



ED DATA II

Ethiopia Early Grade Reading Assessment

Data Analytic Report:
Language and Early Learning



Ethiopia Early Grade Reading Assessment
Ed Data II Task Number 7 and Ed Data II Task Number 9
October 31, 2010

This publication was produced for review by the United States Agency for International Development. It was prepared by RTI International.

Ethiopia Early Grade Reading Assessment

Data Analysis Report: Language and Early Learning

Ed Data Task Order 7
Ed Data Task Order 9
October 31, 2010

Prepared for
CTO: Mrs. Allyson Wainer

USAID/Ethiopia
Office of Education
United States Agency for International Development (USAID/ETHIOPIA)
Riverside Building
P.O. Box 1014
Addis Ababa, Ethiopia

Prepared by:

Benjamin Piper
RTI International
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

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

Table of Contents

List of Figures vi

List of Tables vii

Acronyms ix

Executive Summary	1
Data Collection	1
Language of Instruction Findings	2
EGRA Findings Snapshot	3
Gender Gaps 5	
Reading Materials	5
Factors Predicting Reading Outcomes	7
Findings Summary	8
Recommendations	8
1. Introduction	1
1.1 Ethiopian Context	1
1.2 Objectives 1	
1.3 Rationale – Why EGRA?	2
1.4 General Education Quality Improvement Program	3
1.5 National Learning Assessment Findings	3
1.6 Woliso EGRA	5
1.7 EGRA Tools	6
1.8 Minimum Learning Competencies and EGRA	7
2. Research Design	9
2.1 Research Design	9
2.2 Regional Selection	9
2.3 Instrument Adaptation	10
2.4 Training and Piloting	11
2.5 Sampling 12	
2.6 Achieved Sample	13
2.7 Analytic Strategies	16
3. Descriptive Statistics	17
3.1 Home Background	17

3.2 Other Background Items	18
4. Findings.....	19
4.1 Early Reading Achievement Is Low	19
4.2 Overall Achievement on Oral Reading Fluency by Region	23
4.3 Regional Comparisons for Amharic and Afan Oromo	25
4.4 Comparing Reading Comprehension and Listening Comprehension Outcomes.....	29
4.5 Reading Outcomes and Gender	31
4.6 Grade Difference.....	32
4.7 Accuracy Analysis	33
4.8 Relationships between Fluency Scores.....	35
5. Predictive Factors.....	36
Student, School and Family Level Predictive Factors	36
6. Regional Analysis Comparing Subtask Achievement.....	38
7. Proposed Benchmarks.....	40
7.1 Quantile Regression Results	40
7.2 Results from High Achieving Poor Schools	40
7.3 Oral Reading Fluency and Reading Comprehension Scores	41
7.4 Levels of Fluency Needed for High Comprehension	43
7.5 Draft Fluency and Comprehension Benchmarks	44
8. Interventions in Early Literacy in Sub-Saharan Africa.....	45
8.1 International Experience	45
8.2 Kenya Intervention Findings.....	46
8.3 South Africa Intervention	47
8.4 Liberia Intervention (EGRA Plus).....	49
8.5 Interventions in Ethiopia.....	50
9. Recommendations.....	50
Appendix A. Sampling by Woreda within Regions	1
Appendix B. EGRA Scores by Region	1
Appendix C. Technical Reliability Analysis	1
Reliability Analysis.....	1
Amharic Tool Analysis	1

Tigrigna Tool Analysis	3
Sidaamu Afoo Tool Analysis.....	5
Hararigna Tool Analysis	6
Somali Tool Analysis.....	7
Appendix D. Head Teacher Questionnaire Findings	1
Background of Respondent Head Teachers	1
School-related Findings	6
Appendix E. Teacher Questionnaire Analysis.....	1
Teachers’ Gender Distribution.....	1
Trained Teachers.....	2
Teachers’ Qualification and Experience	2
Sufficient Learning Materials	3
Supervise Pupils on Use of Library	4
Walking Duration to School	5
Teachers’ Guide	6
Appendix F. Amharic EGRA Tool	1
Appendix G. Afaan Oromo EGRA Tool.....	1
Appendix H. Tigrigna EGRA Tools	1
Appendix I. Somali EGRA Tools.....	1
Appendix J. Sidama EGRA Tools.....	1
Appendix K. Harari EGRA Tools	1
Appendix L. EGRA Teacher Questionnaire.....	1
Appendix M. Ethiopia Director Questionnaire	1

List of Figures

ES Figure 1.	Children Learning in Mother Tongue (%)	ES-2
ES Figure 2.	Children in Grade 2 Reading at Different Benchmark Levels (%)	ES-3
ES Figure 3.	Percentage of Children with Reading Comprehension Scores of 0%	ES-4
ES Figure 4.	Impact of Reading Materials and Wealth on Oral Reading Fluency	ES-5
ES Figure 5.	Children with School Textbook and Other Reading Materials at Home (%)	ES-6
ES Figure 6.	Impacts of Family Background, School, and Individual Factors on Oral Reading Fluency.....	ES-7
Figure 1.	ETNLA Mother Tongue Reading Scores.....	4
Figure 2.	Reading Fluency Results from Woliso Study (Destefano & Elaheebocus, 2009, p. 8)	5
Figure 3.	Ethiopia's Regions	10
Figure 4.	EGRA Zero Scores by Grade.....	20
Figure 5.	Percentage of Children Scoring Zero on Reading Comprehension	21
Figure 6.	Percentage of Children Reading 0 and 60 Words in Grade 2 by Region	22
Figure 7.	Percentage of Children Not Meeting Benchmark	23
Figure 8.	Reading Levels by Region for Grade 2.....	24
Figure 9.	Reading Levels by Region in Grade 3	25
Figure 10.	Subtask Comparisons for Amharic-speaking Schools	27
Figure 11.	Radial Plot Comparing Achievement by Region for Amharic	28
Figure 12.	Subtask Comparisons for Afan Oromo Schools	29
Figure 13.	Reading Comprehension, Listening Comprehension and Teaching Gap Scores.....	30
Figure 14.	Gender Gap in EGRA Scores.....	31
Figure 15.	Gender and Urbanicity Interactions, Controlling for Grade	32
Figure 16.	Grade Difference (between Grade 2 and 3) by EGRA Task.....	33

Figure 17.	Letter and Word Accuracy Scores by Language, Including Difference Between Accuracy Scores.....	34
Figure 18.	Correlations between EGRA Tasks, by Sabeen or Latin Script.....	35
Figure 19.	Predictive Factors Relationships with Oral Reading Fluency.....	37
Figure 20.	Amhara Radial Plot Comparing Achievement by Grade and Gender for Six EGRA Tasks.....	39
Figure 21.	Oromiya Radial Plot Comparing Achievement by Grade and Gender for Six EGRA Tasks.....	39
Figure 22.	Oral Reading Fluency and Comprehension Rates for Amharic and Afan Oromo.....	42
Figure 23.	Oral Reading Fluency and Reading Comprehension Scatter Plots for Tigrigna, Sidaamu Afoo, Hararigna, and Somali.....	43
Figure 24.	Oral Reading Fluency Scores Necessary for 80% and 100% Comprehension	44
Figure 25.	Impact of South Africa SMRS on Reading Outcomes, by Task...	48
Figure 26.	Impact of South Africa SMRS Program on Oral Reading Fluency Outcomes.....	49
Figure 27.	Impact of Liberia EGRA: Plus on Oral Reading Fluency Outcomes.....	50

List of Tables

ES Table 1.	Summary of EGRA Data Collection Sample.....	ES-2
Table 1.	Ethiopian Third National Learning Assessment Mean Scores by Year (MOE, 2008, p. 49).....	3
Table 2.	Ethiopian Third National Learning Assessment Achievement Levels by Subject (MOE, 2008, p. 47).....	4
Table 3.	Minimum Learning Competencies in Mother Tongue and Associated EGRA Tasks	7
Table 4.	Sampled Districts for EGRA Study, by Region	12

Table 5.	EGRA 2010 Sample, by Language, School Type, and Location	14
Table 6.	Head Teacher and Teacher Questionnaires by Region.....	15
Table 7.	EGRA 2010 Sampling Frame	16
Table 8.	Descriptive Statistics of Self-reported Figures by Region.....	18
Table 9.	Subtask Comparisons for Amharic Assessed Regions.....	26
Table 10.	Subtask Comparisons for Afan Oromo-assessed Regions	28
Table 11.	Quantile Regression Results for Each Level of Oral Reading Fluency by Language	40
Table 12.	Oral Reading Fluency Scores at the School Level for High Scoring Schools with High Levels of Student Poverty	41
Table 12a.	Draft Oral Reading Fluency Benchmarks by Language.....	45
Table 13.	Kenyan Reading Scores at Baseline	46
Table 14.	South Africa Impacts from Literacy Intervention.....	47

Acronyms

CTE	College of Teacher Education
DIBELS	Dynamic Indicators of Basic Early Literacy Skills
EBNLA	Ethiopian Baseline National Learning Assessment
EMIS	Education Management Information Systems
EGRA	Early Grade Reading Assessment
ES	Executive Summary
ESDP	Education Sector Development Program
ESNLA	Ethiopian Second National Learning Assessment
ETNLA	Ethiopian Third National Learning Assessment
ETQAA	Education and Training Quality Assurance Agency
FDRE	Federal Democratic Republic of Ethiopia
GEQIP	General Education Quality Improvement Program
IQPEP	Improving Quality in Primary Education Program
MLC	Minimum Learning Competencies
MOE	Ministry of Education
NLA	National Learning Assessment
NER	Net Enrollment Rate
PASEC	Programme d'Analyse des Systemes Educatifs de la CONFEMEN
PIRLS	Progress in International Reading Study
PISA	Programme for International Student Assessment
pm	per minute
REB	Regional Education Bureau
RTI	Research Triangle Institute International
SACMEQ	Southern and Eastern African Consortium for Monitoring Education Quality
SIP	School Improvement Program
SMRS	Systematic Method for Reading Success

SNNPR	Southern Nations, Nationalities and People’s Region
TEI	Teacher Education Institution
TIMSS	Trends in International Mathematics and Science Study
TTC	Teacher Training College
T&L	Teaching and Learning
UNICEF	United Nations Children’s Fund
UNESCO	United Nations Educational, Scientific and Cultural Organization
USAID	United States Agency for International Development
WEO	Woreda Education Office
wpm	words per minute

Executive Summary

In May and June 2010, an early grade reading assessment (EGRA) was performed in eight regions in Ethiopia. The EGRA was a collaboration among the Ministry of Education (MOE), RTI International, members of the Education and Training Quality Assurance Agency (ETQAA), the Improving Quality in Primary Education Program (IQPEP), several core processes, and other stakeholders, and was a study of the reading skills in Ethiopia in a variety of areas. Due to the efforts of the MOE, and the generous funding of United States Agency for International Development (USAID)/Washington and USAID/Ethiopia, this EGRA study is the largest of almost 50 performed.¹

The assessment was developed for 6 languages in Ethiopia, such that Grade 2 and Grade 3 students were assessed in Tigrinya, Afan Oromo, Amharic, Somali, Sidaamu Afoo, and Hararigna. The assessments included a variety of subtasks, including letter (or fidel) sound fluency, phonemic awareness, word naming fluency, unfamiliar word naming fluency, oral reading fluency, reading comprehension, and listening comprehension. The assessments were leveled according to the MOE's Minimum Learning Competencies. The sampling included 338 schools and 13,079 students assessed by RTI and the IQPEP with the MOE. The purpose was to investigate the children's reading skills in the context of the General Education Quality Improvement Program (GEQIP) and the rapidly changing primary school environment in Ethiopia. In addition to student literacy assessments, a family background questionnaire was administered to students, and head teacher and teacher questionnaires at the school level. School level and teacher level data were matched with student achievement data to determine how student background, the classroom environment, and community factors were correlated with student outcomes.

Data Collection

Data collection took place between May 10, 2010, and June 16, 2010. Data collectors were trained intensively in the basics of reading assessment, specific to each language by RTI, IQPEP and renowned language experts from many universities, particularly Addis Ababa University. Assessors included experts from Colleges of Teacher Education (CTE), Woreda Education Offices (WEO), and Regional Education Bureau (REB) staff, as well as private data collectors, interrater reliability scores were higher than .94. Supervised by experts from the MOE and REBs, a team of 109 assessors was deployed in the eight regions. Table 1 summarizes the sample.

¹ <http://www.rti.org/page.cfm?objectid=0105C3ED-F254-B0BE-B763260791DE62B6>

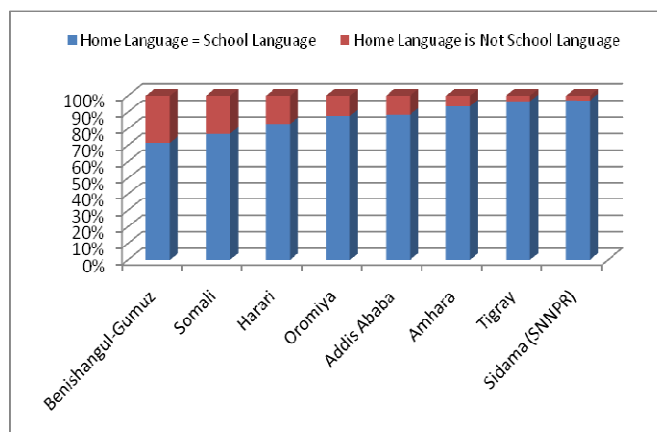
ES Table 1. Summary of EGRA Data Collection Sample

Region	Language	Woredas	Schools	Children Assessed
Tigray	Tigrinya	11	39	1551
Amhara	Amharic	15	60	2316
Oromiya	Afan Oromo	16	63	2491
Somali	Somaligna	4	33	1183
Benishangul-Gumuz	Amharic	6	35	1289
SNNPR(Sidama)	Sidaamu Afoo	11	42	1752
Harari	Hararigna	16 total	2	80
	Amharic		8	320
	Afan Oromo		21	785
Addis Ababa	Amharic	11	33	1312
Total			338	13,079

Language of Instruction Findings

Ethiopia’s primary school language policy is often noted for being the most progressive policy in Sub-Saharan Africa with respect to mother tongue instruction. The EGRA study asked children whether they speak the same language at home as they are taught in at school. Figure 1 shows that, in each region, the percentage of children learning in their mother tongue (home language) ranges from 71.5% (Benishangul-Gumuz) to 97.8% (Sidama zone, SNNPR), with the majority of regions surveyed having more than 85% overlap between language of instruction and mother tongue. This is certainly one of the highest uses of local languages in primary school anywhere in the continent, and likely contributes to literacy acquisition in Ethiopia, though the scores remain lower than

ES Figure 1. Children Learning in Mother Tongue (%)



expected. However, note that in each region a significant proportion of children learn in languages they do not speak at home; e.g., 28.5% in Benishangul-Gumuz and 12.2% in Oromiya.

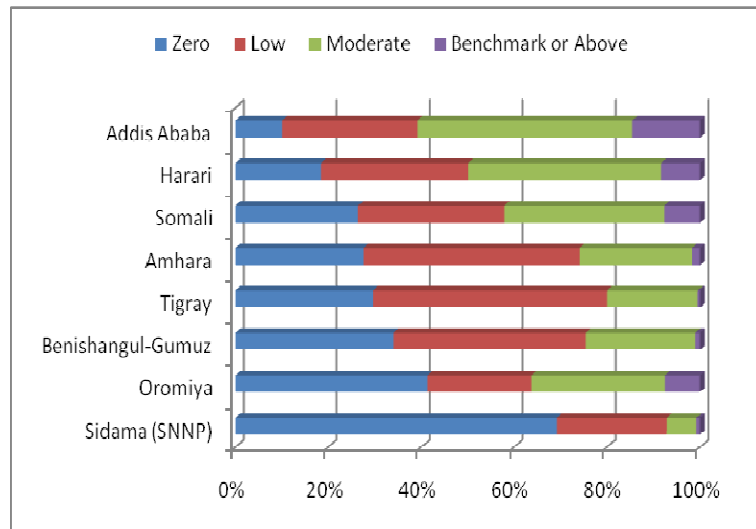
EGRA Findings Snapshot

This dataset provides opportunity for complex analysis of interesting relationships between language, student background, and student reading outcomes. The findings presented in this executive summary are in the areas of oral reading fluency and reading comprehension, as defined below:

1. *Connected text oral reading fluency*: ability to read a passage, about 60 words long. It is timed to 1 minute. The passages were targeted at the early Grade 2 level in vocabulary and complexity.
2. *Comprehension in connected text*: ability to answer several comprehension questions based on the passage read. Each assessment had 5 or 6 questions, and the scores presented are percentage-correct.

Figure 2 shows the percentage of children in each region in Grade 2 reading at different benchmark levels. The blue bars represent children who were unable to read a single word on the connected text oral reading fluency measure; the red bars indicate children who were very weak readers; the green bars represent children who read moderately well; and the purple bars describe children who read at the expected rate for the grade. The figure shows

ES Figure 2. Children in Grade 2 Reading at Different Benchmark Levels (%)

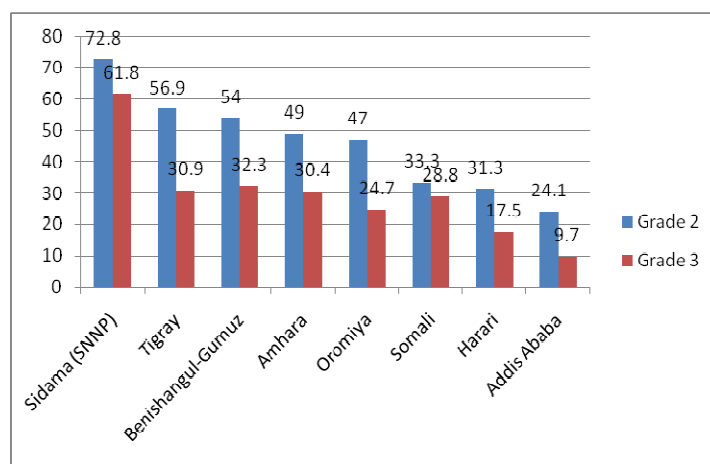


that a **significant percentage of children in Grade 2 read zero words correctly**. In Sidama the percentage of nonreaders was 69.2%, and in Oromiya it was 41.2%. Only Harari (17.9%) and Addis Ababa (10.1%) have percentages of zero scores less than 20%, with the largest regions (SNNP, Oromiya, Tigray, and Amhara) all having Grade 2 zero scores above 25%.

Even in Grade 3, significant percentages of children remained nonreaders. For Somali (21.4%), Amhara (17.0%), Sidama (54.0%), and Oromiya (20.6), it is striking that after 3

years of school, such large proportions of children remained completely illiterate in their mother tongue. Interestingly, it appears that large decreases in the percentage of nonreaders occur between Grade 2 and 3 for Oromiya, Benishangul-Gumuz, and Tigray. However, for those children who were just beginning to learn to read at the end of Grade 3, it was likely too little and too late. These children are likely candidates for dropout or repetition, and they certainly run the risk of not being prepared for the end-of-primary examinations. The purple bars that relate to children reading at the expected rate indicate very low scores. **In each of the 8 regions, at least 80% of children—and in the case of Sidama, 100%—were not reading at the expected oral reading fluency rate.**²

ES Figure 3. Percentage of Children with Reading Comprehension Scores of 0%



The problem of very low achievement exists for oral reading fluency as well as reading comprehension. Figure 3 shows the percentage of children whose reading comprehension scores were 0% correct. It is clear that a large percentage of children did not comprehend what they were reading, though the questions were quite simple. In Sidama (72.8%), Tigray (56.9%) and Benishangul-Gumuz (54.0%), **more than half of the region’s children in Grade 2, did not understand a story at all.** Even in the urban regions (Harari and Addis Ababa), one quarter or more of

children could not comprehend basic questions. There were some improvements between Grade 2 and 3, with less than one third of Grade 3 children scoring zero in all regions (except Sidama at 61.8%). On the other hand, the stories and associated questions were developed such that Grade 2 children should have been able to answer 4 or 5 of the 5 comprehension questions correctly.

These findings show that even though the purpose of mother tongue instruction is to ensure that children understand what they read, the children’s inability to decode the words means they are unable to understand the text, although they are likely to have the vocabulary to understand it. This is confirmed after analysis of the listening comprehension task, which shows that the average child can listen to and comprehend spoken stories quite well. The gap between the reading comprehension and listening comprehension scores is consistently large, and shows that **the problems identified by**

² This is based on benchmarks from other countries and preliminary analysis from Ethiopia. Using these data, the MOE will be able to determine appropriate grade-level benchmarks for children’s oral reading fluency.

this EGRA are specific to teaching *reading*, and not due to language issues in the children.

Gender Gaps

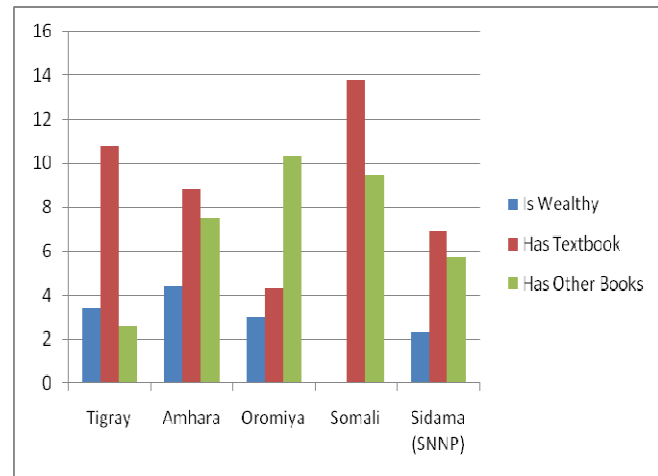
In EGRA administrations across Sub-Saharan Africa, RTI has found that in most instances, girls significantly outperform boys on early reading tasks. We investigated the relationship between gender and urbanicity to determine whether there are systematic gender gaps in reading achievement as measured by the EGRA. The 2007 National Learning Assessment (NLA) literacy results found that boys outperformed girls in rural areas, but there were no differences in urban areas. The EGRA study found almost the same result. Across regions, there was a statistically significant difference between boys' and girls' achievement in all reading tasks (save listening comprehension) that favored boys. On the other hand, in urban areas, girls outperformed boys, though in some cases it was not statistically significant. This relationship mirrors the NLA results. **There seems to be a problem in the education system for rural girls**, since girls can read in urban areas, and in other Sub-Saharan African countries girl (both urban and rural) outperform boys.

