range of the internal consistency statistic is from zero to one, where higher values are desired and a value of zero indicates inconsistency of measurement. As a general guideline, Cronbach's alpha should be at least 0.70 for adequacy, and coefficients closer to 1 indicate a good assessment (Aron, Aron, & Coups, 2013).

The Cronbach's alpha coefficient was computed by using the Stata analytics software, which produced **Tables F-1 and F-2**, separated by reading and mathematics subtasks. The first two columns of the tables provide general subtask information, including the subtask name and the number of pupils accounted for within the subtask. The next three columns of the tables provide interrelationship information, including item-test correlations (i.e., the correlation between a subtask and the entire scale), item-rest correlations (i.e., the correlation between a subtask and the scale that is formed by all other subtasks), and the Cronbach's alpha (previously discussed). Overall, the subtask scores show good reliability statistics (Cronbach's alpha of at least 0.80), with alpha score of 0.84 for reading subtasks and an alpha score of 0.85 for mathematics subtasks.

**Table F-1. Cronbach's Alpha Coefficients for Reading Subtasks**

READING SUBTASKS	NUMBER OF PUPILS	ITEM-TEST CORRELATION	ITEM-REST CORRELATION	CRONBACH'S ALPHA
Initial sound score (percentage correct)	3,738	0.5935	0.3864	0.8624
Letter sound score (percentage correct)	3,792	0.702	0.5551	0.826
Familiar word score (percentage correct)	3,779	0.9017	0.8419	0.765
Invented word score (percentage correct)	3,778	0.5684	0.5047	0.8468
Oral reading score (percentage correct)	3,774	0.9023	0.8241	0.7642
Reading comprehension score (percentage correct)	3,774	0.8388	0.7525	0.7857
			<b>Alpha</b>	<b>0.8392</b>

**Table F-2. Cronbach's Alpha Coefficients for Mathematics Subtasks**

MATHEMATICS SUBTASKS	NUMBER OF PUPILS	ITEM-TEST CORRELATION	ITEM-REST CORRELATION	CRONBACH'S ALPHA
Number identification (percentage correct)	3,774	0.8359	0.7295	0.81
Number discrimination (percentage correct)	3,771	0.7539	0.6041	0.8385
Missing number (percentage correct)	3,766	0.6647	0.5155	0.8506
Word problem (percentage correct)	3,757	0.766	0.627	0.8328
Addition problems (percentage correct)	3,762	0.8328	0.7648	0.8129
Subtraction problems (percentage correct)	3,758	0.7856	0.715	0.826
			<b>Alpha</b>	<b>0.8531</b>

## REFERENCES

Aron, A., Aron, E. N., & Coups, E. (2013). *Statistics for psychology (6th ed.)*. Boston, Massachusetts, USA: Pearson.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.

# ANNEX G: SUMMARY OF BENCHMARKING EXERCISE IN LIBERIA



# Proposing Benchmarks for Early Grade Reading Skills in Liberia

## BACKGROUND

From 2008, beginning with EGRA+ and continuing under the Liberia Teacher Training Program (LTTP), USAID has supported improved early grade reading in Liberia. Since its inception, the LTTP supported intervention to improve early grade reading is being implemented in 1,200 schools. The most recent analysis of that program’s impact (based on a reading assessment conducted towards the end of the 2012-13 school year) showed that, on average, children in grades 1 through 3 in participating schools significantly improved their achievement in all the reading skills tested. From 2011 to 2013, grade 1 students on average almost tripled their ability to read familiar words, more than tripled their decoding skills, and were able to read text more than one-and-a-half times more fluently. Grade 2 and 3 students on average were 10 times better at decoding and had almost three times higher oral reading fluency than before the program. On average, students in all grades improved their listening and reading comprehension.

However, even grade 3 students were still not on average attaining a level of reading skill adequate to ensure full comprehension of what they read. Liberian student performance in reading remains far below U.S. or other developed country standards. This begs the question, “What is an acceptable level of reading achievement for Liberian students in the early grades of primary school?”

