NATIONAL ASSESSMENT SURVEY OF LEARNING ACHIEVEMENT AT GRADE 2

Results for Early Grade Reading in Zambia

April 2015

This study report was made possible by the support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of RTI International and do not necessarily reflect the views of USAID or the United States Government.

USAID
FROM THE AMERICAN PEOPLE

NATIONAL ASSESSMENT SURVEY OF LEARNING ACHIEVEMENT AT GRADE 2

Results for Early Grade Reading in Zambia

Prepared for

Teresiah Wambui Gathenya (PhD), Education Advisor
Education Office, United States Agency for International Development (USAID)/Zambia
U.S. Embassy
Subdivision 694 / Stand 100
P.O. Box 32481
Kabulonga District, Ibex Hill Road
Lusaka, District
Zambia

Prepared by

Aarnout Brombacher, Jennae Bulat, Simon King, Emily Kochetkova, and Lee Nordstrum
RTI International
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

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

ACKNOWLEDGMENTS

The authors wish to acknowledge partnership support from the United States Agency for International Development (USAID); the Government of the Republic of Zambia, through its Ministry of Education, Science, Vocational Training and Early Education (MESVTEE); and the Examinations Council of Zambia. Many departments and staff from the Ministry of Education and the Examinations Council of Zambia also made important contributions to the development and implementation of this study and helped ensure that the instruments and methodologies were appropriately aligned to the Zambia context. Most importantly, this work could not have succeeded without the cooperation and contributions of the District Education Officers, head teachers, teachers, pupils, and communities included in the study who, for obvious reasons, must remain anonymous.

The USAID funding for the Early Grade Reading Assessment (EGRA) and the associated contextual interviews and observations was provided through an Education Data for Decision Making (EdData II) task order, Data Collection Services for the USAID/Zambia Education Project, Contract No. AID-611-M-14-00002 (RTI Task 28). The British Department for International Development (DFID) sponsored the simultaneous administration of an Early Grade Mathematics Assessment (EGMA), whose results are documented in a separate report.

CONTENTS

LIST OF FIGURES

PAGE

LIST OF TABLES

PAGE

ABBREVIATIONS

DFID	British Department for International Development
ECZ	Examinations Council of Zambia
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
G2, G4	grade 2, grade 4
GRZ	Government of the Republic of Zambia
INSET	in-service education and training
IRC	International Rescue Committee
LOI	language of instruction
MESVTEE	Ministry of Education, Science, Vocational Training and Early Education
NAS	National Assessment Survey
ORF	Oral Reading Fluency
PPS	probability proportional to size
PTR	pupil–teacher ratio
R-SNDP	Revised Sixth National Development Plan (2013–2016)
RTI	RTI International (registered trademark and trade name of Research Triangle Institute)
SES	socioeconomic status
SNDP	Sixth National Development Plan (2011–2015)
SSME	Snapshot of School Management Effectiveness
USAID	United States Agency for International Development
WPM	words per minute

EXECUTIVE SUMMARY

Zambia is a diverse nation that comprises approximately 72 ethnic groups. Although the official language is English, there are also seven official languages of instruction—Nyanja, Bemba, Lozi, Tonga, Kaonde, Luvale, and Lunda—which the Ministry of Education, Science, Vocational Training, and Early Education (MESVTEE) has directed be used for teaching initial literacy, from preschool through grade 4. English language is introduced as a subject at grade 2.

However, this arrangement creates challenges for the instruction and learning of reading in the early grades in Zambia. Furthermore, the migration of many Zambians from rural to urban areas means that many pupils are learning to read in a language that is not spoken in their home; and in a number of cases, there are entire districts where all the pupils learn in a language that is not the local language. For example, the Tumbuka people living in Eastern Province are officially learning to speak in Nyanja in the classroom even though their language spoken at home is Tumbuka.

This research report is the first publication of the MESVTEE's Grade 2 National Assessment Survey (NAS) report; the Examinations Council of Zambia (ECZ) will collect data and prepare reports for similar studies in the future. In addition to results from Early Grade Reading Assessment (EGRA) in local languages, this report documents findings regarding selected English competencies, as well as responses from questionnaires on pupil and teacher characteristics.

This study was conducted in November 2014. Its purpose was to measure a basic skill that all pupils need to possess as the foundation of a successful education: being able to read fluently with comprehension. The results can be applied to curriculum development and teacher training in order to increase the focus and quality of early grade reading instruction in Zambia. Future Grade 2 National Assessment Survey reports will inform stakeholders of the progress of such programs.

Under consultation with ECZ and the United States Agency for International Development (USAID/Zambia), RTI International administered a survey to 4,855 grade 2 pupils across all 10 provinces of Zambia and all seven official local languages of instruction. Four hundred and eighty-six schools were included in the sample and were a mix of government (GRZ), grant-maintained, community, and private schools.