Reading Materials

Figure 4 shows the impact that three student characteristics had on oral reading fluency. First, the blue bars show the impact that being a wealthy student had on student achievement across the five languages, with the largest impact of 4.4 words per minute (wpm) in Amhara, and the smallest of 0 words in Somali.

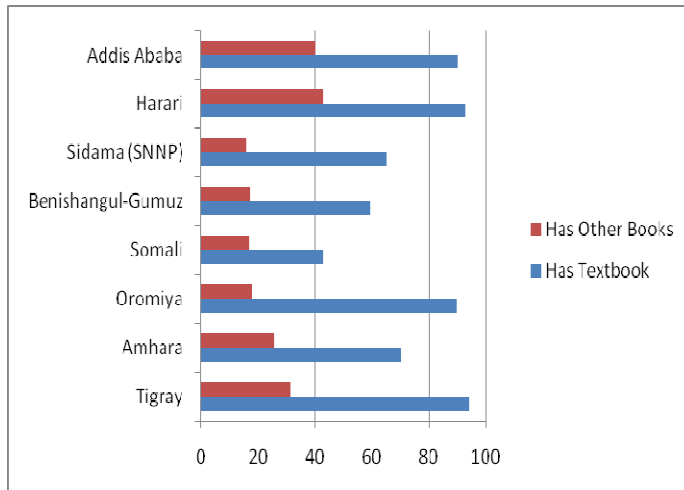
The red bars show the impact that having the language textbook had on student achievement. Having the book increased oral reading fluency by between 4.3 wpm (Oromiya) and 13.8 (Somali) in these large regions. The impact of having a book was larger than being wealthy, and that was true for all regions presented here. Finally, the green bars show the effect of having extra reading materials at home, which was significantly larger than being wealthy (except in Tigray) and larger in some cases than having the school reading textbook. For example, having other books is related with 10.3 words more per minute in Oromiya, much more than the 3.0 words related to being wealthy. Research in Sub-Saharan Africa decades ago confirmed the importance of

ES Figure 4. Impact of Reading Materials and Wealth on Oral Reading Fluency



having a textbook,³ and the findings from this EGRA study mirror what was known several decades ago: Having access to materials to read, both inside and outside of school, encourages achievement in literacy over and above the wealth of individual families.

ES Figure 5. Children with School Textbook and Other Reading Materials at Home (%)



This leads to the question of what percentage of children in Ethiopia have access to the valuable reading materials that make such a large difference on student achievement. Figure 5 shows the percentages of children with the school textbook (blue bar) and other reading materials at home (red bar). There is a wide range of responses as to whether the child had the textbook across the regions: **94.5%**

of Tigray’s children had the book, but only 42.8% of Somali children and 59.3% of Benishangul-Gumuz children had it.

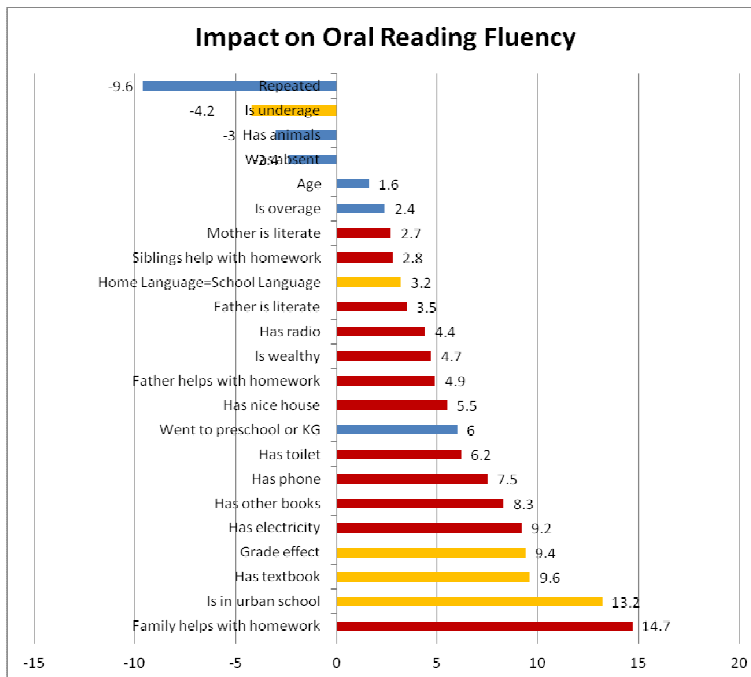
Given the relationship between having the book and student achievement, there is a clear policy implication: Getting the school textbook into the hands of each child is critical. In Sidama (65.1%) and Amhara (70.4%), the percentages of children with textbooks are low. Ensuring that children have other materials to read has also shown to be a critical variable, but this figure shows how few children have any other reading materials at home or at school. In urban regions (Harari and Addis Ababa), 40% or more of the children had other reading materials, but in the rest of the regions, the percentages were much lower. **In Oromiya, Somali, Benishangul-Gumuz, and Sidama, only 1 in 6 children had any other reading materials.** Given that having even a tiny amount of extra material to practice reading was so strongly correlated with policy outcomes, it appears that providing books to children or encouraging families and communities to have books at home is an important next step to improve the quality of literacy outcomes in Ethiopia.

³ Fuller, B. 1986. *Raising School Quality in Developing Countries: What Investments Boost Learning*. Washington, D.C.: The World Bank.

Factors Predicting Reading Outcomes

The very large dataset analyzed in this paper provides a significant opportunity to look at the factors related to student reading achievement (as measured by oral reading fluency scores). Controlling for region, these factors are predictive of student outcomes. The colors in Figure 6 relate to the level of the factors. Red bars show family background factors; gold bars show school level factors; and blue bars show individual factors. Clearly student achievement was highly dependent on all three levels. At the school and system level, it is clear that children learn a significant amount in Grade 3 (9.4 wpm), and if children learn in the same language they speak at home, it has a positive impact on student achievement (3.2 wpm), though a full 11.8% of the children in these regions do not learn in the same language that they speak at home.

ES Figure 6. Impacts of Family Background, School, and Individual Factors on Oral Reading Fluency



Critically, as mentioned above, **having a textbook provided by the school was associated with 13.2 more words per minute**. At the family level, many wealth factors were related to higher achievement (having a radio; having a nice house, a phone, electricity, and family helping with homework). Schools cannot affect these factors. However, whether a child had other books in the home (8.3 wpm), the father helped with homework (4.9 wpm), or if the entire family was available to assist with schoolwork (14.7 wpm) can make a big difference. At the student level, repetition (-9.6

wpm) was a significant problem, as was being underage for the grade (-4.2 wpm). These factors are measured at the student level, but actually are system level factors that can be improved by ensuring adherence to the entry policies of the REBs and the MOE. In short, there are many factors related to student achievement that the school and the system can improve upon.

Findings Summary

The findings suggest that while **children attend school for two or three years, a significant percentage is illiterate**. These findings buttress the work of the Ministry of Education and the National Organization of Examinations in the NLA and show that there is strong evidence that reading achievement is low in all regions sampled, with the urban regions Harari and Addis Ababa modestly outperforming the other regions. The language usage findings show strong adherence to the language of instruction policy, and that most children in the regions sampled learn in the language they speak at home, which increases their ability to understand and to read. Most critically, these findings show that reading achievement is very low in Ethiopia. When asked to read a simple passage at a Grade 2 level, many regions had more than 30% of Grade 2 and 20% of Grade 3 unable to do so successfully, with children in Sidama zone and Oromiya region particularly struggling. When it comes to reading comprehension, scores are extremely low, with **more than 50% of the children in most regions unable to answer a single simple comprehension question**. The exceptions are for urban areas and urban regions, and in some schools in Grade 3, where children are only starting to understand what they read. This appears to be too little, too late, and the current status of reading skills suggests that significant interventions in the quality of reading instruction and the provision of reading materials are necessary.

Recommendations

Recommendations from this study will be shared with the policy workshop to be held soon, and to include stakeholders from across the education sector and mirror the ideas that will be presented to the MOE in September, 2010. The recommendations include the following:

- **Focus resources on reading instruction.** Very few teachers reported any in-service training in reading methods and pedagogical techniques. This should be supported by specific training for teachers on how to appropriately and successfully teach children this content. This will support their ability to help children learn the fundamentals necessary for successful reading, including vocabulary, phonemic awareness, fluency, and decoding. Note that teaching teachers how to teach reading must be language-specific. Amharic and Afan Oromo are very different languages structurally, for example, and precision is needed to focus training on how best to improve primary education.
- **Start early, in Grades 1 and 2.** The findings show that teachers' views of when pupils could read and understand what they read are important for pupil outcomes. They also show that much learning of the fidel and alphabet is occurring primarily in Grade 3. This suggests that in some Ethiopian classrooms teachers wait too long to teach students how to read and expect too little from their young learners, and this has

implications for what children can gain from early primary school. It is recommended, therefore, that Grade 1 and 2 pedagogy focus most heavily on early reading acquisition and outcomes.

- **Encourage reading in the community.** Our findings showed that few classrooms were stocked with reading books, and very few children had many reading materials at their homes. Thus children have limited exposure to the joys of reading engaging and interesting materials appropriate for their developmental stages. A two-pronged effort should be made to increase the amount of reading material in classrooms and encourage families to make reading a part of their daily family activity. This could be accomplished by awareness raising efforts at the regional and woreda level.
- **Review in-service teacher professional development.** The findings from this study clearly indicate that, in many schools, little reading instruction happened, though mother tongue class is a significant part of the day. By this we mean that there was far less interaction between teachers and students around letters, words, sentences, and stories than there should be. This need not be the case, and experiments in Kenya, Liberia, and South Africa show that teachers can be very receptive to focused in-service professional development supporting skill acquisition in early literacy interventions. It is recommended that the in-service teacher professional development programs target the building blocks of reading and where possible, provide targeted lesson plans for teachers.
- **Set literacy benchmarks.** The complex language environment in Ethiopia means that policy makers should think carefully about outcomes they expect children to achieve by the end of Grade 2. This can be added to the reading portion of the Minimum Learning Competencies and will prepare Ethiopia for the indicative frameworks designed by the Fast Track Initiative. The findings suggest that without benchmarks, and work to achieve those benchmarks, children may never reach reading fluency.
- **Improve the quality of reading instruction.** There are some critical areas necessary for immediate intervention.
 - **Use letter sounds and the fidel as building blocks for reading.** There is a strong correlation with a child’s scores on letter sounds with their reading fluency and comprehension scores in languages that use the alphabet. This means that these building blocks for fluency and comprehension are important skills for children to master in Grade 1.
 - **Teaching decoding is critical.** Many of the classroom observations in this sample revealed teaching situations where teachers pointed to words and encouraged the children to call out the word house. However, when faced with very similar words, those same children did very poorly because the pedagogy encourages the children to memorize particular words, and spends much less time training them in how to decode and “solve” new words.

- **Teach formal comprehension strategies.** The children in this sample had very low comprehension levels. This is partially because the children were likely to have limited oral vocabulary skills, in particular, but also because the children did not have much formal training in comprehension strategies. These can be systematically taught. Note, however, that without the ability to read fluently, comprehension is nearly impossible.
- **Expand literacy interventions.** Ethiopia has been very receptive to changes that can impact the quality of reading instruction; yet the evidence suggests that more effort is needed. Combined with scripted lesson plans, material book development and provision, and ongoing teacher professional development, it is clear that improvements to the quality of reading outcomes can be had in Ethiopia. We suggest that the following elements be included in the literacy interventions that are attempted in Ethiopia.
 - **Development of targeted lesson plans.** Teachers should be provided with specific instructions as to how to teach early reading acquisition, since most pre-service programs do not provide the level of detail and precision necessary to do it properly.
 - **Provision of ongoing support for teachers.** In order to support the behavioral changes necessary to help teachers to teach significantly better, they need ongoing support using a combination of new instructional methods and opportunities to discuss how their experiments with the new methods are working. A coaching model might provide effective support.
 - **Development and usage of significant reading materials.** Leveled materials need to be developed to support the graduated instruction in Grade 1 and Grade 2. These materials need to be read easily by burgeoning learners and incorporated into lessons.

Comments or questions on this executive summary are welcome, and should be sent to bpiper@rti.org.

1. Introduction

The Ethiopian Ministry of Education's interest in ensuring the quality of primary education across Ethiopia has led to an exploration of the means by which the quality of early grade reading can be assessed. The Early Grade Reading Assessment (EGRA) is a tool that allows careful investigation of the component parts of early literacy acquisition. Given the interest of the Ministry of Education (MOE) in collecting regionally representative data in a significant number of regions, RTI International was tasked with development of the assessment in 6 languages: Afan Oromo, Amharigna, Tigrigna, Somaligna, Sidamigna, and Hararigna, and the administration of EGRA in eight regions, encompassing almost 96% of Ethiopia's population. This report presents the findings from this assessment data to allow the MOE to unpack the variety of impediments to early grade literacy acquisition and inform the development of interventions to improve the quality of early reading teaching and learning.

1.1 Ethiopian Context

Ethiopia has been at the forefront of the continent's move toward improving access to education. Since the overthrow of the Derg in 1991, Ethiopia's primary education net enrollment rate has increased from less than 30% to more than 90%. This tremendously rapid increase in enrollment has been lauded in a recent United Nations Educational, Scientific and Cultural Organization (UNESCO) Global Monitoring Report as an example of the type of government commitment necessary to make dramatic changes in enrollment. Ethiopia also has one of the most language diverse environments in Sub-Saharan Africa, and has implemented probably the most comprehensive mother tongue education policy on the continent.

1.2 Objectives

The ability to read and understand a text is the most fundamental skill a child learns. Without literacy there is little chance a child will escape the intergenerational cycle of poverty. Yet, in many countries, students enrolled for four or as many as six years are unable to read and understand simple texts. Evidence indicates that learning to read early and at a sufficient rate, with comprehension, is essential for learning to read *well*, and reading well is critical for overall academic success. Children who do not learn to read in the first few grades are more likely to repeat and eventually drop out, or will fall behind others for the rest of their lives, and countries where the population cannot read well will tend to lag behind the more educated countries in terms of student achievement. In order to estimate the levels of reading in Ethiopia, the National Learning Assessment (NLA) program has been investigating the quality of literacy skills for Grade 4 students since 2000. The remarkably low (and declining) scores on the Grade 4 NLAs suggest the need for further assessment instruments that can identify points of weakness and strength not

easily identified by paper and pencil tests, such as NLA. The objective of the EGRA, then, is to provide evidence in these important areas. The purpose, of course, is to inform the efforts in the area of education quality, such as GEQIP.

1.3 Rationale – Why EGRA?

It is in this context that the MOE, including members of the Assessment Sub-Process in the Ethiopian Education Training Quality Assurance Agency (ETQAA) and partnering with RTI, undertook an EGRA in 2010. The strategy behind the EGRA package of assessments is to determine the areas of reading or prereading in which children have particular difficulty, which will allow policy makers in Ethiopia to be able to target interventions at the reading components most likely to increase outcomes. Given the low levels of student achievement in Sub-Saharan Africa as measured by Southern and Eastern African Consortium for Monitoring Education Quality (SACMEQ) and Programme d'Analyse des Systemes Educatifs de la CONFEMEN (PASEC)—and the lower levels of student achievement in developing countries in international assessments, such as Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS)—several international donors (particularly the United States Agency for International Development and the World Bank) and organizations (particularly RTI) collaborated to fund and create the EGRA. Development of the EGRA occurred between 2006 and 2007, drawing on research from other contexts, especially the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), a U.S.-based early literacy assessment.

The need for EGRA is clear. The average child from the low-income countries participating in international tests performed at approximately the third or fourth percentile of a developed country distribution. Unfortunately, it is difficult to disentangle whether a child's knowledge and skills are lacking, or whether the lack of foundational reading ability hinders the child's ability to understand the assessment. In response to this need, EGRA was designed to orally assess literacy acquisition for children in grades 1 through 4. The EGRA instrument measures oral reading fluency, reading comprehension, letter recognition, and phonemic awareness, among other skills predictive of future reading success. The need for EGRA has been echoed in low-income countries across the world, with the instrument implemented in more than 40 countries and 60 languages since 2006. EGRA is used for two main purposes. First, EGRA is designed to provide governments and Ministries of Education with policy making information regarding areas of improvement. Second, early grade reading achievement is a proxy for the quality of the early part of a school system. If reading, particularly in local languages, is not being learned sufficiently by children, it is likely that achievement in other subjects will be similarly low. The international community has been receptive to EGRA and it has become a relatively standard assessment tool with applications in many Sub-Saharan African countries. At the local level, the rationale is that this assessment will provide

REBs and the MOE the ability to look carefully at their policies and programs and to investigate how best to improve the quality of education.

1.4 General Education Quality Improvement Program

The Ethiopian MOE, as part of Ethiopia’s Poverty Reduction Strategy Program, has noted that while education enrollment increases are important, they are not sufficient. As a result, within the Education Sector Development Program III (ESDPIII) and the forthcoming ESDPIV, attention has been and will continue to be paid to the quality of education necessary to ensure that students that graduate from the first and second cycle of primary school leave with the requisite skills to be a contributing citizen and be prepared for secondary school. In addition, the GEQIP, which is at the core of the strategy for improving the quality of education, targets the resources needed to provide education, the curriculum delivered in that education, and the pedagogy by which the curriculum is delivered.

1.5 National Learning Assessment Findings

Ethiopia has a history of producing high-quality national learning assessment materials and results. In 2000, the baseline national learning assessment (EBNLA) was undertaken, in 2004 the second national learning assessment (ESNLA) occurred, and in 2007 the third national learning assessment (ETNLA) was implemented. While scaling issues exist,⁴ the MOE’s (2008) findings were that the quality of the reading comprehension outcomes have decreased since the 2004 ESNLA, as Table 1 shows. The mean score in the 2007 ETNLA was only 43.9, which was much lower than in either the 2004 EBNLA (64.5) or the 2000 EBNLA (64.3).

Table 1. Ethiopian Third National Learning Assessment Mean Scores by Year (MOE, 2008, p. 49)

Table 15. Comparison of mean scores among the three national assessments

Subjects	EBNLA (2000)	ESNLA (2004)	ETNLA (2007)
Mathematics	39.3	39.7	40.3
Reading	64.3	64.5	43.9
English	40.5	38.7	36.5
Env. Science	48.1	51.7	42.6
Composite	47.9	48.48	40.9

With respect to the 2007 ETNLA, the breakdown of Grade 4 scores is as follows in Table 2. Only 14.6% of children were deemed proficient, with a much larger proportion achieving at a below basic level (51.7%). Note that the scores presented here are from a Grade 4 assessment. Ethiopia’s NLA scores are similar to those of other countries, in that the scores are much lower than expected. However, while the MOE analysis team should

⁴ The ETNLA report notes that the reading comprehension test was dramatically changed between 2004 and 2007, so comparisons must be made carefully.

be lauded for its analysis of the raw data and their ability to understand the school and student level factors related to the low achievement, it remains a matter for research what are the requisite skills that children have or do not have in Ethiopia that might be related to the low scores in reading comprehension identified in the ETNLA.

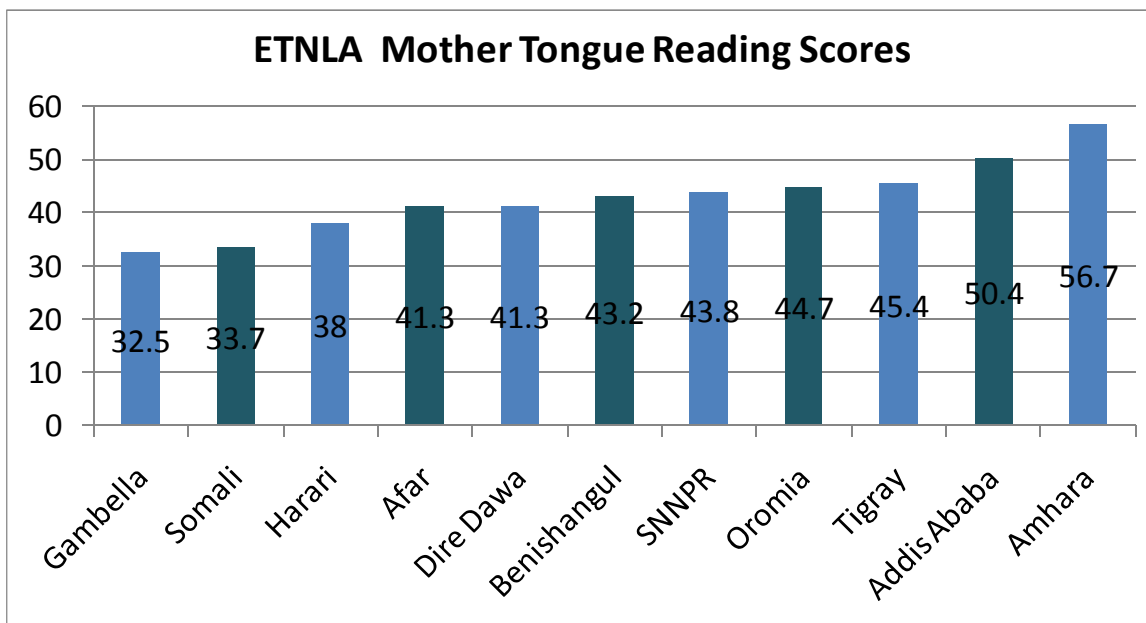
Table 2. Ethiopian Third National Learning Assessment Achievement Levels by Subject (MOE, 2008, p. 47)

Table 12. Achieved performance level for each subject

Subject	Proficient	Basic	Below Basic
Reading	14.6	33.7	51.7
English	16.9	31.5	51.6
Mathematics	17.1	29.0	53.9
Env. Science	16.3	38.5	45.1
Composite	14.7	37.8	47.4

Figure 1 presents the ETNLA’s average Grade 4 mother tongue reading scores by region, in Grade 4. Note that the highest scores were found in Amhara, Addis Ababa, Tigray, and Oromiya, but that only Amhara and Addis Ababa scored more than the expected level of 50%. The lowest scores were found in Gambella (32.5%), Somali (33.7%), and Harari (38.0%).