With additional assistance from USAID a workshop to address this question took place during the first week of March, 2014.

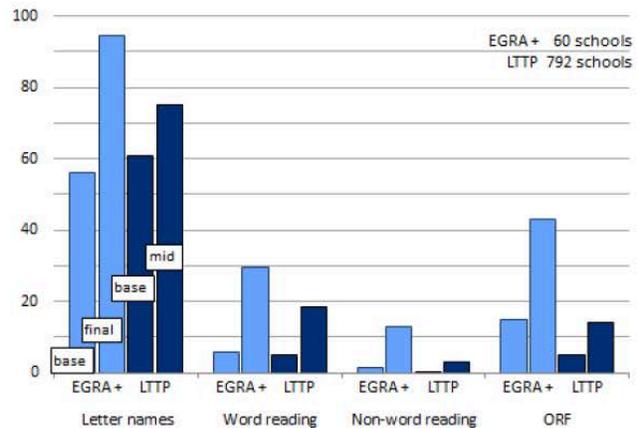
Ministry of Education officials, district education officers, and a cross section of stakeholders attended this workshop over two days to begin

Liberia’s first ever effort to define standards for student performance in key areas of reading skill development in grades 1, 2 and 3.

## LIBERIA IS UNIQUELY POSITIONED TO UNDERTAKE AN INFORMED PROCESS OF SETTING READING BENCHMARKS

Only a handful of developing countries have taken on the challenge of setting benchmarks for reading skills in early grades. Mexico did so several years ago. And more recently Kenya and Egypt have defined benchmarks, with Kenya officially adopting a standard for oral reading fluency in both English and Kiswahili. Liberia has the distinct advantage of having a bounty of data to inform the setting of reading skill benchmarks.

Grade 2 Performance and Improvement Over Time in EGRAPlus And LTTP Schools



As shown in the graph above, the EGRA+ and LTTP programs have baseline and subsequent measures of reading performance that show not just how well students perform in different skill areas, but also how much improvement can be achieved through a targeted instructional intervention. This provides a realistic foundation from which to discuss what benchmarks may be most appropriate for the current Liberian context.

## THE BENCHMARK SETTING WORKSHOP

A two day workshop on March 6 and 7, 2014 brought together 60 MOE officials, district education officers, donor agency representatives, NGOs active in the education sector, LTTP project staff, and outside experts, to begin a process of defining benchmarks for specific skill areas of early grade reading. The objectives of the workshop were to:

- Share the most recent assessment results from LTTP's reading intervention
- Orient and engage a cross section of Liberian stakeholders in a participatory process of setting reading benchmarks for grades 1, 2 and 3.

During the first morning of the workshop, data from the LTTP midterm assessment and from the EGRA+ end line assessment were shared and discussed. Data from international (PIRLS 2011) and from U.S. assessments of reading were also shared. In addition to providing points of comparison, these data helped illustrate that reducing the percentages of students scoring at the lowest levels is a key strategy for improving a country's overall performance.

Following the presentation, participants were engaged in a discussion of benchmarks – what they are and how to set them by combining empirical data both from other countries and from Liberia, working knowledge of Liberia's education sector, and common sense. Overall, the ratio of presentation to participatory work was about 1 to 2.5.

Small working groups took on the challenge of analyzing the available information, discussing and debating what seemed possible, and then defining an initial set of benchmarks for grade 3.

Those results were shared and discussed, prior to moving on to setting benchmarks for the other grades.



At the end of the workshop, all the groups' points of view were recorded and areas of convergence and divergence in recommended benchmarks were identified and discussed so as to generate further convergence. This report shares those results, showing the full range of points of view advocated by the participants, and concluding with the recommendations of the authors as to what may be useful "converged" benchmarks for early grade reading in Liberia at the present time.