Overall, the EGRA showed that grade 2 pupils, on average, were struggling to read fluently; the average oral reading fluency rate for the local languages ranged from 1.84 to 8.40 words per minute, indicating that the typical grade 2 pupil could sight-recognize a few words but struggled to string the words from a passage into a coherent sentence. This finding is not surprising, as pupils were able to produce the correct sounds of only between 3.68 and 9.63 letters per minute across languages, indicating they lacked the foundation needed to decode unfamiliar words. This finding also was reflected in the Reading Comprehension subtask, for which most pupils were challenged to answer the comprehension questions of the passage they had just read. However, the issue was not one of inability to comprehend at all, because the pupils could sufficiently comprehend passages being read to them, scoring on average between 43% and 78% correct responses across languages.

While these results will help inform policy for curriculum and instruction for early grade reading, it is important to stress that other contributing factors also make it challenging for children to learn to read. Pupils who said they spoke the same language at home as the one in which they were instructed were able to read, on average, 2.5 words per minute (or 0.23 standard deviations) more than those who did not. This difference was even more evident in locations where Nyanja was the language of instruction, namely Lusaka district, where many pupils speak Bemba at home; and Eastern Province, where pupils in Lundazi district mostly speak Tumbuka at home.

Parents'/guardians' support for their child's reading and education is clearly important, as pupils who read to someone at home every day could read 8.9 words per minute more on average than those who had no one to read to (controlling for all other factors).

Similarly, access to materials in the classroom is critical to reading success; pupils who had a language-of-instruction reader on average read 3.9 words per minute more than those who did not have a reader. Critically, over 80% of pupils reported not having a language-of-instruction reader (a reading text critical for foundational reading development).

For the English assessment, pupils could identify on average 15.9 English letter names per minute. However, they scored on average only 1 out of 5 on the Listening Comprehension subtask, which was much less than pupils' scores on Listening Comprehension for the local languages. Pupils who lived in urban areas, where English is spoken more, had a clear advantage in comprehending the language, such that their Listening Comprehension scores were on average 7 percentage points higher than those of pupils living in rural areas, after controlling for all other measured pupil indicators. Socioeconomic differences were a much bigger factor for English than they were for local language reading; for example, pupils who attended private school scored 17% higher in English Listening Comprehension than pupils in community schools, and pupils identifying with the

highest socioeconomic quintile scored on average 19% more on Listening Comprehension than those from the lowest quintile. These values were irrespective of other factors, such as residence in an urban or rural area.

1 BACKGROUND AND PURPOSE OF THE STUDY

1.1 BACKGROUND

Assessments of pupil learning in the primary grades, such as the Early Grade Reading Assessment (EGRA), offer an opportunity to determine whether children are developing foundational skills upon which other literacy skills build, and, if they are not, where instructional efforts might be best directed. This is vital information for countries that are working to improve the quality of education in their schools. Indeed, the fact that the EGRA has been adapted and used around the world is evidence of growing international concern for children's learning outcomes, as opposed to attendance or completion rates.

Similar to the pattern in many countries, since Zambia abolished school fees for grades 1 through 7 in 2002, enrollment has increased and the policy focus has shifted from access to school to the quality of teaching and learning that takes place in schools. This focus is articulated in the nation's Sixth National Development Plan (SNDP), which spanned 2011 to 2015; as well as in the Revised Sixth National Development Plan (R-SNDP), from 2013 to 2016. The R-SNDP especially acknowledges the challenges facing educators and pupils in rural areas, and notes, "Regional comparisons suggest that the quality of primary education in Zambia is one of the worst in Southern Africa."[1] Reasons given for poor learning outcomes include high pupil–teacher ratios and insufficient time on task due to double- or even triple-shift schools. The R-SNDP concludes that Zambian pupils "exiting primary education have not acquired the basic literacy and numeracy skills they need to be productive in the labour market,"[2] and reinforces the importance of early childhood education and teacher training, among other policy endeavors.

After subnational applications of grade 2 reading assessments in 2011 and 2012, in 2014, grade 2 was officially added to Zambia's National Assessment Programme in order to identify the state of learning outcomes in the early grades, when pupils are expected to gain the foundational reading skills upon which all later learning is built. Prior to this, the National Assessment Programme—which

[1] Revised Sixth National Development Plan, Ministry of Finance, Republic of Zambia, p. 97.
[2] *Ibid.,* p. 98.

began in 1998—included assessments every two years in grade 5. Grade 9 was added in 2013.

For grade 2, a nationally representative sample of pupils was selected to take locally developed versions of the EGRA and the Early Grade Mathematics Assessment (EGMA), which together comprised the Grade 2 National Assessment Survey (NAS). Outcomes on the EGMA portion of the survey, sponsored by the British Department for International Development (DFID), are documented in a separate findings report.

This analysis report for the reading portion of the Grade 2 NAS presents the results for pupil performance on the EGRA in seven local languages, each of which also had a small set of subtasks in English. Additional information about schools, teachers, head teachers, and pupils' socioeconomic background that was collected alongside the EGRA provides context for many of the results.