Figure 1. ETNLA Mother Tongue Reading Scores



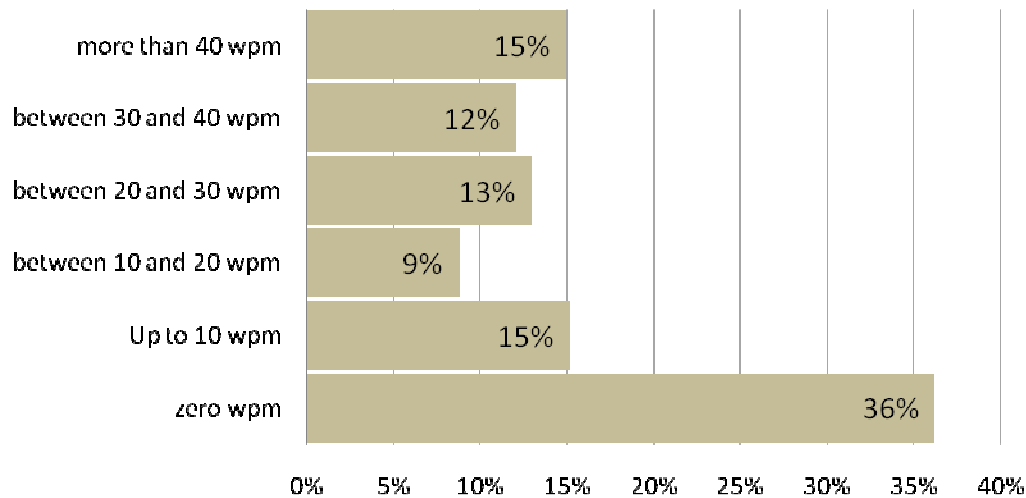
It appears that, given the low reading scores found in the ETNLA in particular, more research is necessary at early levels (prior to Grade 4) to determine the areas of quality improvement that might support the GEQIP efforts of the MOE.

1.6 Woliso EGRA

A small EGRA study took place in Woliso woreda in 2009 in Grade 3.⁵ The sample was quite small and is not thought to be representative of Ethiopia or Oromiya region. The findings were striking, though, as Figure 2 indicates. In the study, 36% of the sampled children could not read a single word of a simple text, and only 15% of children could read 40 words per minute (wpm) or more.

Figure 2. Reading Fluency Results from Woliso Study (Destefano & Elaheebocus, 2009, p. 8)

Reading Fluency Results for All Students, % of students who could read:



⁵ DeStefano, J & Elaheebocus, N. (2009). School quality in Woliso, Ethiopia: Using opportunity to learn and early grade reading fluency to measure school effectiveness. EQUIP2 report submitted to USAID.

1.7 EGRA Tools

EGRA is an orally administered assessment targeted at measuring the prereading and reading skills foundational to later reading (and academic success). EGRA takes approximately 15 minutes to administer and is often combined with a questionnaire measuring a variety of student background variables to assist in explaining some of the reading outcome findings. The Ethiopian EGRA consists of the following components, which have been found to be highly correlated with one another.

1. *Letter-naming (or fidel identification) fluency*: ability to read the letters of the alphabet (or the fidel) without hesitation and naturally. This is a timed test that assesses automaticity and fluency of letter or fidel sounds. It is timed to 1 minute, which saves time and also prevents children having to spend time on something that is difficult for them.
2. *Phonological awareness*: awareness of how sounds work with words. This is generally considered a prereading skill, and can be assessed in a variety of ways. In some Ethiopian languages, this task might be designed to determine whether children could differentiate the first syllable (or fidel) in a word, or whether they could identify all of the fidels in a word.
3. *Familiar word fluency*: ability to read high-frequency words. This assesses whether children can process words quickly. The lists of words were derived from the 50 most frequently used words in Grade 2 and 3 textbooks in each language. It is timed to 1 minute.
4. *Non-familiar or non-sense word fluency*: ability to process words that could exist in the language in question, but do not. The words were derived from the list of familiar words and follow the common patterns of the language. This component assesses a child's ability to "decode" words fluently. It is timed to 1 minute.
5. *Connected text oral reading fluency*: ability to read a passage, about 60 words long, that tells a story. The stories were created to be appropriate for particular regions and targeted at Grade 2 and Grade 3 children. The component is timed to 1 minute.
6. *Comprehension in connected text*: ability to answer several comprehension questions based on the passage read.
7. *Listening comprehension*: being able to follow and understand a simple oral story. This assesses a child's ability to concentrate and focus to understand a very simple story, assessed by asking simple noninferential (factual) questions. It is considered a prereading skill.

Note that each EGRA task was always adapted and essentially redesigned uniquely for each of the 6 languages in Ethiopia.

1.8 Minimum Learning Competencies and EGRA

While EGRA has gained international acceptance and has been undertaken in more than 40 countries worldwide, it is useful to determine whether EGRA relates to the Ethiopian curriculum at grade level. EGRA does not test whether children have learned an appropriate amount of the curriculum, but rather whether they have the basic skills required for a particular level. However, it is fair to assess whether EGRA is in line with the curricular goals of Ethiopia. To that end, a close analysis of the Minimum Learning Competencies (MLCs) document,⁶ allows an analysis of whether and how the EGRA is correlated with the expected tasks for a particular grade level. In order to do this, the portion of the MLC document related to mother tongue instruction in Grades 1-4 was referenced, with particular attention to the Listening and Speech and Reading sections. Table 3 makes clear that the EGRA fits well into the expected learning competencies for Grade 2 and Grade 3. In fact, EGRA appears to be targeted slightly below level for Grade 3 students in some tasks (letter/fidel fluency, word fluency, decoding, phonemic awareness) and for Grade 2 students in some tasks (letter/fidel fluency, phonemic awareness). Each EGRA task finds its match in an MLC competency.

Table 3. Minimum Learning Competencies in Mother Tongue and Associated EGRA Tasks

Minimum Learning Competency Statement	Grade	Content	EGRA Task
Ask simple questions related to the lessons they learned.	1	Listening	Listening Comp, Reading Comp
Listen to simple community story and speak about the characters	2	Listening	Listening, Reading Comp
Listen to stories and histories and tell one main idea	2	Listening	Listening
Speak sequentially the action of characters in a story	2	Listening	Listening, Reading
Listen to simple traditional stories and tell main ideas and characters	3	Listening	Listening, Reading
Ask questions that require explanations and reasons from the lesson they learned	3	Listening	Listening, reading
Explain events related to cultural customs and traditional practices	3	Listening	Listening, reading
Identify the sounds and read the alphabets excluding hybrids	1	Reading	Fidel/letter fluency, phonemic awareness
Read about 15 words at a glance without counting letters	2	Reading	Word fluency, decoding
Read silently and apply 2–3 sentences long guidelines	2	Reading	Oral reading fluency
Read silently about 3–4 sentences & match with pictures	2	Reading	Oral reading fluency
Read and comprehend short passages from textbooks	2	Reading	Oral reading fluency
And give appropriate answer for selected questions	2	Reading	Oral reading fluency, reading comprehension

⁶ Ministry of Education (2009). Minimum learning competencies Grades 1-4. Addis Ababa, Ethiopia.

Minimum Learning Competency Statement	Grade	Content	EGRA Task
Read loudly 1–2 sentences properly	2	Reading	Oral reading fluency
Read longer words at once	3	Reading	Word fluency, decoding
Identify special words and find out meanings	3	Reading	Decoding
Understand the meaning and concepts of new words	3	Reading	Decoding, oral reading fluency
Read incomplete story and guess its end	3	Reading	Oral reading fluency, reading comprehension
Read and understand 4–6 paragraphs	3	Reading	Oral reading fluency, reading comprehension

2. Research Design

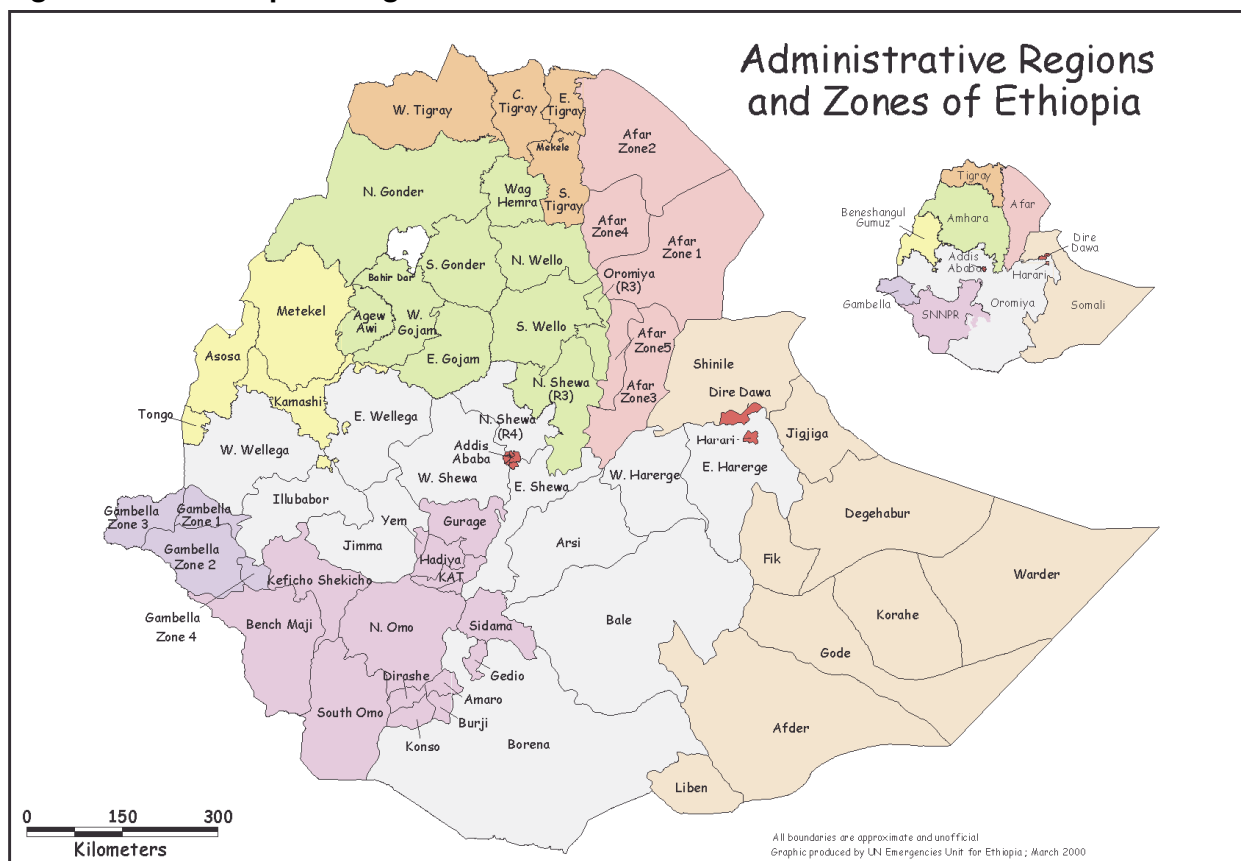
2.1 Research Design

The size of the Ethiopian EGRA allowed for the most complex and representative research design for any EGRA in Sub-Saharan Africa or elsewhere in the world. The initial stage, explained below, was to allow the MOE to determine the selection of regions. This means that any analysis made is representative to the region itself, rather than to the entirety of Ethiopia. This section then presents in some detail the development of the individual EGRA assessments. While EGRA is an internationally accepted assessment, at its core is an understanding that language and reading is primarily a local phenomenon. This is followed by a description of the piloting and revision processes, which were very systematic in the Ethiopian case. Then the section closes with some discussion of the analytic strategies employed in this study. While this is not meant as an overly technical report, some discussion is made here about the sampling design and analytic methods employed, particularly multiple regression methods.

2.2 Regional Selection

An important first step in this EGRA project was selection of the regions for the study. Given that EGRA was to be used for dual purposes—to provide an assessment of reading levels for a significant portion of the Ethiopia population, and also to provide a baseline for the IQPEP—a large geographic and regional spread was necessary. The MOE chose a panel to explore and debate the possibilities. The panel decided that Tigray, Amhara, Oromiya, Somali, Benishangul-Gumuz, SNNPR (Sidama zone), Harari, and Addis Ababa would be chosen. These regions cover over 96% of Ethiopia’s population and include a significant amount of language and cultural diversity. There is hope that future work will include the remaining three regions and additional language groups in SNNPR.

Figure 3. Ethiopia's Regions



2.3 Instrument Adaptation

To ensure that EGRA instruments are appropriate for the Ethiopian contexts and each of the selected languages, significant preparation was necessary. This included several critical items. First, several language consultants were identified from Addis Ababa University and regional academic institutions to prepare some of the linguistically more complex portions of the assessment for review before the assessment development workshop. Second, the Grade 2 and Grade 3 textbooks were obtained from the regions for each of the 6 languages. Then a representative sample of pages were typed and examined in Microsoft Excel to determine the most frequently occurring words and letters. This proved a challenge for the languages that used the Sabeen script (Amharic, Tigrigna, and Hararigna), because Excel is not yet capable of ordering the fidel alphabetically. However, it is an important task to ensure that the tasks reflect the textbooks and language complexity that children are exposed to in classrooms. Finally, the instruments were adapted to the Ethiopian contexts during workshops in May 2010. The Amharic, Tigrigna, and Afan Oromo tasks were developed in one week-long workshop, and Hararigna, Somali, and Sidamigna in another workshop.

This adaptation workshop was attended by experts from the ETQAA, the Assessment sub-process, the Curriculum process, the Planning process, the Teacher Education

process, several other portions of the MOE, UNICEF, the World Bank, Save the Children USA, International Rescue Committee, USAID, IQPEP, Addis Ababa University, the Regional Education Bureaus (REBs) from all 6 regions, and the relevant woreda education offices. In addition, an international reading consultant, Dr. Sandra Hollingsworth, supported the development of the reading components of the assessments. The local experts at that panel worked hard to ensure that the instruments matched local realities, developing appropriate stories; creating word frequency, letter frequency, and syllable frequency lists from actual textbooks used in the regions; and reviewing pretest data. These workshop participants were very hard-working and committed to the production of leveled instruments capable of identifying the true status of early reading skills in the 6 languages of interest for the study: Amharic, Tigrigna, Afan Oromo, Sidamigna, Somali, and Hararigna. They were also involved in the pretests of the instruments in the field.

2.4 Training and Piloting

The assessors were trained during a two week training session held from 19-23 April, 2010 and 10-14 May, 2010. Both RTI and IQPEP data collectors were involved in the training.⁷ Each assessor was given four interrater reliability tests and the lowest scoring assessors were not allowed to be a part of the pilot training, nor the full data collection. Final interrater reliability scores were higher than 94% for the entire group of assessors, which is extremely high, but similar to what was found in Kenya and Uganda for EGRA studies there.

In order to test the reliability and validity of the various subtasks in the 6 languages, a full pilot took place in Amhara and Oromiya regions 26-29 April, 2010, and in Tigray, SNNP, and Somali regions 17-20 May, 2010. Assessing children in several chosen schools in each region, pilot findings were analyzed from 77 Sidamigna assessments, 78 Tigrigna assessments, 90 Somali assessments, 90 Amharic assessments, and 105 Afaan Oromo assessments.⁸ The pilot data were entered the same day they were collected using an RTI-developed Excel-based data-entry system. These data were cleaned and coded immediately and sent to an RTI psychometrician. Findings from the pilot psychometric

⁷ IQPEP assessors were primarily from within the education sector, housed at Woreda Education Offices, Colleges of Teacher Education and other education offices. RTI assessors were primarily data collectors engaged in other assessment work in the system, but also included educators of various sorts. There were no statistically significant differences in assessment outcomes for the two groups across the 8 regions, except for Amhara, likely due to the very different woredas assessed. This shows that the fidelity of the assessment was not adversely impacted by the assessor selection, and the EGRA results remain robust.

⁸ Note that some colleagues point out that the EGRA assessment was not done by teachers. While this does present some challenges, in that the assessors were in some cases not pedagogical experts, it also makes it less likely that the teacher assisted the child with the assessment. Our findings in other countries suggest that teachers are often very weak assessors with respect to reliability and validity.

Rasch analysis that had implications for portions of each language assessment were adapted appropriately. Many of the changes related to particular items that were more or less discriminating than they should have been, and the response was to confer with language experts and assessors to determine how best to improve the assessments. Changes were made to improve each of the instruments before the full data collection and were included in the updated assessment versions. In nearly all cases, the changes necessitated by the pilot results were cosmetic; yet the exercise was important to determine items that were inappropriate.

2.5 Sampling

To ensure regional representativeness, RTI, with the support of the MOE, developed a sampling framework, employing three-stage stratified sampling, using proportional to population sampling at the regional and school levels and systematic sampling at the classroom level. We sampled from several woredas from within the 6 regions using the 2001 E.C. EMIS school data. Note that in Addis Ababa, two-stage sampling was employed because there was no need to sample at the woreda (or subcity) level. In Harari, sampling was done from the entire list of schools, stratified by language of instruction. In Somali, due to issues existing during the period of the election, convenience sampling was done at the woreda level and random sampling at the school level. Note that the IQPEP schools were sampled purposively; therefore the analysis below is stratified by IQPEP and RTI. Table 4 presents the number of sampled woredas and number of schools. Appendix A presents more detailed information about the sample of schools by woreda.

Table 4. Sampled Districts for EGRA Study, by Region

Region	RTI or IQPEP	Woredas	Schools
Tigray	RTI	6	25
	IQPEP	8	14
	Total	11	39
Amhara	RTI	8	40
	IQPEP	7	20
	Total	11	39
Oromiya	RTI	8	40
	IQPEP	7	23
	Total	15	63
Somali	RTI	4	25
	IQPEP	2	8
	Total	5	33

Region	RTI or IQPEP	Woredas	Schools
Benishangul-Gumuz	RTI	3	25
	IQPEP	3	10
	Total	6	35
Sidama	RTI	6	25
	IQPEP	5	17
	Total	11	42
Harari	RTI	8	22
	IQPEP	3	9
	Total	11	31
Addis Ababa	RTI	10	25
	IQPEP	4	8
	Total	11	33
Total	RTI	53	227
	IQPEP	39	109
	Total Sample	84	336

To ensure representativeness within regions, these random selections were kept as much as possible. In some cases, where EMIS data was inaccurate,⁹ replacements were made by the next school on the sampling list. The number of replacement schools differed by region.

2.6 Achieved Sample

Data collection took place in all 6 regions between May 10 and June 16, 2010. The woreda education officers, head teachers, and teachers were very cooperative and interested in the results. While many challenges existed, in particular the inconvenient timing of the data collection vis a vis the election, and the end of school examinations, the assessment data collection was successful. The student sample is summarized in Table 5. The table is disaggregated by grade and gender, as well as by RTI and IQPEP samples.

⁹ Several schools were noted as government schools, and were therefore part of the sampling frame, but the woreda education office informed the research team that the schools were actually private schools. In other cases,

Table 5. EGRA 2010 Sample, by Language, School Type, and Location

Region	Languages	RTI/IQPEP	Woredas	Schools	Grade 2 Boys	Grade 2 Girls	Grade 3 Boys	Grade 3 Girls	Children Assessed
Tigray	Tigrigna	RTI	6	25	245	250	248	249	992
		IQPEP	8	14	140	139	140	140	559
		TOTAL	11	39	385	389	388	389	1551
Amhara	Amharic	RTI	8	40	373	383	392	373	1521
		IQPEP	7	20	195	200	200	200	795
		TOTAL	15	60	568	583	592	573	2316
Oromiya	Afan Oromo	RTI	8	40	378	402	394	400	1574
		IQPEP	7	23	230	230	230	227	917
		TOTAL	15	63	608	632	624	627	2491
Somali	Somali	RTI	4	25	229	224	228	182	863
		IQPEP	2	8	81	79	86	74	320
		TOTAL	4	33	310	303	314	256	1183
Benishangul-Gumuz	Amharic	RTI	3	25	232	238	235	242	937
		IQPEP	3	10	92	86	85	89	352
		Total	6	35	314	324	320	321	1289
SNNP	Sidamigna	RTI	6	25	248	249	248	250	995
		IQPEP	5	17	190	189	188	190	757
		Total	11	42	438	438	436	440	1752
Harari	Hararigna	RTI		2	20	20	20	20	80
	Amharic	RTI		8	79	81	79	81	320
	Afan Oromo	RTI		12	124	101	119	102	446
	TOTAL	RTI	8	22	223	202	218	203	846
	Afan Oromo	IQPEP	3	9	87	77	95	80	339
		Total	11	31	310	279	313	283	1185
Addis Ababa	Amharic	RTI	10	25	242	256	248	252	997
		IQPEP	4	8	80	80	80	75	315
		Total	11	33	322	336	327	327	1312
EGRA	6 languages	RTI	53	227	2160	2204	2210	2151	8725
		IQPEP	39	109	1095	1080	1104	1075	4354
		Total	84	336	3255	3284	3314	3226	13,079

Note that the achieved sample was quite close to the planned sample. For example, for the RTI portion of the study, we planned to sample in 230 schools and 9200 students, and actually sampled 227 schools and 8725 students. The reason for the discrepancy in the number of students sampled was due to some grades and schools not having the sampled 20 students per grade level. In those situations, we sampled all available students. The following statistics show our success rate. The achieved school sample was 98.7% of the planned sample and the achieved student sample was 96.1% of the planned sample for the sampled schools.

In addition to the student data above, teacher and head teacher data were compiled for each school and relevant grade level, as indicated in the sample below.

Table 6. Head Teacher and Teacher Questionnaires by Region

Questionnaire	Sample	Tigray	Amhara	Oromiya	Somali	B-G	Sidama (SNNP)	Harari	Addis Ababa	Total
Teachers	RTI	79	91	109	36	59	125	61	69	629
	IQPEP	41	69	81	18	20	99	20	22	370
	Total	120	160	190	54	79	224	81	91	999
Head teachers	RTI	25	40	40	25	25	25	22	25	227
	IQPEP	14	20	23	8	10	19	9	8	109
	Total	39	60	63	33	35	44	31	33	336

Throughout this report, results are reported both for the population of students for each group of interest. The sample results, including frequencies and percentages, should be interpreted as representative of the students in each regional sample. Estimations, including means and regression results, allow for interpretation of results for the entire population of interest, as described in Table 7. The estimations of means and significance levels were calculated in STATA using the survey command (svy) to establish the parameters for each level of selection. Similar to other national assessments such as NLA, ours did not draw a simple random sample of the population of students in each group of interest, for cost and efficiency reasons. But to enable us to make inferences about the performance of the entire population and not just those sampled, we weighted our results.