## THE READING SUBTASKS

The policy workshop helped define benchmarks for three reading subtasks evaluated using the Early Grade Reading Assessment (EGRA) in grades 1, 2 and 3. The three EGRA subtasks include:

- **Non-word fluency.** This subtask evaluates a student's ability to decode unfamiliar words. Short combinations of three letters (often consonant, vowel, consonant) that do not form words (e.g., "nak") are used so that the assessment can distinguish the skill of decoding from the skill of whole word reading. The subtask is timed, so the resulting measure is the number of non-words decoded correctly per minute.
- **Oral reading fluency.** This subtask evaluates how well a child reads out loud a coherent, short passage of text. It is also timed, and therefore produces a measure that is the number of words of text correctly read per minute.
- **Reading comprehension.** Students are asked five questions relating to the text which they would have read aloud for the oral reading fluency portion of the assessment. The resulting measure is a number or percent of correct responses out of five.

## THE BENCHMARK SETTING PROCESS

Working in seven separate small groups, participants received data tables showing how students performed on these subtasks at the start and end of EGRA+ (2008 and 2010) and at the start and midpoint of LTTP (2011 and 2013).

Data that expressed the relationship between these subtasks were also shared. For example, a scatter plot of oral reading fluency and comprehension showed that students who demonstrated comprehension at 80% or better (answering 4 out of 5 questions correctly) were for the most part

reading with oral fluency of between 45 and 65 words per minute. Similar data were used to demonstrate the relationship between students' decoding abilities (as measured by non-word reading) and their levels of oral reading fluency. In addition, some international data were shared. These data helped participants to use a benchmark set in one area – say comprehension – to define the benchmarks in the other skill areas.

Each group, armed with these data and their own vast working knowledge of the education system in Liberia, was asked to:

- For each subtask, define three aspects:
  - The benchmark value for the indicator for that subtask,
  - The percentage of students that would be meeting that benchmark in five years, and
  - The percentage of students who would be scoring zero on that indicator in five years.
- Define the above values for grade 3
- For grade 3, define first the values for reading comprehension (benchmark, percentage meeting the benchmark and percentage scoring zero) and then use that to inform the "needed" values for the two other skill areas.
- Having completed this work for all three subtasks (comprehension, oral reading fluency and non-word reading) for grade 3, the groups reconvened in plenary to compare, discuss and arbitrate among their responses.
- Following the plenary discussion, the groups were charged with first revisiting what they had proposed for grade 3, then, using the grade 3 values they decided on, define the levels that would propose for each of the three subtasks for grade 2 and then grade 1.
- Group work was interspersed with short plenaries to clarify concepts and document convergence.

### THE BENCHMARKING RESULTS

The table below summarizes the results of the group work defining benchmarks for all three subtasks for grade 3. Each subtask skill is shown in a column in the table. In addition to the output from the groups,

relevant data from the LTTP midterm is shown as a point of comparison to the standards proposed by the workshop participants.

<b>Grade 3</b>	Reading Comprehension <sup>1</sup>	Oral Reading Fluency <sup>2</sup>	Non-word Reading/decoding <sup>3</sup>
<i>Average performance at LTTP midterm</i>	22%	20.2	4.5
<b>Proposed benchmark</b>	<b>60-80%</b>	<b>45-50</b>	<b>10-20</b>
<i>% of students meeting proposed benchmark (LTTP midterm)</i>	8%	17%	18%
<b>Proposed % of students at benchmark in 5 years</b>	<b>45-50%</b>	<b>40-50%</b>	<b>30-40%</b>
<i>% of students scoring zero (LTTP midterm)</i>	44%	10%	52%
<b>Proposed % of students scoring zero in 5 years</b>	<b>20-30%</b>	<b>5-15%</b>	<b>25-35%</b>

<sup>1</sup> % of questions answered correctly; <sup>2</sup> correct words per minute;

<sup>3</sup> correct non-words per minute

Concerning reading comprehension, the workshop participants discussed a benchmark target reading score of between 60% and 80% correct. This compares to the average performance at LTTP midterm of 22%. The group in general wanted to set a standard well above what grade 3 students are scoring now, reasoning that the standard should reflect a decent level of comprehension of grade appropriate text by the end of grade 3. They were also informed by what EGRA+ was able to achieve, while conscious that a pilot project is quite different from a national scale-up.