Our data needed to be weighted because the sample design did not give each individual an equal chance of selection. If we did a simple random sample of all students in Ethiopia, we would have had to send the assessment teams to thousands of schools throughout the country. Instead we grouped students within schools, schools within woredas, and woredas within regions, and corrected for this grouping using weights. (The weights increase the power of the individuals who were sampled, making them represent

the estimated population within each group.) Based on the estimated total population and students in the final sample, we calculated a weight for each level of selection (woredas, schools, students) and for each observation (student). STATA takes this final student level weight to determine the best estimation of the results for the entire population of students. This weighting requires that instead of reporting the standard deviation (which tells the average of the difference from our sample mean) we must report a standard error, or the accuracy of our estimation (mean) for the population.

The sampling frame was defined as described in Table 7.

Table 7. EGRA 2010 Sampling Frame

Level	Sampling Unit	Stratified by
First	Woredas	
Second	School	RTI/IQPEP
Third	Students	Gender and grade

2.7 Analytic Strategies

The estimates and percentages provided here are from the weighted data, as specified above. This allows our estimates to be regionally (rather than sample-based) representative, and provide a significant improvement over other estimates, that are often cited, that are not in fact representative at the regional level. Note that the sample sizes achieved here by region are much larger than the expected levels for statistical significant differences at the grade level, as indicated in the EGRA toolkit (Gove, 2009). This means that the analyses in this report are well within what are expected to be conservative measures for identifying differences in reading outcomes.

Analyses are performed using t-tests and multiple regression techniques. This is important since simple comparisons are often made between groups without employing tests to ensure that any differences identified are statistically significant. This is done in three ways, first are the standard t-tests performed using the `ttest` command in Stata, or using post-hoc mean tests on weighted data using the `svy` command in Stata. Second are the standard multiple regression tests (including t-tests and other significance tests) performed in Stata, using the `reg` procedure. This is often also performed using weighted data, and to account for the nesting structure of the data, additional tests are performed using post-hoc significance tests. Given the corrections made for the nesting of the data and the corrections to the standard errors, we are able to show with confidence where the differences are statistically significant and where they are not.

Our ability to make comparisons is buttressed at the language level, particularly for Amharic and Afan Oromo, where more than one region is assessed using the same language. We refuse to make comparisons across languages (other than zero scores) given the fundamental language differences that make fluency scores uncomparable. On

the other hand, we are able to exploit the fact that some of the assessments were performed in the same languages, and we are able to see if the differences between regions are significant.

Note that the Findings section does not concern itself with whether there are statistically significant differences between regions. Such an analysis would be cumbersome and not particularly useful, since there is no interest in comparing between regions at this level. The more complex analyses, such as those found starting in Section 4.3 present “tested” mean scores by t-tests and multiple regression analyses. The basic scores are presented, but where differences are identified, they are ones that are statistically significantly different. The gender and grade comparisons in Section 4.5 and 4.6 are performed using multiple regression analysis and the differences are significant where noted. In Section 5, where predictive factors are presented, the entirety of the analysis using multiple regression. These models are simple ones, with the outcome variable compared against predictors presented in the section. In many of the models, grade and gender were also controlled for. In Section 7, the benchmarking section, more sophisticated regression techniques such as quantile regression and regressions within sub-samples were used. Where these differences are significant, they are noted. Then, in Section 7.3, scatterplots were employed using oral reading fluency as the predictor and reading comprehension as the outcome. These models explained very high percentages of variation, as the section explains. In short, the report employs standard statistical analyses in standard ways, and notes where models and particular variables were statistically significant. More details on the outcomes are found in Appendix A and B.

3. Descriptive Statistics

Descriptive statistics for the sample of children whose achievement data are described below are presented in Table 8. We find that the average age of children in the region is oldest in Somali (Grade 2, 10.6 years, and Grade 3, 11.6 years old), and youngest in Harari (Grade 2, 8.9 years old, and Grade 3, 9.8 years old). When the percentages of children whose home language matches the school language are compared across regions, the percentages are high in Tigray (96.7%), Amhara (93.6%), and Sidama zone (97.2%). The lowest percentage of children with a match between their language and the language of the school is found in Benishangul-Gumuz (71.5%) and Somali (76.9%). Even in Harari, where the schools teach in three languages, a relatively low percentage of children claim that the school language is the same as their home language (82.9%).

3.1 Home Background

The next several items asked children if their houses had a variety of commodities in order to estimate the relative family wealth. The percentage of children responding that their families had a radio ranged from 88.5% in Addis Ababa to 40.8% in Benishangul-Gumuz. The percentage of families with a mobile phone or landline ranged from 88.9%

in Addis Ababa to 10.2% in Benishangul-Gumuz and 10.7% in Amhara. The percentage of families that had electricity ranged from 12.5% in Harari to 85.5% in Addis Ababa. For televisions, the percentages ranged from 3.1% in Amhara to 78.1% in Addis Ababa. The percentage of children with televisions in Oromiya was surprisingly high (26.1%). The percentage of children with access to toilets either inside or outside of the house ranged from 37.5% in Sidama zone to 89.6% in Harari. Bicycles were relatively rare, with rates ranging from 1.1% in Tigray to 17.1% in Sidama. Motorcycles and cars were also low, with Harari having the highest percentages of both item (6.1% and 11.3%, respectively). Significant percentages of the children sampled had animals, with Tigray, Amhara, Oromiya, and Sidama all having more than 80%, while Addis Ababa had only 12.7%. The number of animals that a family had ranged from 0.9 in Addis Ababa to 12.1 in Tigray. Some regions had more children attending preschool or kindergarten, with rates in Addis Ababa (69.1%) and Harari (46.4%) particularly high. On the other hand, low rates for preschool/kindergarten were reported for Somali (10.0%), Amhara (5.9%), and Benishangul-Gumuz (11.3%).

3.2 Other Background Items

In addition to the wealth questions, children were asked other indicators about their families and school backgrounds. The findings showed that repetition rates were highest in Somali (7.0%) and lowest in Addis Ababa (1.7%). The textbook ratios were also quite different. Only 41.7% of children in Somali said that they themselves had the language textbook, while textbook rates were above 90% in Harari (92.7%) and Tigray (94.4%), and almost 90% in Oromiya (89.6%) and Addis Ababa (89.8%). Many fewer children had any other reading books at home, with Oromiya (17.8%), Somali (17.3%), Benishangul-Gumuz (17.2%), and Sidama (16.1%) reporting particularly low rates. Mother's literacy varied quite a bit with low rates in Somali (22.7%), Amhara (27.2%), and Benishangul-Gumuz (27.2%). High rates of father's literacy were found in Harari (73.7%) and Addis Ababa (72.5%).

Table 8. Descriptive Statistics of Self-reported Figures by Region

Item	Tigray	Amhara	Oromiya	Somali	B-G	Sidama	Harari	Addis Ababa	Total
Grade 2 Age (average)	9.1	9.7	9.6	10.6	9.3	9.2	8.9	9.4	9.5
Grade 3 Age (average)	10.2	11.1	10.8	11.6	11.0	9.9	9.8	10.4	10.6
Home Language =School Language (%)	96.7	93.6	87.8	76.9	71.5	97.2	82.9	88.4	88.2
Has Radio (%)	57.2	51.8	70.0	47.6	40.8	55.0	80.2	88.5	61.1
Has Phone (%)	15.3	10.7	43.7	29.5	10.2	15.2	72.5	88.9	34.1

Item	Tigray	Amhara	Oromiya	Somali	B-G	Sidama	Harari	Addis Ababa	Total
Has Electricity (%)	36.4	25.5	15.0	37.1	30.8	18.0	12.5	85.5	96.3
Has Television (%)	5.6	3.1	26.1	16.2	4.7	7.3	70.7	78.1	23.3
Has Toilet (%)	85.0	74.5	87.6	37.9	63.7	37.5	89.6	83.4	71.2
Has Bike (%)	1.1	4.0	15.0	4.7	6.2	17.1	14.5	4.9	8.8
Has Motorcycle (%)	0.4	0.8	3.3	2.3	0.7	4.2	6.1	0.7	2.3
Has Car (%)	0.5	0.8	3.3	3.6	1.3	1.3	11.3	9.1	3.3
Has Animals (%)	84.4	88.8	81.0	73.5	76.0	80.1	44.4	12.7	71.3
Number of Animals (average)	12.1	10.3	7.6	7.9	4.3	5.2	3.1	0.9	6.9
Went to Preschool/KG (%)	12.7	5.9	22.1	10.0	11.3	18.5	46.4	69.1	22.4
Was Absent (%)	25.2	40.3	26.7	21.3	21.0	64.4	15.1	13.3	30.5
Repeater (%)	3.2	5.0	3.2	7.0	5.4	5.8	5.1	1.7	4.4
Had Textbook (%)	94.4	69.6	89.6	41.7	59.4	65.0	92.7	89.8	76.4
Has Other Books (%)	31.3	25.5	17.8	17.3	17.2	16.1	43.0	40.3	25.0
Mother Is Literate (%)	65.3	27.2	40.6	22.7	27.2	35.1	58.0	55.2	40.5
Father Is Literate (%)	30.8	54.8	62.0	39.3	56.2	63.5	73.7	72.5	56.7

4. Findings

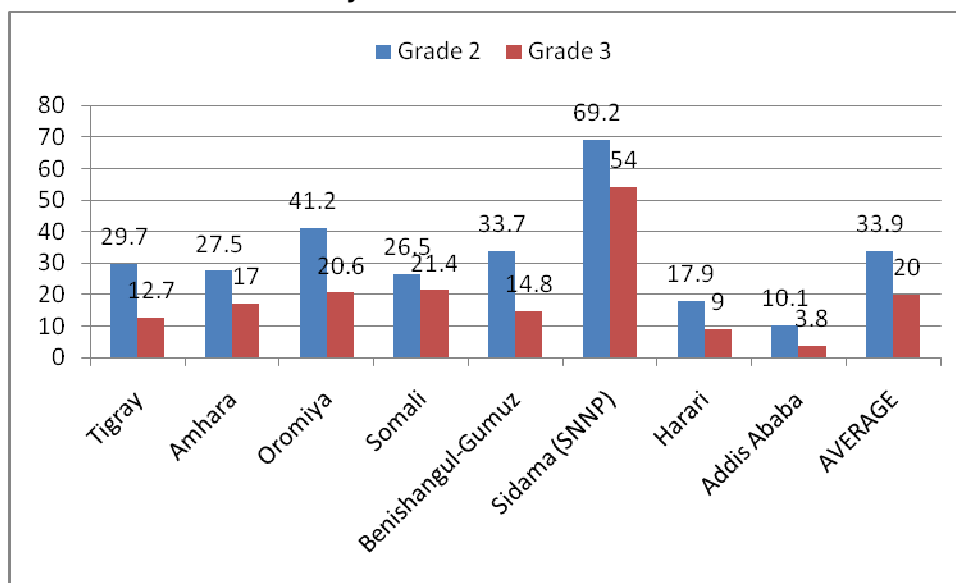
4.1 Early Reading Achievement Is Low

This section presents findings on the regional levels of reading. Of course, as each language in this study is different, comparisons are not easily made among them. Therefore, sections for each region present much more specific findings.¹⁰ While comparisons among various levels of reading are impossible, it is possible to identify the numbers of children who are unable to complete particular portions of tasks, and this section provides these comparisons.

¹⁰ These regional sections are found in the regional annex.

Each region sampled showed evidence of very low reading achievement. Figure 4 presents the percentage of children in each region who were unable to read a single word of a simple story targeted at Grade 2. It shows that in Grade 2, between 10.1% (Addis Ababa) and 69.2% (SNNP) of the children in these 8 regions could not read one word in Grade 2. The larger regions of Tigray (29.7%), Amhara (27.5%), Oromiya (41.2%), and Sidama (69.2%) had the highest percentages of children unable to read at all at the end of Grade 2. The scores for Oromiya and SNNP in particular were concerning, more so when we note that 54.0% of SNNP and 20.6% of Oromiya children were unable to read anything, even at the end of Grade 3. The region with the fewest zero scores was Addis Ababa, and Harari had the next fewest at both Grade 2 and Grade 3. The bar marked Average should be considered only as illustrative, and not as a proxy for Ethiopia, because no effort was made to weight the regional selections. If we had created an Ethiopia level in the weighting framework, it appears that the percentages of zero scores would be higher than this figure shows, because the lowest performing regions are the population-heavy ones of Oromiya, Amhara, Tigray, and Sidama (SNNPR).

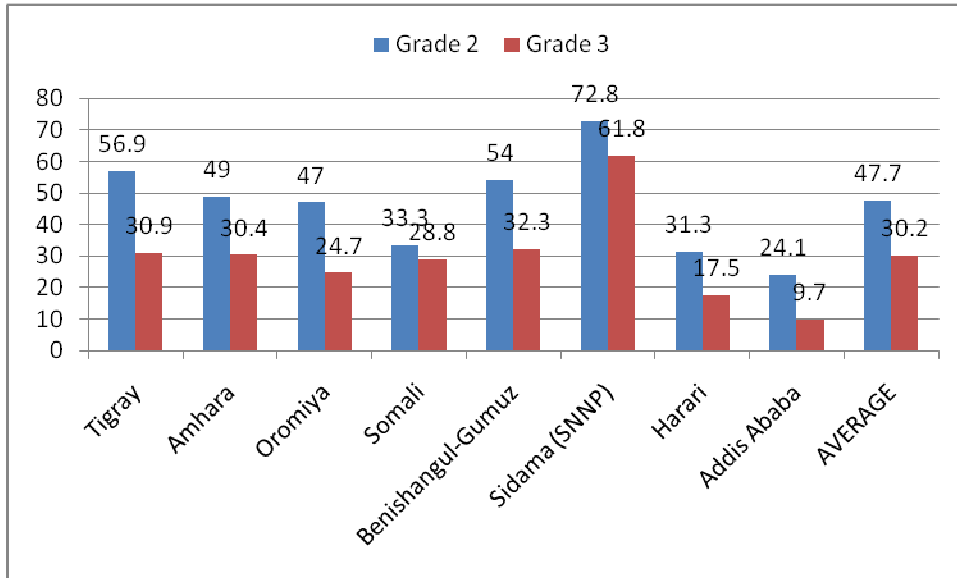
Figure 4. EGRA Zero Scores by Grade



The low achievement is not only evident in a child’s ability to read words fluently, but is even more stark in the percentages of children who were unable to correctly answer a single word of the reading comprehension test, as is evident in Figure 5. This table presents more frustrating results about the levels of comprehension across the regions. In Tigrinya, Amhara, Oromiya and SNNP, nearly 50% (or in the case of SNNP, nearly 75%) of children were unable to correctly answer one comprehension question. Note that the comprehension questions were written quite simply, with 4 of the 5 comprehension questions for each language serving as factual recall directly from the story. If

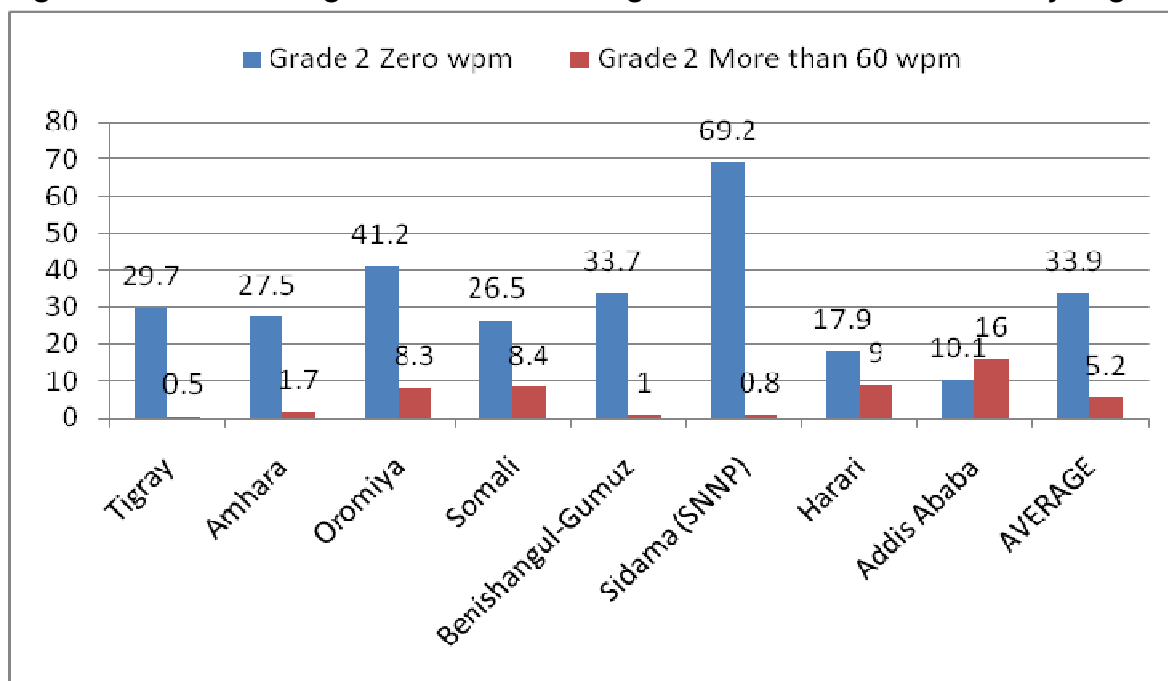
comprehension is the ultimate goal of reading, then it is remarkable to note that such high percentages of children are unable to comprehend simple stories.

Figure 5. Percentage of Children Scoring Zero on Reading Comprehension



As the regional analysis sections explain, the underachievement is similar even if the languages and regions are different. Each regional section will present the region’s scores on each of the subtasks. In this section, however, another comparison is useful. The U.S. reading benchmarks are not applicable to the Ethiopian context, and certainly not to the diverse languages in this study. In fact, part of this report focuses on the development of local benchmarks for Ethiopia’s languages. However, the U.S. and international benchmarks do shed some illustrative light on where Ethiopia is in the area of reading. Figure 6 shows the percentage of Grade 2 children who were reading at least 60 wpm (wpm), the absolute lowest benchmark for reading difficulties in the U.S., as well as the number of children who were reading zero words. Figure 6 presents some sobering facts. Only in Addis Ababa did more children read 60 words than 0 words, and that comparison was quite close (16.0% against 10.1%). Note that this is a grade where the majority of children should be reading at least 60 wpm. Of the other regions, only Harari, Oromiya, and Somali had more than 8% of children reading more than 60 wpm. On the other hand, all of the other regions (save Harari) had 25% or more of children reading zero wpm. The comparison suggests that while a significant number of children in each region are nonreaders (scoring 0), very few in any region are fluent readers.

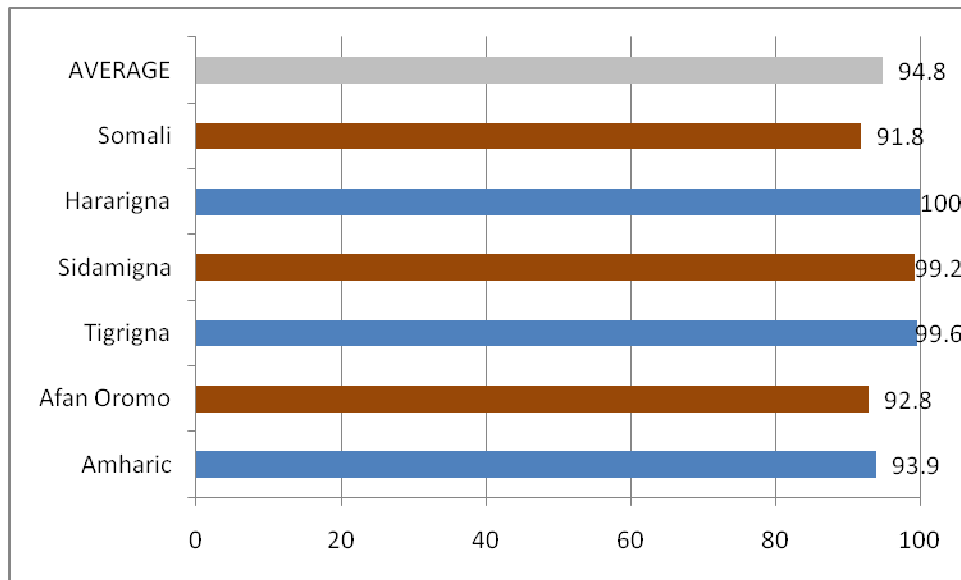
Figure 6. Percentage of Children Reading 0 and 60 Words in Grade 2 by Region



The final comparison in this section is the percentage of children in each language group who are meeting the minimum benchmark of 60 wpm. Note that this shows that the underachievement in reading is neither regionally specific nor limited to particular languages. Figure 7 presents the percentage of children assessed in each language who did not meet the benchmark of 60 wpm. The orange lines represent languages that use the Latin alphabet; blue lines languages that use the Sabean script; and the gray line shows the average of all language groups. The figure shows clearly that neither region nor language is sufficient for children to meet basic benchmarks in oral reading fluency, for in no language do more than 9% of children meet the benchmark. The reading deficiencies, therefore, exist in each of the languages assessed in this EGRA.¹¹

¹¹ Note that this figure includes both Grade 2 and 3, though the benchmark used is a Grade 2 benchmark.

Figure 7. Percentage of Children Not Meeting Benchmark

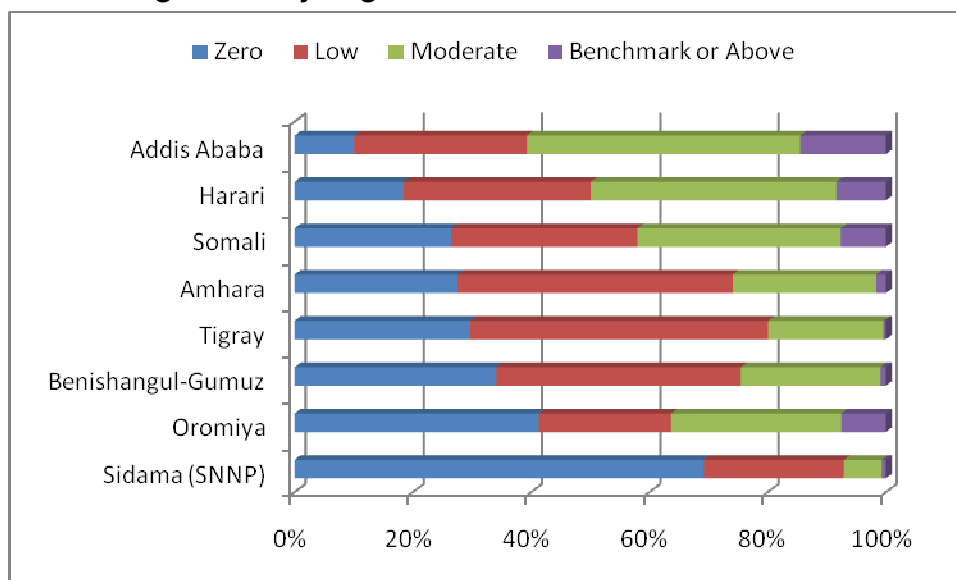


4.2 Overall Achievement on Oral Reading Fluency by Region

In the Figure 8 below, the regional achievement of children in the sample in Grade 2 is examined in more detail. The blue bars represent the zero scores; red bars show the percentage of children in the region with low oral reading fluency (between 1 and 29 wpm); green bars represent the percentage of children with moderate oral reading fluency (between 30 and 59 wpm); and purple bars indicate the percentage of children who read at least 60 wpm.

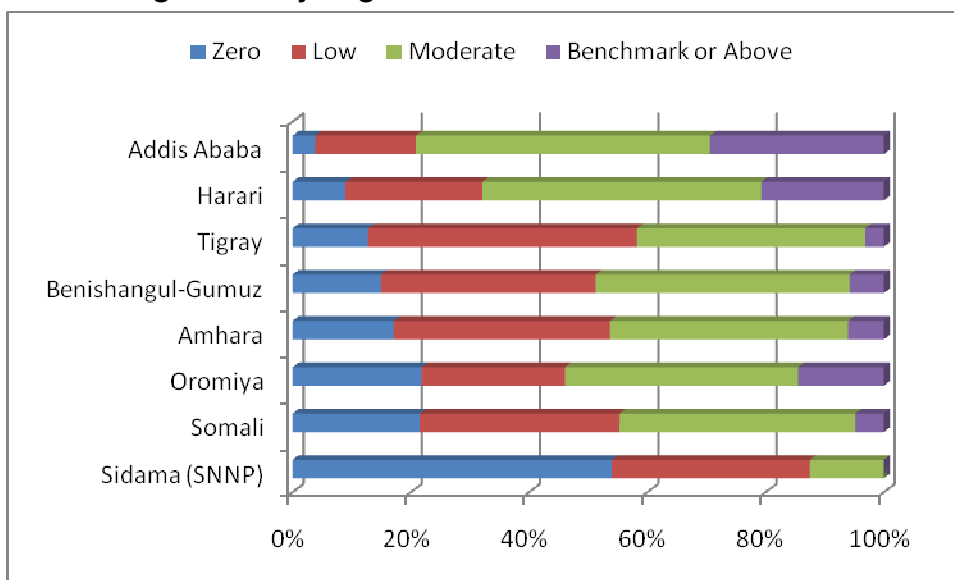
For Grade 2, the figure shows the underperformance of each region's children. In the highest scoring two regions, only 14.5% of Addis Ababa children and 8.2% of Harari children read at the international benchmark. In Sidama, Benishangul-Gumuz, and Tigray, the percentage of children reading 60 wpm was less than 1%. Only two regions had 50% of their children reading at least 30 wpm, Addis Ababa (70.0%) and Harari (58.1%). The regions with the largest percentage of children scoring less than 30 wpm are Tigray (68.7%), Benishangul-Gumuz (58.8%), and Sidama (90.3%). Interestingly, while Oromiya has a very large proportion of Grade 2 children who were not able to read at all (30.5%), a comparatively large percentage of children read 30 wpm or more (43.8%). This seems to indicate a tremendous amount of regional variation within Oromiya. In summary, this analysis shows that the distribution of Grade 2 reading skills is heavily skewed toward the low end of reading fluency.

Figure 8. Reading Levels by Region for Grade 2



To continue with the analysis above, Figure 9 provides the breakdown of reading scores in Grade 3 across regions. It makes clear that children were reading more fluently in Grade 3 than in Grade 2. The decreases in zero scores (Grade 2 zero scores – Grade 3 zero scores) range from 5.1% (in Somali) and 20.6% (Oromiya), except in Harari and Addis Ababa, where the decreases were smaller (because the number of zero scores in Grade 2 were much smaller). In short, in Grade 3 a significant percentage of children who could not read at all learned at least the basics of reading, particularly in Oromiya region. Similarly, there was a moderate increase in the number of children who could read 60 wpm. The gains in Grade 3 are significant, yet they show that children in Grade 3 remain, for the most part, quite far from reading fluency. Note that in SNNP, a full 87.5% of Grade 3 children read less than 30 wpm. In Tigray, 58.0% of children read less than 30 wpm. There remains, even at the end of Grade 3, a significant amount of progress necessary for children to read with sufficient fluency to ensure comprehension.

Figure 9. Reading Levels by Region in Grade 3



4.3 Regional Comparisons for Amharic and Afan Oromo

While it is impossible to compare reading outcomes across regions, due to the language differences, it is possible to compare regions that used the same language assessment. For Afan Oromo, both Oromiya and Harari regions used the same assessment. For Amharic, Amhara, Benishangul-Gumuz, Harari and Addis Ababa regions used the same assessment. This section uses multiple regression analyses to determine, for each task, which regions scored higher. Table 9 presents the scores for each region by the subtask.¹² When comparing within each task, the pattern is consistent. Harari students (assessed in Amharic) outperformed Addis Ababa students. Both of those regions outperformed students in Amhara and Benishangul-Gumuz.

¹² Post hoc general linear hypothesis tests were performed to determine whether the differences among regions were statistically significant. For each comparison, there is no statistical difference between Amhara and Benishangul-Gumuz. For phonemic awareness, the differences between Amhara and Harari, Amhara and Addis Ababa, and Harari and Addis Ababa are not statistically significant. All other comparisons are different in the direction shown in Table 9.

Table 9. Subtask Comparisons for Amharic Assessed Regions

	Schools	Students	Fidel Fluency	Phonemic Awareness	Word Fluency	Decoding Fluency	Oral Reading Fluency	Reading Comp.	Listening Comp.
Harari	8	320	92.9	82.5	52.5	28.2	46.4	50.9	69.2
Addis Ababa	33	1312	76.1	82.9	46.3	25.1	40.9	43.7	69
Amhara	60	2316	47.7	73.6	24.7	15.5	23.4	28.5	55
Benishangul-Gumuz	35	1289	37.5	61.2	20	14	22	24.2	55

Differences in regional achievement by subtask are evident in this Figure 10. Gaps between the urban regions (Harari and Addis Ababa) and Amhara/Benishangul-Gumuz were quite large, with fidel fluency as an example. Children in Harari and Addis Ababa were about twice as fluent with their fidel as the average children in Amhara and Benishangul-Gumuz. The gap was less wide for phonemic awareness, as all groups scored above 60% correct. It appears that Harari and Addis Ababa were both 20% more accurate with phonemic awareness than Benishangul-Gumuz. For word reading fluency, children in Harari and Addis Ababa were twice as fluent with identification of words as they are in the other two regions. The comparison is nearly as stark in decoding fluency, as well. For oral reading fluency, we find that Harari children were twice as capable at reading stories fluently as are those in Amhara and Benishangul-Gumuz. Addis Ababa and Harari children were 70% and 100% more fluent than children in Amhara, respectively.

The comparisons were also wide for reading comprehension (as is expected, given the differences in oral reading fluency). For listening comprehension, the gap was much more limited, with Amhara and Benishangul-Gumuz both scoring 55% correct. Addis Ababa (69%) and Harari (69.2%) scored only 14% higher. This has interesting implications. It appears that with respect to listening skills and oral vocabulary, there are only small differences between Amhara/Benishangul-Gumuz and Harari/Addis Ababa. The differences are large at the fidel fluency level, so from the initial skill that children learn, the gaps between the two sets of regions starts at the beginning of primary school. Those gaps exist throughout the rest of the tasks, where the urban regions often double the performance of Amhara and Benishangul-Gumuz. This seems, then, to be a story of initial differences (fidel identification fluency) continuing and expanding over time.

Figure 10. Subtask Comparisons for Amharic-speaking Schools

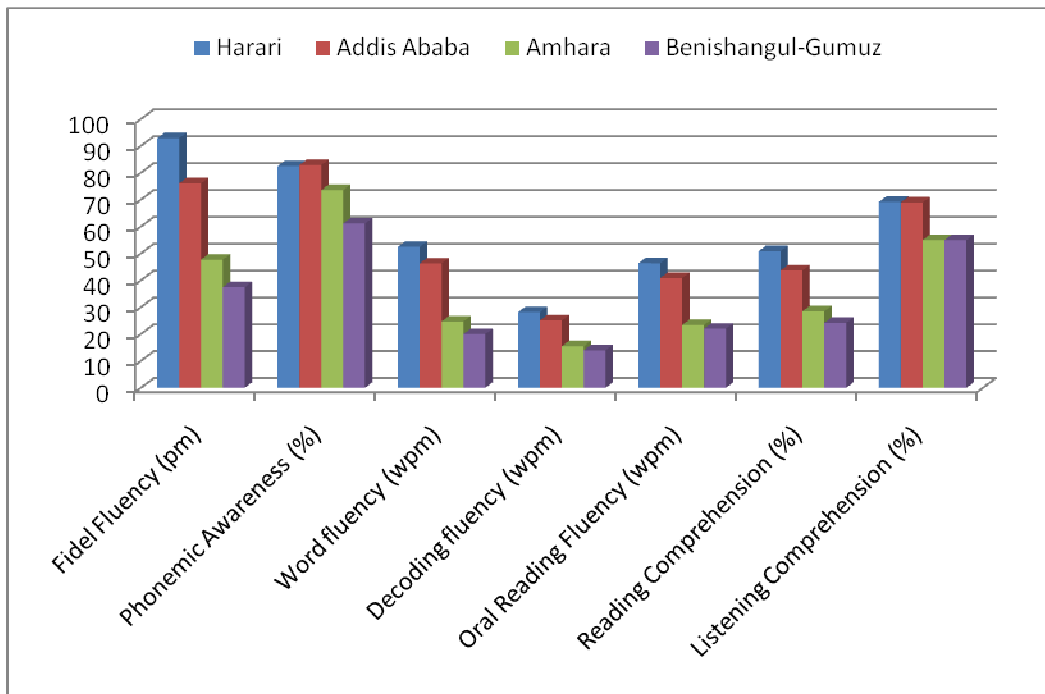
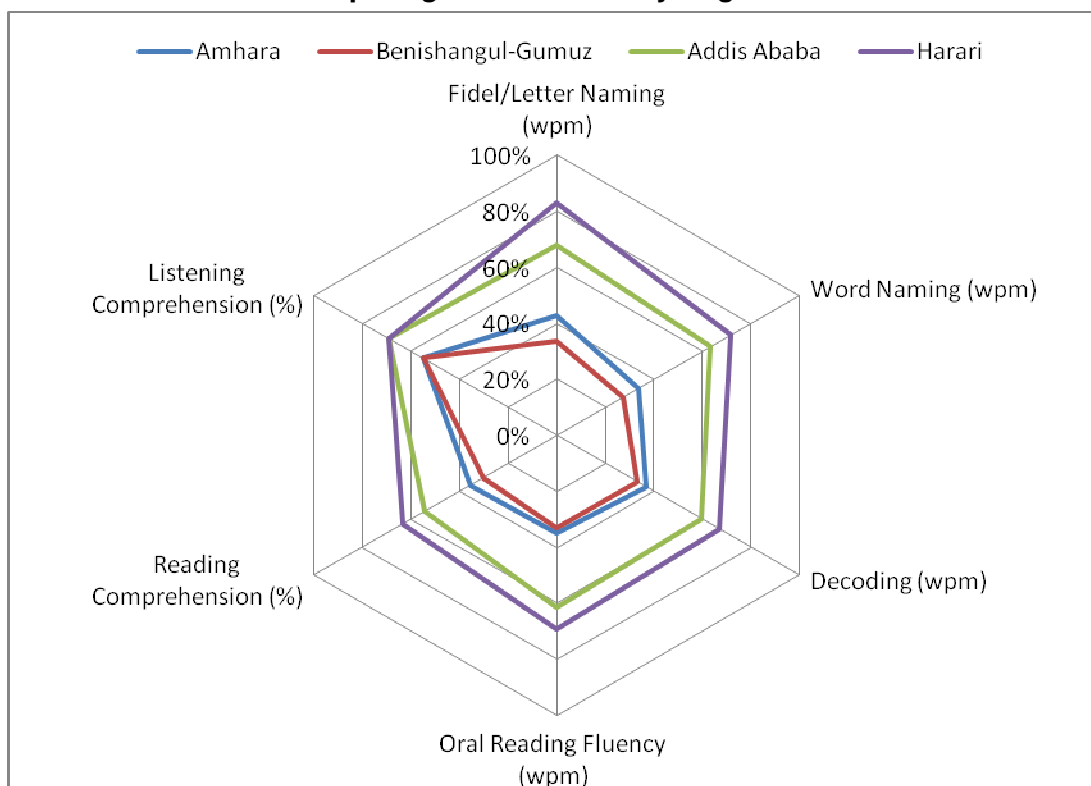


Figure 11 presents the achievement of a particular region assessed in Amharic against the benchmarks for that language. This allows us to investigate where the strengths and weaknesses of each region are, compared against a benchmark for the entire language group. The findings are clear. Children in Benishangul-Gumuz and Amhara are achieving lower than those in Addis Ababa or Harari. In particular, while listening comprehension scores are moderate in all four regions, and the letter naming fluency scores are reasonable in Amhara and Benishangul-Gumuz, the word level tasks (word naming, decoding, oral reading fluency) are all very low. In other words, while the children in those regions have some limited skills in identifying the fidel, combining the fidel to make words proves much more difficult, which shows itself in every word naming task. For Harari and Addis, the scores are closest to the benchmark for letter naming, but still far from it, and especially for reading comprehension.

Figure 11. Radial Plot Comparing Achievement by Region for Amharic



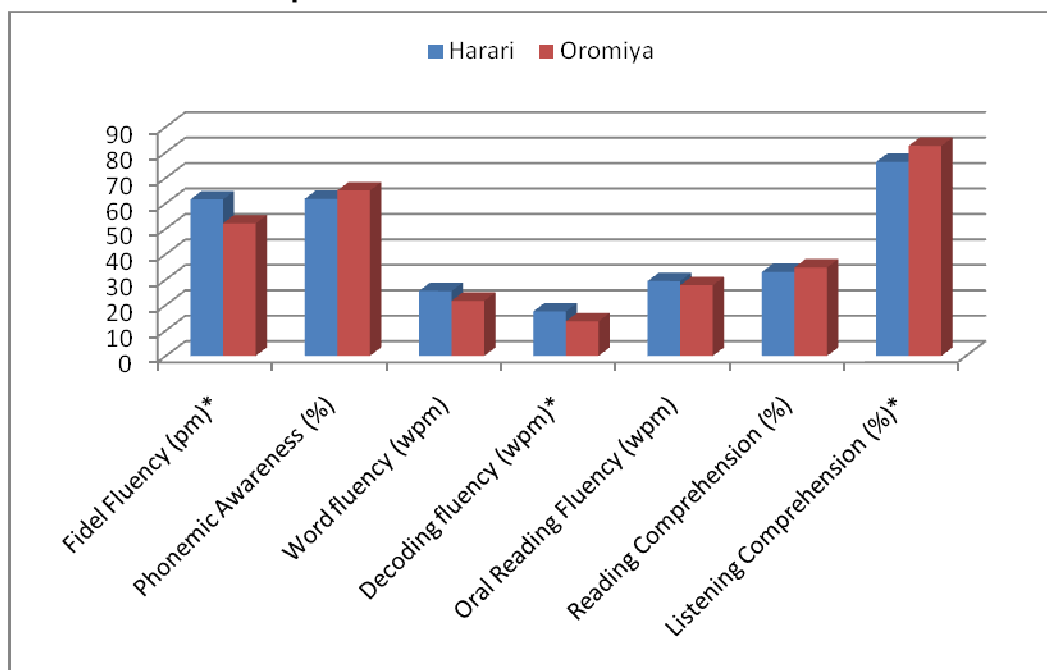
It is possible to compare the achievement of Harari children and Oromiya children on Afan Oromo reading. Table 10 presents the average scores for each task. Not all of the differences in this table are statistically significant, but the gap between fidel fluency (favoring Harari by 9.4 fidel per minute), decoding fluency (favoring Harari by 3.9 wpm), and listening comprehension (favoring Oromiya by 6.3%) are significant. The patterns here are much less consistent than the relationship between Amharic-speaking regions. In some cases Harari outperforms Oromiya, and in other cases, Oromiya does better than Harari. Listening comprehension is higher in Oromiya, but the skills of fidel fluency and decoding fluency are higher in Harari. This suggests that perhaps the oral language facility is more developed in Oromiya, but the pedagogy in Oromiffa-speaking portions of Harari are more sophisticated. More analysis is necessary to understand why these gaps exist and what can be done to ameliorate them.

Table 10. Subtask Comparisons for Afan Oromo-assessed Regions

	Schools	Students	Fidel Fluency	Phonemic Awareness	Word fluency	Decoding fluency	Oral Reading Fluency	Reading Comp.	Listening Comp.
Harari	12	785	61.5	61.7	25.5	17.5	29.4	33	76.3
Oromiya	63	2491	52.1	65	21.7	13.6	27.8	34.8	82.6

Figure 12 presents the comparisons between Harari and Oromiya students on Afan Oromo. For the most part, the substantive differences are small between the 2 regions, on average, with not many of the differences between scores by region being statistically significant. It is notable that the biggest gap by percentage is in decoding fluency, with Harari children doing better. Decoding fluency is often the task most responsive to specific teaching strategies, so it might be that Harari teachers spend more time helping children to decode new words.

Figure 12. Subtask Comparisons for Afan Oromo Schools



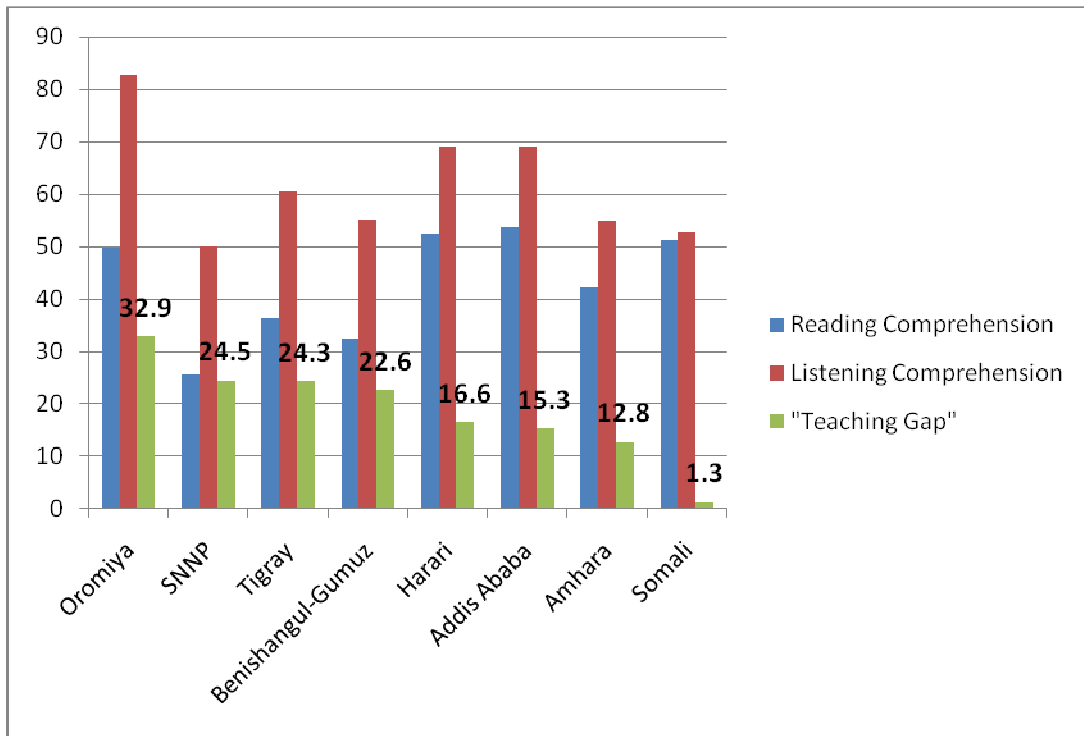
4.4 Comparing Reading Comprehension and Listening Comprehension Outcomes

It is important to determine the reason for the low achievement across the 8 regions in the study. In this section, the relationship between reading comprehension and listening comprehension is examined. Theoretically, if children understand what they read with ease, their scores on reading comprehension and listening comprehension should be very similar. Likewise, since listening comprehension examines the extent of a child’s oral language skills (including vocabulary, grammar, comprehension, and synthesis), it stands to reason that the difference between reading and listening comprehension is not due to oral skills, such as oral vocabulary. The Ethiopian curriculum laudably focuses on listening as part of the target of its learning competencies. It appears that the gap between reading and listening comprehension can be attributed to the skills a child receives primarily at school—such as a child’s ability to identify letters, combine letters to make words, read with speed and accuracy, decipher with meaning the words, and read

sentences read fluently. A simple analysis, then, allows us to estimate the “teaching gap.” It can be loosely identified as the gap between listening and reading comprehension scores, and due to children’s inability to match their oral skills with their reading skills.

In Figure 13, the reading comprehension scores (percent correct out of up to 5 questions attempted), are presented in blue bars, and the listening comprehension scores (percent correct out of 5 questions attempted), are presented in red bars.¹³ The gap between listening and reading comprehension scores are shown as green bars. The regions are organized in order from the region with the largest gap between listening and comprehension (Oromiya – 32.9%) to the smallest gap (Somali – 1.3%). Oromiya, Sidama zone, Tigray, and Benishangul-Gumuz have the largest “teaching gap,” each above 20%. The implication of these large gaps is that children’s scores in these regions could increase quite significantly with better teaching.