While setting the benchmark somewhat high, the group was more modest in their estimation of the percentage of students who would be at that benchmark in 5 years' time. The groups converged around 40-50% of students being able to meet the benchmark of 60 to 80 percent comprehension. Compared to LTTP midterm –only 8% of grade 3 students achieved 60% or better – the target of 40-50% would represent significant improvement. The same could be said for zero scores – the groups all proposed a 5 year target of fewer students scoring zero in reading comprehension than did so on the LTTP midterm.

Results for grades 2 and 1 are presented below.

<b>Grade 2</b>	Reading Comprehension <sup>1</sup>	Oral Reading Fluency <sup>2</sup>	Non-word Reading / decoding <sup>3</sup>
<i>Average performance at LTTP midterm</i>	14%	14.2	3.2
<b>Proposed benchmark</b>	<b>40-60%</b>	<b>35-40</b>	<b>10-15</b>

<i>% of students meeting proposed benchmark (LTTP midterm)</i>	10%	19%	23%
<b>Proposed % of students at benchmark in 5 years</b>	<b>40-50%</b>	<b>40-50%</b>	<b>25-40%</b>

<i>% of students scoring zero (LTTP midterm)</i>	59%	27%	64%
<b>Proposed % of students scoring zero in 5 years</b>	<b>25-30%</b>	<b>15-20%</b>	<b>25-40%</b>

<sup>1</sup> % of questions answered correctly; <sup>2</sup> correct words per minute; <sup>3</sup> correct non-words per minute

<b>Grade 1</b>	Reading Comprehension <sup>1</sup>	Oral Reading Fluency <sup>2</sup>	Non-word Reading / decoding <sup>3</sup>
<i>Average performance at LTTP midterm</i>	22%	20.2	4.5
<b>Proposed benchmark</b>	<b>40-60%</b>	<b>30-40</b>	<b>5-10</b>

<i>% of students meeting proposed benchmark (LTTP midterm)</i>	3%	6%	7%
<b>Proposed % of students at benchmark in 5 years</b>	<b>30-45%</b>	<b>30-50%</b>	<b>25-40%</b>

<i>% of students scoring zero (LTTP midterm)</i>	44%	10%	52%
<b>Proposed % of students scoring zero in 5 years</b>	<b>30-50%</b>	<b>15-35%</b>	<b>25-50%</b>

<sup>1</sup> % of questions answered correctly; <sup>2</sup> correct words per minute; <sup>3</sup> correct non-words per minute

As was the case for grade 3, the workshop participants proposed benchmark targets that surpass the levels of performance seen on the LTTP midterm. The groups set benchmarks above the LTTP midterm averages in each skill area and proposed more students meeting those benchmark levels of performance than had previously did so. Also, the standards reflect a progression of increasing levels of achievement from grade 1, to grade 2, to grade 3.

During the benchmarking process, there was much lively discussion and debate about how much improvement over the present levels of performance one could expect to see. Groups vacillated between being ambitious and setting standards well above current levels of achievement and being realistic, if not pessimistic.



Often the question was raised as to what the groups should assume the MOE and its partners would be doing during the next five years to improve reading instruction. This led to considerable debate about whether the education sector in Liberia had sufficient resources, capacity and know-how to bring about dramatic improvements in reading performance. The encouraging fact was that the LTTP project (and before it EGRA+) had demonstrated that it is possible to improve reading outcomes in Liberia. Whether concerted effort can be continued and in fact broadened to address the needs in all schools across the country is the paramount concern. The proposed benchmarks have to assume that concerted effort will be made—otherwise, in a sense, there is no point in setting benchmarks. In fact, non-achievement of the benchmarks would be a tell-tale sign that not enough resources and effort are being mobilized

## CONCLUSIONS AND RECOMMENDATIONS

After careful consideration of the work produced by the participants in the Benchmarking Workshop, the following table summarizes what we would recommend to the MOE, its partners and stakeholders as standards for reading performance in the three skill areas across the three grades.