Figure 13. Reading Comprehension, Listening Comprehension and Teaching Gap Scores

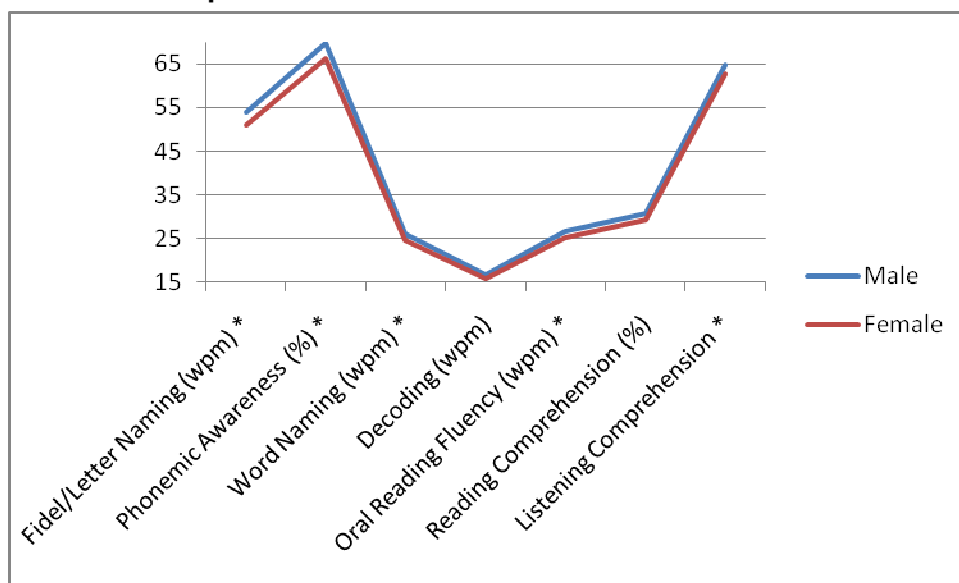


¹³ Note that the reading and listening comprehension tasks in Oromiya (and in parts of Afan Oromo-speaking Harari region) have 6 questions rather than 5. This has no impact on this analysis, since it is percentage-correct. Note that this analysis uses the reading comprehension scores for questions attempted rather than all 5 (or 6) questions.

4.5 Reading Outcomes and Gender

In most of the countries in Sub-Saharan Africa where EGRAs have been administered, RTI has found gender differences in reading outcomes. Differences also exist between urban and rural schools, as well as across grades. Unlike many other gender gaps in Sub-Saharan Africa education, the early reading scores identified by EGRA consistently favored girls. We examined the Ethiopia data to see whether there was a gender gap in achievement. Unlike in other countries, Ethiopian boys consistently outperformed girls on all 7 tasks. Figure 14 shows that the average scores for each task was slightly higher for boys. Note that the gender gap is statistically significant (noted by an asterisk) for all tasks except non-word decoding and reading comprehension.¹⁴

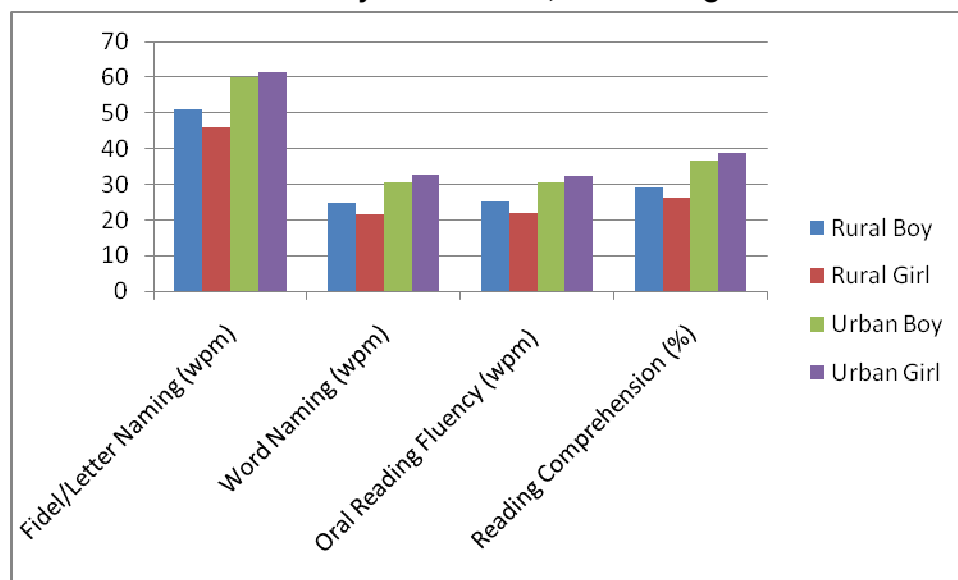
Figure 14. Gender Gap in EGRA Scores



However, the story is more complex than this simple gender comparison shows. Given that there might be differences in performance of boys and girls in urban and rural settings, multiple regression analyses controlling for region were performed to identify whether the gender differences were the same in urban and rural schools. Results are shown in Figure 15. While rural boys outperformed rural girls (on all tasks except for listening comprehension), the opposite relationship was true in urban schools. Urban girls outperformed urban boys on every task except for listening comprehension. In Ethiopia, it appears that rural girls have lower achievement than girls in urban areas and much lower than girls in other African countries.

¹⁴ The magnitude of the gender gap ranges from .05 standard deviations for reading comprehension to .09 standard deviations for phonemic awareness and fidelity/letter naming fluency. Therefore the effect size of the gender gap is quite small.

Figure 15. Gender and Urbanicity Interactions, Controlling for Grade



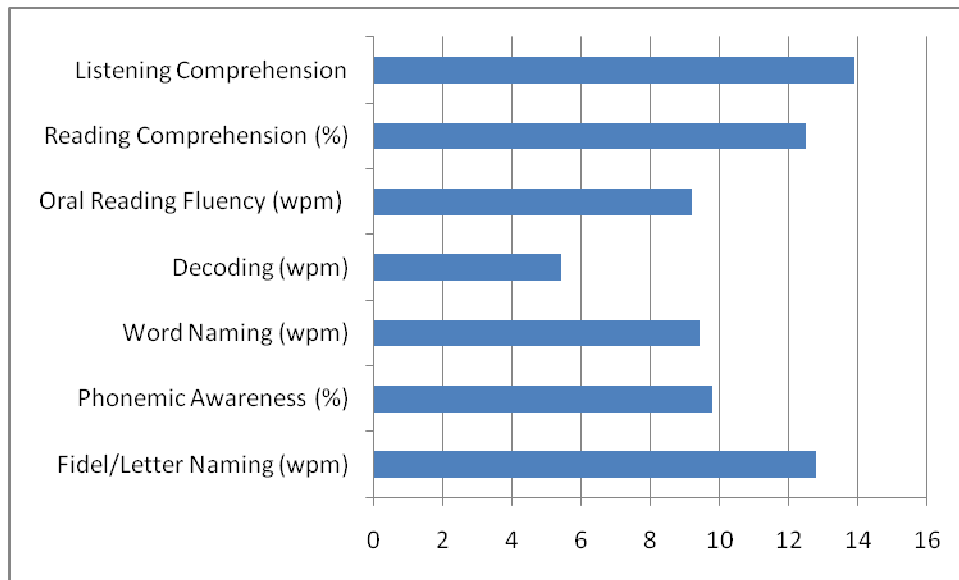
4.6 Grade Difference

The data presented here show how much reading skills children acquire from the end of Grade 2 through the end of Grade 3.¹⁵ Figure 16 shows the average grade gain (controlling for language and region) in Grade 3. The first finding is that children are learning quite a bit in Grade 3. For example, the average child increases the number of letters they can identify by 12.8 per minute, the words they can name by 9.4 wpm, the words they can decode by 5.4 wpm, the words they can read in a story by 9.2 wpm, and their reading and listening comprehension scores by 12.5% and 13.9%, respectively. The wpm increases are slightly lower than what has been identified in other countries (9.2 wpm compared to 12). Compared to US DIBELS benchmarks, the increase in wpm read in Grade 3 is lower in Ethiopia (20 wpm in the US, 9.2 wpm in Ethiopia).

An additional comment is that it would be expected that most children should have mastered their alphabet/fidel by the end of Grade 2. The fact that the average child in Grade 3 can identify 12.8 more letters per minute than can children in Grade 2 shows that many children are not fluent at letter identification after two years of schooling, and spend significant time in Grade 3 still learning the fidel. For these children, this is of course a critical skill, but seems to be an inefficient use of time in Grade 3, when we would expect that children are mastering words, learning to read sentences and stories, and comprehending what they read. This provides support for the types of church schools that Ethiopia has had in abundance in earlier times and that ensured children mastered the fidel before enrollment in Grade 1, and certainly before Grade 3.

¹⁵ Note that this is under the assumption that the children in Grade 2 are not dramatically different from those in Grade 3. There is no way to assess this, given the cross-sectional (non-longitudinal) nature of the dataset analysed here. Without drastic interventions, however, this is a reasonable assumption.

Figure 16. Grade Difference (between Grade 2 and 3) by EGRA Task



4.7 Accuracy Analysis

When investigating literacy outcomes it is important to understand the differences among languages and how differences relate to approaches to improve the quality of reading outcomes. Languages are different, and Ethiopia is particularly language-diverse, with more than 80 languages. Ethiopia is even more language-complex than its neighbors given the different scripts used in writing the languages. Of the languages in this particular EGRA study, 3 use the Sabean script (Amharic, Tigrigna, and Hararigna) and the other 3 use the Latin alphabet (Afan Oromo, Sidamigna, and Somali). This section presents the accuracy scores of children in order to determine whether the language scripts present different implications for improving student achievement in reading. In order to produce accuracy scores, each item was analyzed to determine first whether a child attempted it, and if she did, whether she got it right. These scores are then converted to percentages of the items attempted, or how likely children were to get an item correct.

It stands to reason that, all things equal, children learning in languages using the Sabean script might be less accurate in their responses to letter identification, given the larger number of letters in the script. It also stands to reason that the difference between letter accuracy and word accuracy is likely to be larger for children reading with the Latin alphabet, since each word is a combination of more letters. Put another way, children using the Sabean script are more likely to decode after they learn the fidel successfully. Therefore, for languages that use the Sabean script, the ability to read words accurately is not likely to differ significantly from the ability to read the fidel accurately.

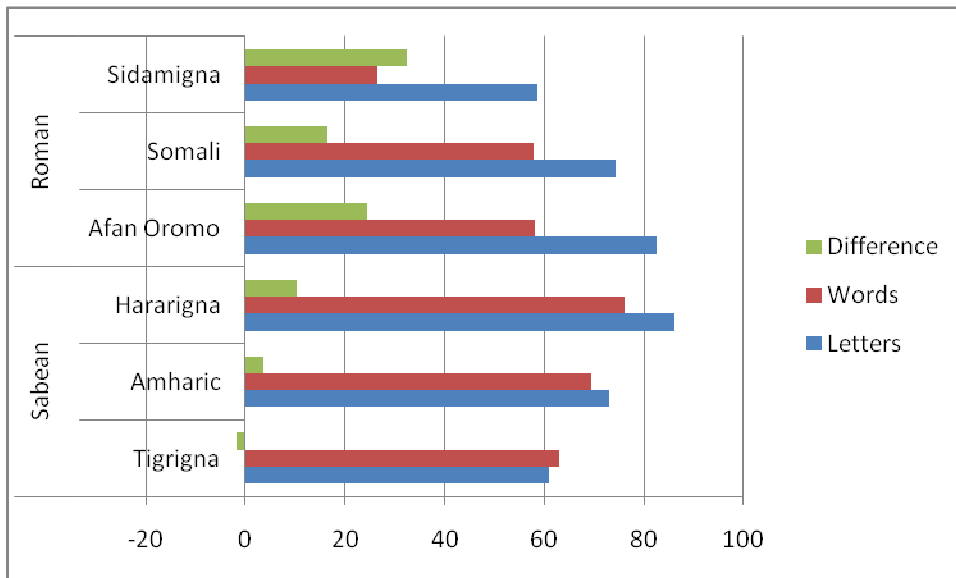
Figure 17 presents the findings regarding these questions. Scores for Sabean script languages (Tigrigna, Hararigna, and Amharic) are presented against scores for Latin alphabetic languages (Sidamigna, Somali, and Afan Oromo). The blue bars show that

children are particularly accurate at letter identification in Afan Oromo, with accuracy scores of 82.6%, and Hararigna (86.2%), and in Somali (74.3%). Children are less accurate in Tigrigna (73.0%) and Amharic (61.0%). This indicates that the first hypothesis is correct, that children reading in the Latin alphabet are more accurate at identifying letters.¹⁶

The red bars, on the other hand, show that the advantage that Latin alphabet learners have in identifying letters disappears when it comes to words. Word reading accuracy scores are very close between Afan Oromo (58.2%), Tigrigna (62.9%), and Amharic (69.4%). It is possible, therefore, that much of the controversy that has existed in Ethiopia regarding the benefits of various scripts is academic rather than practical, since for these three major languages, children can identify words at essentially the same level of accuracy. The exception, once again, is in Sidamigna, where children’s accuracy scores were an abysmal 29.1%. This again appears to be the result of low reading skills overall, rather than the impact of the language script itself.

Finally, it is of interest that the gap between letter and word accuracy scores (as indicated by the green bars) is negligible for Tigrigna (-1.9%) and Amharic (3.9%), but is quite substantial between Afan Oromo (34.4%) and Sidamigna (32.3%). It is as if children using Sabeen scripts have mastered the decoding skill once they have mastered the fidel, and those skills continue to serve them as they read words, while those using Latin alphabets require more instruction and practice to move from identifying letter sounds to combining those sounds to make words.

Figure 17. Letter and Word Accuracy Scores by Language, Including Difference Between Accuracy Scores



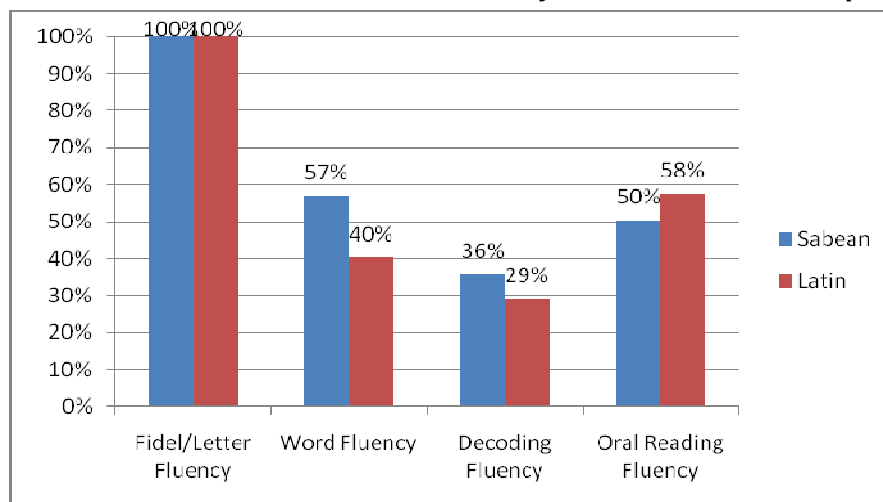
¹⁶ Note that this is not the case for Sidaamu Afoo at the aggregate level, which appears to be because children in Sidama zone have very low reading skills overall, starting with letter identification.

4.8 Relationships between Fluency Scores

The relationship between Sabean script and Latin alphabet can be analyzed another way. It is possible to investigate whether and how letter fluency is related to scores on other fluency tasks, namely word naming fluency, decoding fluency, and oral reading fluency. Figure 18 compares the relationships between these factors for Sabean script languages (Amharic, Tigrigna, and Hararigna) and Latin alphabet scripts (Somali, Sidamigna, and Afan Oromo). The word fluency, decoding fluency, and oral reading fluency scores are expressed as percentages of the scores on the letter fluency tasks.

Previous research has identified a linear relationship between these tasks in other countries that teach decoding and word attack strategies. In those countries, letter fluency scores are highest, followed by word fluency, then decoding, and finally oral reading fluency. This is not the pattern identified in the Ethiopian data, since decoding scores are so low. The drop from letter reading fluency to word reading fluency is extreme, particularly for Latin alphabet languages. It seems that while many children are very comfortable with the alphabet, they struggle with combining the letters into words, given that word reading scores are only 40% of what the letter scores are. The gap between fidel fluency and word naming fluency is smaller for the Sabean script languages (57%). One implication of this is that many children are simply letter readers and are not yet word readers. The decoding scores are quite different from the letter frequency scores. This shows, quite clearly, that children are not skilled in combining letters into new words. This is problematic for when children face any words to which they have not yet been introduced, and points to the importance of teaching decoding and word reading strategies.¹⁷

Figure 18. Correlations between EGRA Tasks, by Sabean or Latin Script



¹⁷ Surprisingly, oral reading fluency scores were more tightly correlated with letter fluency scores in Latin alphabet languages (58%). This might indicate that the stories on the assessments may have been too easy, since children were more able to read the stories than the familiar words.

5. Predictive Factors

Student, School and Family Level Predictive Factors

This section presents the results of multiple regression models to examine which factors predict achievement on reading outcomes. These models were fit at the regional as well as national level, and focus on the relationships between oral reading fluency outcomes and student level predictors. While dozens of models were fit, those presented here and in Appendix A, EGRA Scores by Region as well as the regional annex are the relationships that are statistically significant.¹⁸ Figure 19 presents the magnitude of the relationship on oral reading fluency and these predictors. The colors of the bars identify the “location” of the factor. Red signifies the family background characteristics; gold signifies the school and system level factors; and blue signifies the student characteristics.

Several patterns are evident from this figure. First, many of the predictive factors that are of the greatest magnitude are at the school level (having a textbook is 9.6 wpm) and system level (child attends an urban school is 13.2 wpm, the grade effect is 9.4 wpm). The implication is that schools and teachers matter a great deal.

Second, some of the significant factors are family characteristics that are out of the control of individual children, such as having a telephone or mobile phone (12.1 wpm) or electricity (9.2 wpm), or having other reading materials (8.3 wpm). That is, even if a child is extremely hard working, if she is an orphan, she is unable to have the entire family help her with her homework, and so she will not benefit from the very large increase in student achievement that having family support would provide. This becomes the responsibility of the school to help vulnerable children from less well off backgrounds to overcome their lack of family support. Moreover, some of the factors at the family level can be supported by the school. For example, the very large effect of the family helping with homework (that is the combined effect of fathers, mothers, and siblings helping, 14.7 wpm) could be encouraged by schools by providing training for families on how best to support children in their homework. Similarly, the family factor that determines whether a child has other materials to read (8.3 wpm) can be supported by the school system by providing simple readers to emerging readers’ families.

The student level factor that has the largest relationship with student achievement is attendance in kindergarten and/or preschool. This is clearly related to the MOE’s policy of supporting private kindergartens and the building of zero classrooms in schools. Given the extremely complex fidel that many children learn in Ethiopia, this suggests that the MOE continue providing kindergarten services to assist in the quick acquisition of the fidel. Similarly, for languages using the Latin alphabet, the important basic skills acquired by children in pre-school and kindergareten are (1) knowledge of the alphabet, and (2) the ability to combine letters to make words. Providing kindergarten and zero

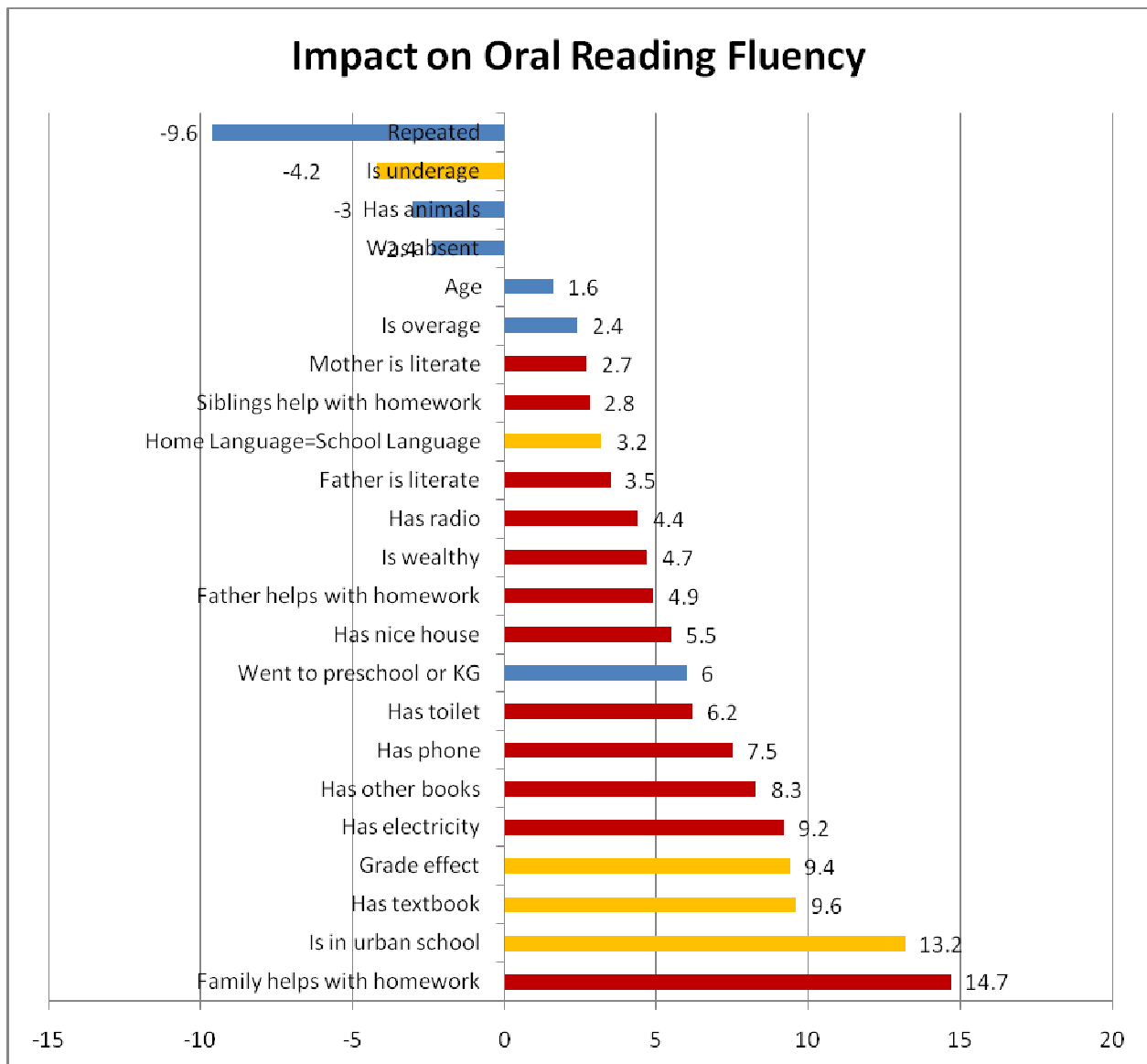
¹⁸ Significance is defined as the .10 level when the magnitude was large and the .05 level otherwise.

classes is, of course, very expensive, so encouraging the private sector development is a logical policy.