<b>Recommended Benchmarks</b>	Reading Comprehension <sup>1</sup>	Oral Reading Fluency <sup>2</sup>	Non-word Reading / decoding <sup>3</sup>
Grade 3 Benchmark	75%	50	15
Grade 2 Benchmark	50%	35	10
Grade 1 Benchmark	40%	30	8
Grade 3, % at benchmark in 5 years	50%	50%	50%
Grade 2, % at benchmark in 5 years	50%	50%	50%
Grade 1, % at benchmark in 5 years	50%	50%	50%
Grade 3, % scoring zero	15%	5%	15%
Grade 2, % scoring zero	20%	15%	20%
Grade 1, % scoring zero	20%	15%	25%

<sup>1</sup> % of questions answered correctly; <sup>2</sup> correct words per minute; <sup>3</sup> correct non-words per minute

Note that all benchmarks and indicators are proposed based on continuing to use a single assessment aligned to grade 2 in all three grades. The approach of using a common assessment across the three grades allows us to easily evaluate differences in reading performance from grade 1 to grade 2 to grade 3. This is what we recommend continue to be the approach in Liberia for monitoring of progress, fully recognizing that at the classroom level the expectation would be that teachers and students are working with grade appropriate materials and that teachers would evaluate their students accordingly. It is at a system level that it makes sense to monitor progress (for the time being) against a fix, single grade level of material.

The recommended benchmark levels of performance for reading comprehension, oral reading fluency and non-word reading for grade 3 are at the lower end of the ranges proposed by the working group participants. We are recommending the less ambitious benchmark because current achievement is so low and because we prefer being more ambitious regarding the other two indicators – the percentage of students meeting the benchmark and the percent scoring zero in five years.

Beginning with comprehension, we reasoned that by the end of grade 3, students should be attaining a reasonable level of comprehension. Therefore they

should be getting at least 75% of comprehension questions correct. The level of oral reading fluency that is associated with 75% comprehension in the Liberia data is 45 to 65 words per minute, so a standard of 50 wpm seems appropriate for assuring the desired level of comprehension. In a similar manner, the benchmark for non-word reading can be defined - the level of decoding skill students need to be reading with fluency approaching 50 words per minute.

Where the recommendations diverge from what was put forth in the workshop is in the standards for the percentage of students meeting the benchmark and in the percent scoring zero. We recommend a somewhat higher percentage of students meeting the benchmark levels and foresee that as being consistent across all subtasks and grades. The reasoning being that the system should strive to have at least half the students meeting benchmark performance in all skill areas.

In relation to zero scores, we recommend targets for the percentage of students scoring zero that are at the low ends of the ranges proposed in the workshop because we think this is where the system should target its improvement efforts. Overall performance is best raised by improving the achievement of students at the lowest ends of the distribution. Both EGRA+ and LTTP (and similar interventions in other countries) have been successful at reducing zero scores so we recommend a slightly more ambitious approach to this key indicator.



# ANNEX H: INTRACLASS CORRELATION COEFFICIENTS AND STANDARD DEVIATIONS

The intraclass correlation coefficient (ICC) is a measure of how much variability lies *between* schools and how much lies *within* schools (RTI International, 2009). The ICC can be used to estimate sample size in a school-based cluster design. Accounting for the ICC means that a more accurate estimate of power can be made for the design.

The LTTP passage reading (oral reading fluency) data were used to derive the ICC and standard deviation values as presented in **Table H-1**.

**Table H-1. ICC and Standard Deviation Values for Oral Reading Fluency, by Grade**

GRADE	ICC	STANDARD DEVIATION
1	0.269	10.0
2	0.301	18.9
3	0.264	27.8

## REFERENCE

RTI International. (2009). *Early Grade Reading Assessment toolkit*. Prepared for the World Bank, Office of Human Development, under Contract No. 7141961. Research Triangle Park, North Carolina, USA: RTI International. Retrieved from <https://www.eddataglobal.org/countries/index.cfm?fuseaction=pubDetail&ID=149>