Two other school factors are of policy relevance. First, there is a large negative impact on student achievement of classroom repetition. Unlike in other countries where children who repeat outperform their counterparts on assessments (typically end-of-primary school assessments), the Ethiopian children in this study do much worse after repeating (-9.6 wpm) than nonrepeaters. Similarly, children who are underage for their grade perform significantly worse (-4.2 wpm) than children who are of age or slightly older.

Absenteeism is another problematic student level characteristic, with a negative relationship with oral reading fluency of -2.4 wpm.

Figure 19. Predictive Factors Relationships with Oral Reading Fluency



6. Regional Analysis Comparing Subtask Achievement

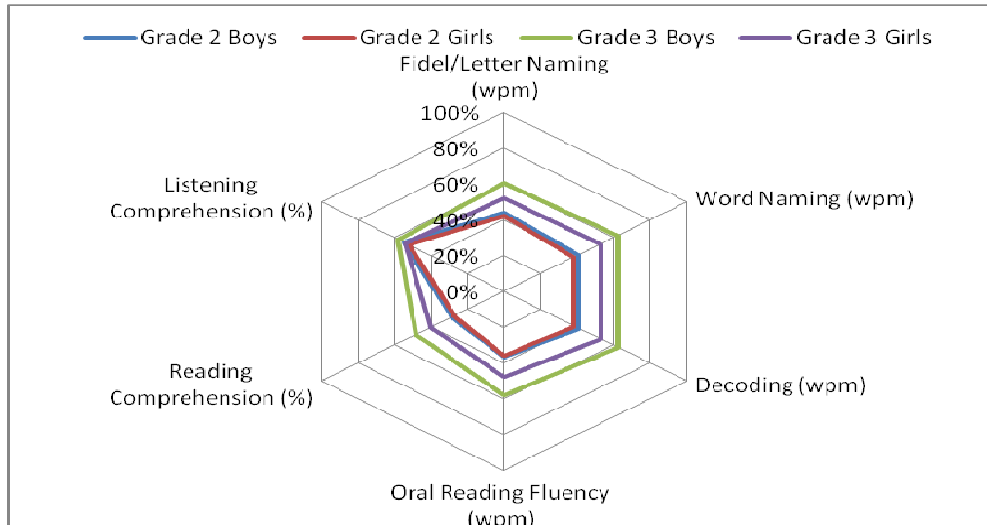
Given the unique nature of Ethiopian languages, and the inability to make comparisons across languages, even within Ethiopia, RTI has taken care to assess the quality of reading outcomes against Ethiopia-specific benchmarks, using the benchmarks that RTI has identified in the many countries where EGRA has been undertaken. Two regional figures are used to investigate whether there are gender gaps in achievement, and to determine the portions of reading that are most at risk, as well as whether there are significant differences between genders and grades.¹⁹

Figure 20 (for Amhara region) shows several things. First, comparisons can be made between the scores for Grade 2 (blue and red bars) and the scores for Grade 3 (green and purple bars). There is a strong relationship at Grade 3 which shows that for all tasks save listening comprehension, boys outperform girls. The Grade 3 scores are also significantly higher than those of Grade 2, which means that children are continuing to learn basic literacy skills in Grade 3. This is even for very basic skills such as fidel naming fluency. This means that some children are still learning the fidel at Grade 3, when the assumption is that they are already fluent with the tasks much earlier.

For Figure 20, it is important to understand what 100% indicates. This is the 90th percentile score for Ethiopian children in Amhara for each of these tasks. While not quite to the levels that we have found for baselines in other countries, the 90th percentile score as a benchmark shows that children in Amhara are capable of high achievement across these measures, and compares Amhara children to others in the same region taking the same language assessment. Note how far the average child is from the Amhara benchmark, however. Most scores in Grade 2 are close to the 40th percentile, and those for Grade 3 are between the 40th and 60th percentiles. This shows that for all tasks (letters naming, word naming, decoding, oral reading fluency, reading comprehension, and listening comprehension, the average child (boy or girl) in both Grade 2 and Grade 3 remains quite far from the levels necessary for grade level literacy scores, as defined by the minimum learning competencies, particularly in reading comprehension.

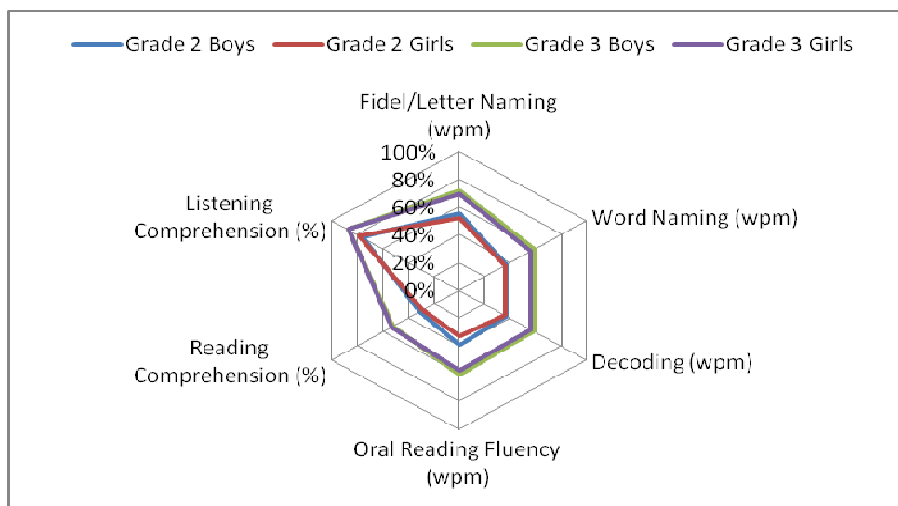
¹⁹ These regional radial plots are each presented in full detail in the regional sections below.

Figure 20. Amhara Radial Plot Comparing Achievement by Grade and Gender for Six EGRA Tasks



The Oromiya radial Figure 21 shows relatively similar achievement scores across levels. However, the gender gaps are less pronounced in Oromiya as they are in Amhara, meaning less gender inequity. Interestingly, the Oromiya children (both Grade 2 and Grade 3) do quite well on listening comprehension and well on letter naming. Their weak areas are in word naming, oral reading fluency, and reading comprehension. These are all tasks that require decoding skills. This is clear evidence for Oromiya, that while children know the language (as identified by listening comprehension) and know their alphabet (as identified by letter naming), they struggle with reading basic words, combining those words into a story, and reading comprehension.

Figure 21. Oromiya Radial Plot Comparing Achievement by Grade and Gender for Six EGRA Tasks



These regional analyses are presented in more detail in the regional annex.

7. Proposed Benchmarks

In this section, we take advantage of the rich amount of Ethiopia-specific data collected by the EGRA study to investigate what information is available to support the development of basic benchmarks for oral reading fluency for each language. To do so, we present several sets of data in this section. First, quantile regression methods are used to show potential markers for oral reading fluency scores. Second, analysis of the average reading scores for schools in the lowest 25th percentile of wealth variables is used to show that schools in poor areas can do quite well in oral reading fluency. Third, scatter plots matching oral reading fluency and reading comprehension scores are presented to investigate the fluency levels necessary to ensure high levels of reading comprehension. Fourth, multiple regression results are used to determine the levels of fluency for the expected levels of reading comprehension.

7.1 Quantile Regression Results

In Table 11, quantile regression methods are used to estimate the 50th, 75th, 80th, and 90th percentile scores for each language. For some of the low-scoring languages—Sidamigna in particular—the median score in the sample was zero wpm read. This means that even the 90th percentile scores were quite low. For the languages with the largest sample sizes (Amharic and Afan Oromo), the 90th percentile scores were the highest.

Table 11. Quantile Regression Results for Each Level of Oral Reading Fluency by Language

	Median	75 th percentile	80 th percentile	90 th percentile
Amharic	27	45	49	59
Afan Oromo	19	42	49	60
Tigrigna	18	34	37	44
Sidamigna	0	15	18	29
Hararigna	27	38	42	50
Somali	31	50	52	57

7.2 Results from High Achieving Poor Schools

In Table 12, we present the highest scoring schools (on oral reading fluency) that had a predominately poor student population.²⁰ This table shows that for Amharic, Somali, and

²⁰ This was done by averaging the socioeconomic status measures at the student level to a school level average. Each

Afan Oromo, there are schools with very poor student populations that also have relatively good average oral reading fluency scores. This is important because setting benchmarks from wealthy populations only is problematic. Instead, in these regions, there are schools with poor children who are successful at teaching those children to read. On the other hand, in Tigray, the highest scoring schools are all relatively wealthy or middle class. For Sidama, all of the average scores for schools are quite low, and for Hararigna, the sample size is too small to perform this analysis. It is encouraging, however, that relatively high achievement levels exist in some poor schools, even before reading interventions have been undertaken.²¹

Table 12. Oral Reading Fluency Scores at the School Level for High Scoring Schools with High Levels of Student Poverty

Language	Socio-economic Status	Region	Woreda	School	ORF
Amharic	Less 25 th	Benishangul-Gumuz	Dibati	Manden	42.6
	Less 25 th	Amhara	Dembecha	Yechereqa Tsion	40.9
	Less 25 th	Amhara	Sayint	Ewa	39.4
Somali	Less 25 th	Somali	Jijiga	Ceel Baxay	48.2
	Less 25 th	Somali	Jijiga	Ceel Amxaar	46.3
	Less 25 th	Somali	Jijiga	Gelbob	43.9
	Less 25 th	Somali	Jijiga	Harre	43.5
Afan Oromo	Less 25 th	Oromiya	Seka Chokorsa	Dabo Yaya	41.7
Tigrigna	Less 25 th	Tigray	Enda Mokoni	Shimta	23.5
Sidaamu Afoo	Less 25 th	Sidama	Chere	Shiko Genet	17.4

7.3 Oral Reading Fluency and Reading Comprehension Scores

In the scatter plots presented in Figure 22, the oral reading fluency scores for each language are presented on the X-axis with the associated reading comprehension scores presented on the Y-axis. The idea is to investigate what levels of oral reading fluency are associated with particular levels of reading comprehension. This is under the assumption

school cited here had less than the 25th percentile score for socioeconomic status.

²¹ Scores in most poor schools are quite low, yet these schools with poor populations show that it is possible for children to do quite well.

that reading comprehension is the ultimate and final goal of reading. The two largest languages in the sample, Amharic and Afan Oromo, are presented below. It is notable that the relationships look remarkably similar, and that a strong linear relationship between reading fluency and comprehension is evident. If and when a child reads at an oral reading fluency rate of 50–70 wpm in both Amharic and Afan Oromo, that child is likely to be able to comprehend at a rate of 80% or 100%.

Figure 22. Oral Reading Fluency and Comprehension Rates for Amharic and Afan Oromo

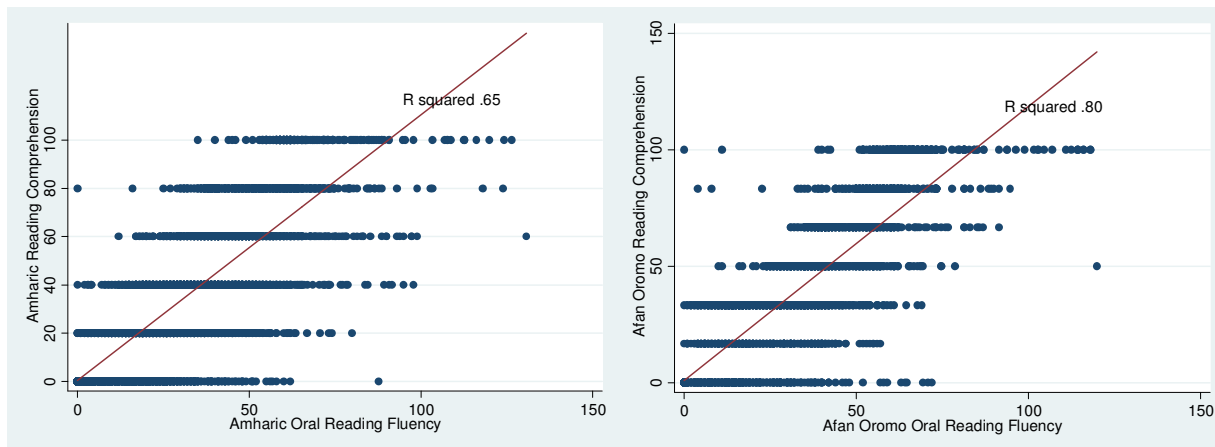
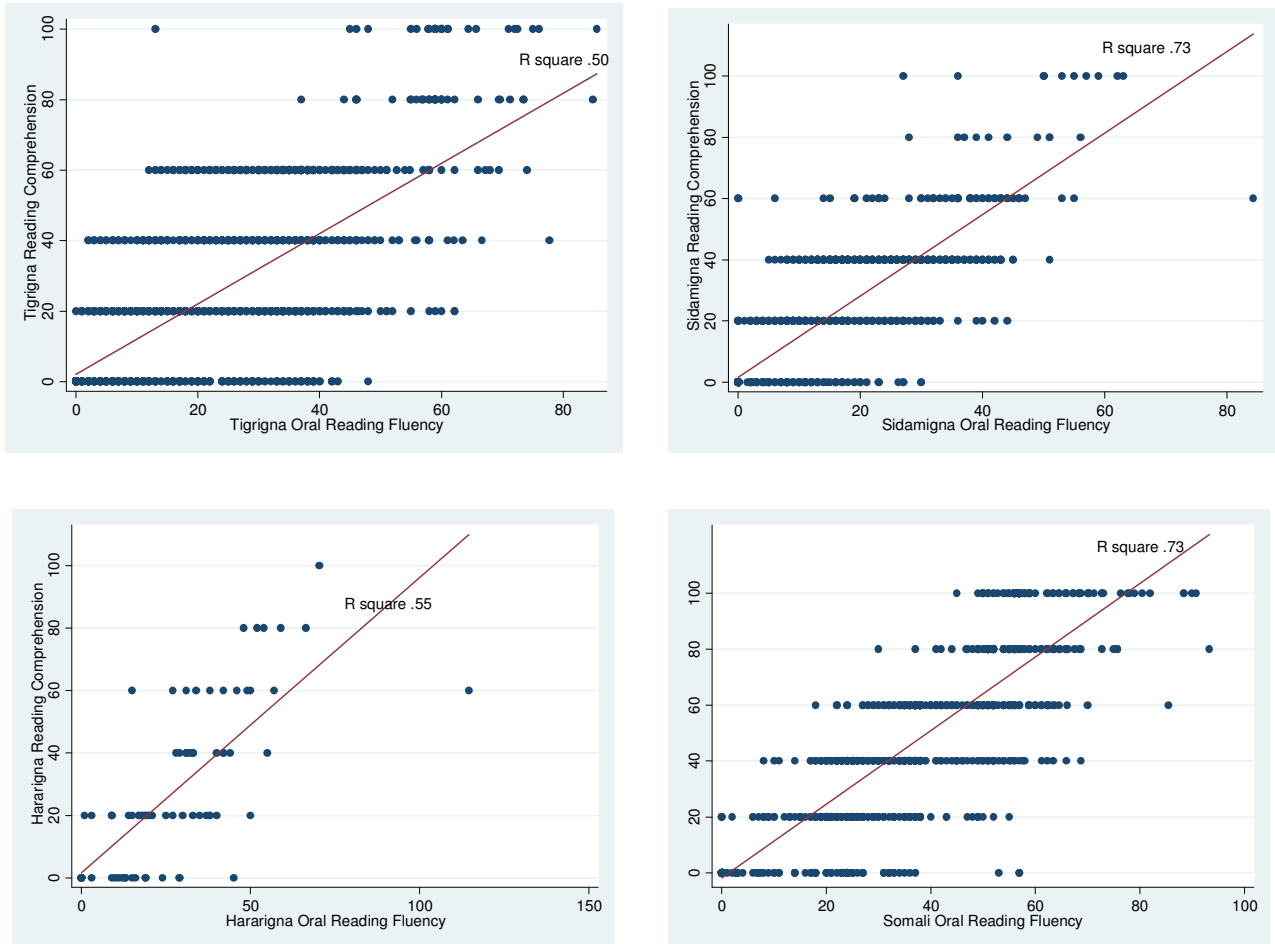


Figure 23 presents the same relationships for each of the other four languages in the study (Tigrigna, Sidamigna, Hararigna, and Somali). Though the sample sizes vary, the general point remains quite similar to what was found above for Afan Oromo and Amharic. The more fluent the child reads, across each language, the higher his or her comprehension scores. Moreover, in order to reach 80% of 100% comprehension scores, oral reading fluency levels need to be somewhere between 50 and 80 wpm. Similar findings exist regardless of which definition of reading comprehension is used.²²

²² Reading comprehension is assessed in two ways: (1) the number of comprehension questions correct out of the total (5 or 6); and (2) the number of comprehension questions correct out of the number attempted (1 to 6, depending on the child). Since children who read with low fluency are unlikely to understand what they read, the computations do not differ very much.

Figure 23. Oral Reading Fluency and Reading Comprehension Scatter Plots for Tigrigna, Sidamigna, Hararigna, and Somali



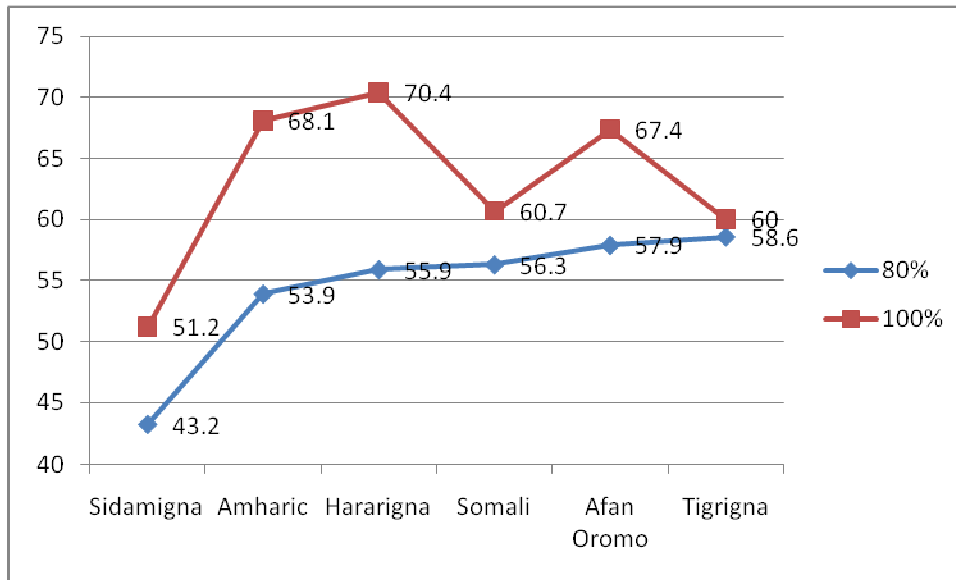
7.4 Levels of Fluency Needed for High Comprehension

Figure 24 specifies the levels of oral reading fluency necessary for high levels of comprehension. In order to examine more closely what levels of oral reading fluency are necessary to reach particular levels of comprehension, we performed simple multiple regression analyses for each of the 6 languages. These regressions allowed us to determine the average predicted level of oral reading fluency associated with each level of reading comprehension. Compared to analyses of this type from other countries, the Ethiopia results were quite similar. The range of oral reading fluency needed for 80% to 100% comprehension was between 53 wpm to 59 wpm for 80% comprehension.²³ For

²³ Note that the oral reading fluency scores for Sidaamu Afoo are lower than what is mentioned here. That is because the extremely low average scores for this language make regression analyses difficult to perform with accuracy.

100% comprehension, on the other hand, oral reading fluency scores are somewhere between 60.0 wpm and 70.4 wpm.

Figure 24. Oral Reading Fluency Scores Necessary for 80% and 100% Comprehension



7.5 Draft Fluency and Comprehension Benchmarks

At the EGRA findings workshop held in Adama on 3-4 November, 2010, regional language groups met to investigate the current levels of reading fluency and comprehension outcomes in each region and to develop draft benchmarks for further use as targets for policy reform. The Table 12a below presents the fluency benchmarks for children to read with at least 80% comprehension by the end of Grade 2. Each group was provided the information presented in this section, and then when requested, were provided with the current percentage of children reading at the benchmark proposed. Note that all of the groups reading comprehension benchmark was at least 80% comprehension rates. The differences depended on the oral reading fluency targets deemed critical to meet. The two different groups for Amharic (expressed as Amhara region and Amharic Group 2) were asked to independently determine a benchmark for the Amharic language, and then come together to agree on one figure. The groups had fundamentally different perspectives on the issue, with the first group (focused on Amhara region) preferring to set a modest fluency benchmark reachable by a large percentage of children, while the second group thought that it was more important for the

Note also that the oral reading fluency scores for 80% comprehension when the alternative definition of reading comprehension (percent correct of attempted, rather than of the total) is used are very similar. The reading fluency scores for 100% comprehension are different, because some children could barely read, but were able to answer one or two questions correctly. This does not change the results of this analysis with respect to the levels of reading fluency necessary for high scores on reading comprehension.

benchmark to be set high. The Table presents an important first step at determining Ethiopian language specific targets for oral reading fluency outcomes.

Table 12a. Draft Oral Reading Fluency Benchmarks by Ethiopian Language

	Proposed Benchmark (words per minute)	Percentage of children at benchmark (2010)	Target percentage of children at benchmark by 2015
Amhara region	60 wpm	4.3%	80%
Amharic (Group 2)	90 wpm	1.0%	50%
Afan Oromo	70 wpm	4.2%	60%
Tigrigna	60 wpm	0.4%	45%
Sidaamu Afoo	75 wpm	0.1%	70%
Hararigna	60 wpm	0.6%	50%

8. Interventions in Early Literacy in Sub-Saharan Africa

8.1 International Experience

This section presents the findings from analyses of interventions in literacy in the last two years in Sub-Saharan Africa in Kenya, South Africa, and Liberia in order to investigate how Ethiopia could respond to the findings of the EGRA presented here. The EGRA findings in Kenya are illuminating. While Kenya has made notable progress in ensuring access for all children, achieving a Net Enrollment Rate (NER) of 92.5% with near gender equity, the quality of education experienced by the youngest pupils remains disparate. For example, the 2007 analysis of Early Grade Reading in Malindi district in Coast province found that pupils in Standard 2 could read 11 wpm on average, far below the international benchmark of 60 wpm, and more importantly, much lower than policy experts at the Kenyan MOE believed was acceptable for Standard 2 (at least 45 correct wpm [EGRA Kenya Stakeholder Workshop, April 2007]). As a result of that low achievement, the Kenyan MOE supported a nationwide baseline assessment of learning achievement in Kiswahili, English, Kikuyu, and Dholuo. In addition to the expansion of the assessment system in Kenya, including EGRA, the MOE and District Education Officers were heavily involved in the implementation of a reading intervention in Malindi managed by Aga Khan Foundation and RTI. This intervention was found to increase the number of words read by children by nearly 100%, and had an impact on pupils’

knowledge of letter names, letter sounds, and word identification. These large impacts occurred in both Kiswahili and English.

8.2 Kenya Intervention Findings

Research provides a great deal of knowledge about the reading skills of Kenyan children. Table 13 presents the results of the Kenyan baseline in 2007, when Grade 2²⁴ children were assessed in Kiswahili and English. In Kiswahili, children were able to properly identify only 4.7 letters in a minute²⁵, and in English, they identified 22.7 letters in one minute. From a list of 50 commonly used words, children were able to identify only 11.7 Kiswahili words and 7.5 English words, showing a limited reading vocabulary and limited fluency even within the vocabulary.

Most concerning, however, are the scores from the oral reading fluency tasks, where children were asked to read a short story of around 60 words in one minute. In Kiswahili, children were only able to read 1/6 of the story correctly (10.2 words), and the English score of 11.4 words was only a little bit higher. As far as comprehension is concerned, it is unsurprising that children who were unable to read most of the passage also exhibited low comprehension, with the Kiswahili scores and English scores showing that children answered less than one half of a question correctly out of a total of 5 questions. One explanation for the low achievement scores is the importance of reading fluency on comprehension, such that children’s ability to decode words is necessary but not sufficient for their ability to understand what they read. In this context, then, children were able to read only a few words and were unable to turn their ability to decode into comprehension.

In summary, Table 13 shows that Grade 2 children in Malindi had low reading skills.

Table 13. Kenyan Reading Scores at Baseline

Task	Kiswahili		Task	English	
	Mean	Standard Deviation		Mean	Standard Deviation
Letter Naming Fluency	4.7	10.7	Letter Naming Fluency	22.7	19.9
Word Naming Fluency	11.7	13.7	Word Naming Fluency	7.5	11.5
Oral Reading Fluency	10.2	14.0	Oral Reading Fluency	11.4	16.2
Reading Comprehension	0.45	1.1	Reading Comprehension	0.39	0.95

²⁴ Kenyan grades are called standards or classes and roughly correlate with grades in Ethiopia.

²⁵ It appears that part of the reason for the low scores on letters in Kiswahili is that children confused the English and Kiswahili letter names.

The results from interventions in South Africa and Liberia show similarly dismal baseline findings. In Liberia, children in Grade 2 were able to read 14.5 wpm in a short story at the baseline. In South Africa, at the end of Grade 1, pupils in control schools could read only 3.9 words per minute in a short story. Preliminary findings from the current EGRA in Kenya, assessed in Grade 3, do not deviate from the basic message: Pupils leave the early grades with very low levels of reading skills.

While much of the research evidence presented above presents a bleak picture of reading achievement in Kenya, the intervention in Malindi and, even more so, findings from recent analyses of experimental interventions in South Africa and Liberia show that targeting reading skills early in primary school can have large impacts on student achievement. In Malindi, after only seven months of intervention, children in treatment schools increased their scores in Kiswahili by 335% (letter recognition), 96% (word recognition), 100% (oral reading fluency), and 106% (reading comprehension). Gains were slightly less for English. The project was well-received by the community and the schools, so much so that since experimental, control, and treatment schools were fairly close one to another; control schools increased their scores almost as much. Qualitative analysis shows that these increases were due to exchanges of ideas between treatment and control schools in the experiment, and show that Kenyan teachers are particularly interested in learning new methods for teaching reading.

8.3 South Africa Intervention

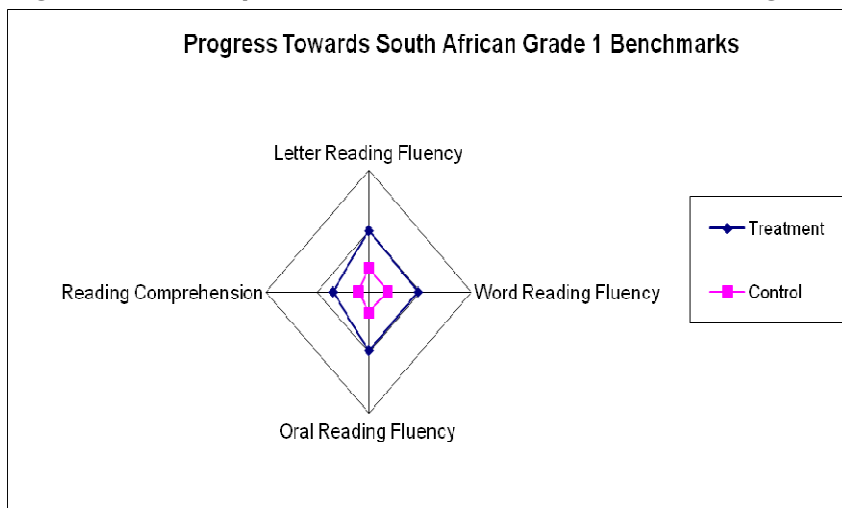
The findings from South Africa show remarkably large impacts on student achievement in all of the major categories of reading skills. In Table 14, the left column shows the three different strategies used to analyze the data, with remarkable consistency that the program increased the number of letter sounds identified correctly by 14.1 per minute. This is an increase over the baseline group. Similarly, being in the treatment group increased word naming fluency by 4.6 wpm. Notably, the program increased oral reading fluency by 7.3 words, and reading comprehension by more than 8 percentage points. Compared against the baseline and the standard deviations, the magnitude of these impacts are, in some cases, more than 2 standard deviations, remarkably large for social science research in general, and huge for educational interventions. The program works, and quite well.

Table 14. South Africa Impacts from Literacy Intervention

	South Africa EGRA Estimates (Units)				
	Letter Sounding Fluency	Word Fluency	Naming Fluency	Oral Reading Fluency	Reading Comprehension
Program Impacts	14.13 (letters pm)	4.61 (words pm)		7.22 (words pm)	8.24 (percent correct)

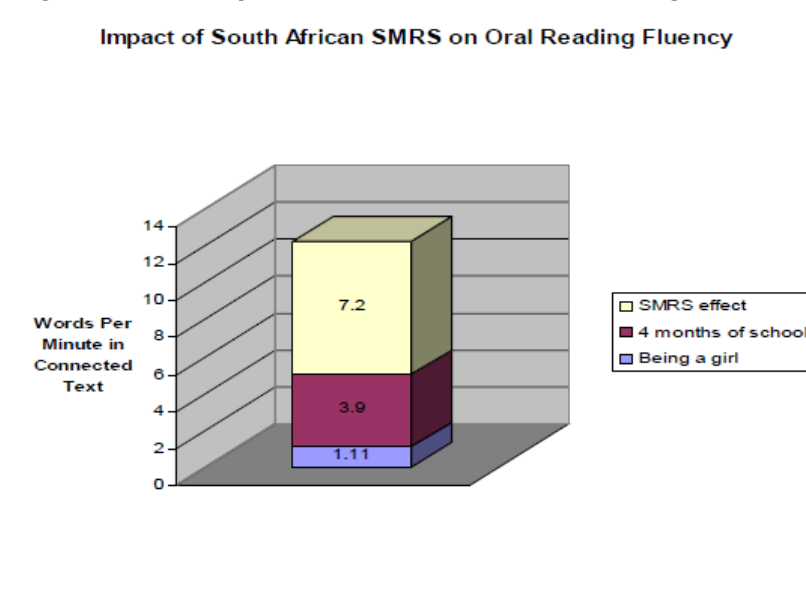
Figure 25 below graphically presents the impact of the South African intervention on Grade 1 outcomes. The radial plots show how much more close the treatment schools are to high level benchmarks after just a few months of interventions. The bar graph on the right shows the very large impact of the program on oral reading fluency, with the effect of the program (called Systematic Method for Reading Success (SMRS)) nearly twice as large as the effect of being in school. In other words, participating in the program basically doubled how much children could read. The impacts were similarly large for letter reading fluency, word reading fluency, and reading comprehension. The program seemed to be particularly successful because of its focus on the development of lesson plans targeted at particular skills and using particular materials.

Figure 25. Impact of South Africa SMRS on Reading Outcomes, by Task



The program in South Africa was also effective because of its targeted population of Grade 1. Children in the program basically learned twice as much as children who did not attend school in the area of oral reading fluency, as Figure 26 indicates. This shows that with careful design, in mother tongue, and in lower levels, children’s outcomes can be drastically improved quite quickly.

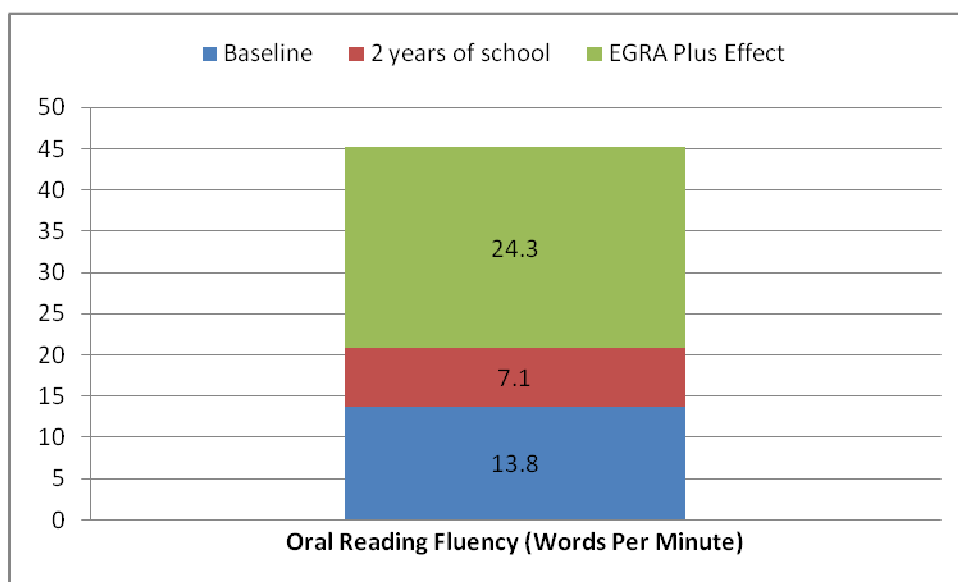
Figure 26. Impact of South Africa SMRS Program on Oral Reading Fluency Outcomes



8.4 Liberia Intervention (EGRA Plus)

In Liberia, a program was designed that used an adapted form of EGRA to support teachers in monitoring education quality in its schools. This program used EGRA also as a means of informing the community about the quality of education by using report cards. While very short in duration, at least at the mid-term assessment, the program had significant impacts after 4 months of intervention. The Liberian results show that a program targeted at Grade 2 and Grade 3 children increased letter naming fluency by 21.0 letters per minute, phonemic awareness cores by 17.7%, familiar word fluency by 15.7 words per minute, unfamiliar word fluency by 12.8 words per minute, oral reading fluency by 24.3 words per minute on connected text, and reading comprehension by 30.2%. Note that for most tasks, being in treatment was worth two or even three years of school. For small experimental programs, these increases are remarkable, as Figure 27 shows.

Figure 27. Impact of Liberia EGRA: Plus on Oral Reading Fluency Outcomes



8.5 Interventions in Ethiopia

Ethiopia has an opportunity to take the programs from these experiments to scale. However, care must be taken, since early grade reading interventions have not yet been brought to national scale anywhere. With significant strategic investments, and commitment to quality, an early grade reading program could have significant impacts on educational quality in primary school.

9. Recommendations

Ethiopia's commitment to improving the quality of education is quite clear, with the MOE engaged in a long-term GEQIP to identify and focus on quality improvement areas across the educational system. The study findings presented in this report represent the largest administration of the EGRA methodology to date. The findings supplement the work of the Ethiopian Training Quality Assurance Agency in the National Learning Assessment studies of 2000, 2004, and 2007. These reports suggested that the quality of reading outcomes is declining slightly; yet the NLA studies were not designed to determine what prereading and early reading skills might be necessary to improve in order to support reading instruction outcomes. The EGRA study was assessed against the Minimum Learning Competencies in mother tongue and EGRA is in line with the Ethiopian curriculum. The findings show that students are not learning at the level expected at either Grade 2 or Grade 3. This results in very low oral reading fluency levels and, for the most part, extremely low reading comprehension. This low achievement in reading comprehension is very concerning, given that listening comprehension skills were moderately strong across regions. The gap, then, is not in language skills or

vocabulary acquisition, but in children’s ability to read, decode, and comprehend written text.

With respect to language of instruction, Ethiopia should be lauded for its success on encouraging adherence to what is the most progressive language of instruction policy in Sub-Saharan Africa. More than 70% of the children in these 6 regions report they learn in the language they speak at home, and many of the larger regions have percentages closer to 90%. The student questionnaires in this study reveal that most children have access to the school reading textbook, but small percentages of children have access to reading materials of any kind at home. The study also shows that these materials are highly correlated with higher achievement in reading. Critically important seems to be whether family members (mothers, fathers or siblings) are available to support children in their homework, and whether that support has created a focus on reading instruction in the home. If so, children’s reading outcome scores are significantly higher.

Although overall scores in oral reading fluency and reading comprehension are low, an interesting result was identified. The correlations between oral reading fluency and reading comprehension were very high, across all languages and regions. This suggests that while most children do not comprehend what they read, the likely reason is that they are unable to read with sufficient fluency to encourage comprehension. The implication is that if Ethiopian classrooms are able to increase children’s ability to decode words, the impact on comprehension will be significant and positive. It is encouraging to know that if children have the skills to read text fluently they also have the skills to comprehend what is read. The following are policy recommendations in response to these findings:

Focus resources and policy efforts on reading instruction. Two-thirds of teachers report not having any in-service teacher professional development in reading methods and pedagogical techniques, and 61% report not having any in-service training at all. Critically, Ethiopia’s regions do have dedicated class time for mother tongue reading. This should be supported by specific training for teachers on how to appropriately and successfully teach children this content. This instruction will support their ability to help children learn the fundamentals necessary for successful reading, including vocabulary, phonemic awareness, fluency, and decoding. Teaching teachers how to foster these skills is of paramount importance for the improvement of reading outcomes in Ethiopia. Teaching teachers how to teach reading must be language-specific. Languages with Sabeian scripts present challenges with respect to learning the fidel fluently, while languages with Latin alphabets present challenges of learning the alphabet, then issues of combining letters to make words. These languages are also very different, structurally, so the guidance should be technically precise in how to support this instruction. Moreover, given that Amharic and English are often taught as second (and third languages), care must be taken to deal with potential language interference issues. In short, in order to enjoy the quick and deep changes necessary to reading instruction outcomes, the Ministry of Education and Regional Education Bureaus should focus their considerable expertise on the improvement of reading outcomes in the country.

Start early, in Grade 1 and 2. The findings show that teachers' views of when students can read and understand what they read are important for student outcomes. This suggests that in some Ethiopian classrooms teachers wait too long to teach students how to read and expect too little from their young learners, and this has implications for what children can gain from their time in early primary school. It is recommended, therefore, that Grade 1 to Grade 2 pedagogy focus most heavily on early reading acquisition and outcomes. This is in line with the expectations from the Minimum Learning Competencies document.

Encourage reading in the community. The findings showed that few classrooms were stocked with reading books other than the textbook, and very few children had any reading materials at their homes, let alone reading materials appropriate for their reading skills. This means that children have limited exposure to the joys of reading engaging and interesting materials appropriate for their developmental stages. A two-pronged effort should be made to *increase the amount of reading material* in classrooms and encourage families to *make reading a part of their daily family activity*. A primary strategy for this change must come from taking advantage of the richness of the intertwined Ethiopian society to encourage families to expect their children to read after being in school, and to provide the children with adequate literate environments. Strategies by which this could occur include woreda-level reading contests, book drives at the local level, and awareness raising activities at the regional education bureau and woreda levels.

Review teacher in-service and pre-service professional development. The findings from this study clearly indicate that little reading instruction is happening in many schools, although literacy class is a significant part of the day. By this we mean that there is far less interaction between teachers and students around letters, words, sentences, and stories than there should be. This need not be the case, and experiments in Kenya, Liberia, and South Africa show that teachers can be very receptive to focused in-service professional development supporting skill acquisition in early literacy interventions. It is recommended that the in-service teacher professional development programs target the building blocks of reading and, where possible, provide targeted lesson plans for teachers, closely related to the textbooks and supplementary reading materials. Ideally, these skills can be incorporated into the College for Teacher Education curriculum to help new teachers prepare students to read. Specifically, we recommend that the in-service program include a specific module focused on the development of basic reading teaching skills, specific to each language. This would include a focus on letter and fidel sounds, the alphabetic principle, decoding new words, and comprehension strategies, amongst other skills.

Review reading curriculum and textbooks. The recent curriculum reform under GEQIP provides an opportunity for Ethiopia's regions to develop reading curriculum that can systematically address the weak levels of reading. This can be done by carefully reviewing reading acquisition in each language, with particular attention to the frequency of letters and words in the language. This will determine in what order particular letter

sounds (or fidel) should be taught, and then support the development of the necessary decoding skills, all in the context of consistent and relevant exposure to interesting and engaging stories. The existing textbooks seem to be written at a level beyond the average learner in the studies, and lesson plans that can support the textbooks, lesson by lesson, would prove a critical component to reading improvement in Ethiopia.

Set literacy benchmarks. The Minimum Learning Competencies document provides guidelines for the expected levels of reading for children. The lack of specific guidelines at the word and sentence level means, however, that it is difficult for teachers and educators to determine whether children read at the appropriate levels. The Fast Track Initiative is creating indicative frameworks with wpm read (with comprehension) as an important marker for education quality. The data presented in this report can be used to create language-specific oral reading fluency benchmarks, such that policy makers determine what outcomes they expect children to achieve by the end of Grade 2. The findings suggest that without benchmarks, children are unlikely to be able to read fluently and with comprehension.

Review teacher deployment. The assignment of teachers by subject should be based on skill level, including language facility, and interest. Teachers trained as generalists, with skills in mathematics, will have less facility with teaching reading than will teachers who have specific subject knowledge and interest in language issues.

Improve the quality of reading instruction. Using some of the recommendations below, including pre-service and particularly in-service teacher professional development, the findings show that reading instruction must be improved. There are several pedagogical areas to focus on.

- **Differentiate reading instruction methods by language.** The findings in this report show that once children have learned the fidel in Sabean script languages, reading words is relatively simple. Many children in regions that use Sabean script languages take far too long to master the fidel, however. On the other hand, decoding (combining letter sounds into words) remains a difficult challenge for Ethiopian languages using the Latin alphabet. The findings show that there is little teaching of decoding in Latin alphabet languages. It is critical that the methods for teaching reading be differentiated by language, and teachers be provided a full and ongoing training of how young children should grasp the basics of that language.
- **Teach decoding.** The findings from the unfamiliar word reading (decoding) tasks show that children in all languages have limited skills in reading new and unfamiliar words. This is logical given that much of Ethiopian reading instruction focuses on reading words that children already know. However, while Ethiopian children have the ability to memorize a small number of words, reading new words is a critical skill that most teachers are not providing. Increasing children’s ability to combine letters (or fidel) into new

words is critically important to increase their reading fluency and comprehension. In Latin alphabet languages, we found that teachers pointed to words (e.g., “house”) and encouraged the children to recite the word aloud. And these children became proficient in reading house. However, when faced with the word “mouse,” those same children would do very poorly, since the pedagogy encourages the children to memorize particular words and spends much less time training them how to decode and “solve” new words to which they have not yet been exposed.

- **Teach formal comprehension strategies.** The children in these 8 regions had very low comprehension levels. This is not likely due to low levels of oral vocabulary, given the relatively high scores in listening comprehension. Instead, it appears that children have weak skills in the metacognitive processes that ensure comprehension. These include the skills of reviewing, questioning, and predicting. This appears to be because children have not had much formal training in comprehension strategies. These strategies can be systematically taught. Note, however, that without the ability to read fluently, comprehension is nearly impossible. Therefore, increased ability to read fluently is critical to the ability to comprehend.

Expand literacy interventions under GEQIP. Given the existing quality improvement initiative, the findings suggest that the MOE and GEQIP respond to the low findings in this report and expand literacy interventions in particular languages. Clearly, there needs to be improvements in teachers’ techniques in teaching literacy. Combined with scripted lesson plans, material book development and provision, and ongoing teacher professional development, it is clear that improvements to the quality of reading outcomes can be had in Ethiopia. Note that literacy interventions with these elements have been able to increase oral reading fluency scores by more than 100% in South Africa, Liberia, and Kenya. We suggest that the following elements be included in the literacy interventions that are implemented:

- *Development of targeted lesson plans.* Teachers should be provided with specific instructions on how to teach early reading acquisition, since most pre-service programs do not provide the level of detail and precision necessary to do it properly.
- *Provision of ongoing professional development.* To support the behavioral changes necessary to help teachers teach better, teachers need ongoing support using a combination of new instructional methods and opportunities to discuss how their experiments with the new methods are going with colleagues. A coaching model would support this effectively, particularly if the cluster and supervisor system is employed.
- *Development and use of significant reading materials.* Leveled materials developed to support the graded instruction in Grade 1 and 2 that can be read easily by burgeoning learners and incorporated into lessons are necessary.

- *Support from the community.* The community and their support are critical to the improvement of learning outcomes. Whether it is asking parents to ensure that their children read 20 minutes a day, or engaging children in reading competitions and holding teachers accountable, the community is critical to any drastic improvement of literacy.