1.2 RESEARCH QUESTIONS

This study was driven by the following four questions:

1. How well are pupils in Zambia learning foundational reading skills, in their local language of instruction, by grade 2?
2. How well are grade 2 pupils able to name the letters of the English alphabet and understand what they hear in English?
3. What effects do various individual, school-related, and socioeconomic factors have on reading in any language?
4. What comparisons, if any, can be made between the grade 2 NAS reading results and results on earlier reading assessments in Zambia?

Note that the first three questions are consistent with a snapshot of early grade reading performance in 2014. The fourth question was a retrospective exercise that was requested as part of the task order effort.

1.3 LEARNING OUTCOMES MEASURED

As noted, pupil learning outcomes and proficiencies in reading were measured by applying the EGRA in English and the local languages of instruction: Nyanja, Bemba, Lozi, Tonga, Kaonde, Luvale, and Lunda.

The EGRA instrument for the local languages contained subtasks specifically designed to collect data on how well grade 2 pupils would perform on Letter Sound Identification, Oral Reading Fluency (ORF), Reading Comprehension, Listening Comprehension, and Invented (Nonword) Reading. The English EGRA instrument yielded data on Letter Name Identification, Listening Comprehension, and English Vocabulary.

Section 2 provides more detail about the instruments used to gather data on pupil learning outcomes for the study; Section 3 explains the survey methodology; and Section 4 presents pupils' average results for local languages and English.

1.4 BASELINE REPORT

Whereas the 2014 EGRA was intended to give USAID and the ministry a sense of overall pupil reading performance in Zambia, USAID and the MESVTEE also wanted to know whether pupil reading performance in Zambia had increased in recent years (more specifically, since the 2011 EdData II Bemba EGRA assessment and since the 2012 Read to Succeed EGRA).

Section 5.3 and two annexes discuss the results and restrictions of this "baseline" comparison exercise alongside other suggested conclusions and recommendations in Section 5.

2 EVALUATION APPROACH AND DESIGN OF THE ASSESSMENT

It is important that the sample of pupils selected for this NAS was designed such that pupil performance in reading could be reported in seven local languages. The sample design used was probability proportional to size (PPS) with oversampling for subpopulations that are a very small proportion of the population—for example, schools where Lunda is the language of instruction. This design assured good pupil score estimates with an acceptable level of precision. In all, 486 schools were randomly selected for the study; ***Table 1*** shows the number of pupils sampled. Because the study designers wanted to report the EGRA results by local language, it was important to design the sample in consideration of this. Particular attention was paid to North Western province, where the Lunda, Luvale, and Kiikaonde languages are all languages of instruction. For example, District Education Officers were approached to identify the language of instruction for the individual schools in Zambezi district where both Lunda and Luvale are spoken.

Table 1. Number of sampled pupils, by language and province

PROVINCE	LANGUAGE							TOTAL PUPILS BY PROVINCE
	CHITONGA	CINYANJA	ICIBEMBA	KIIKAONDE	LUNDA	LUVALE	SILOZI	
Central	180	10	289					**479**
Copperbelt			613					**613**
Eastern		503						**503**
Luapula			403					**403**
Lusaka		496						**496**
Muchinga			451					**451**
North Western				200	131	254		**585**
Northern			425					**425**
Southern	500							**500**
Western							400	**400**
Total pupils by language	**680**	**1,009**	**2,181**	**200**	**131**	**254**	**400**	**4,855**

2.1 OVERVIEW OF EGRA

2.1.1 WHY TEST EARLY GRADE READING?

The document *Guidance Notes for Planning and Implementing EGRA* (RTI International and International Rescue Committee [IRC], 2011, p. vii) offers the following succinct rationale for the testing of pupils' reading skills in the early grades of schooling:

Reading is a fundamental skill for children. It is also a foundational skill, upon which acquisition of other critical skills and knowledge depend. Nevertheless, it is one that the majority of children in low-income countries are not acquiring as improvements in student learning lag significantly behind improvements in access to schooling (World Bank: Independent Evaluation Group, 2006). Assessments in dozens of low-income countries have revealed that the majority of students cannot read even one word of a simple test and even students who have been enrolled in school for as many as six years are unable to read well (Gove & Cvelich, 2011).

Without basic literacy skills, children are unable to learn, and therefore have little chance of succeeding in school and beyond. Research indicates that children who do not learn to read in the early grades (grades 1–3) are likely to fall further and further behind as they grow older. Moreover, students who do not learn to read in the first few grades are more likely fall behind in other subjects, to repeat grades, and eventually to drop out (Stanovich, 1986).

2.1.2 PURPOSE OF EGRA

Again from the *Guidance Notes* (RTI International & IRC, pp. vii–viii):

The EGRA instrument is designed to assess foundational reading skills that a child must have to read fluently with comprehension.

The EGRA instrument is composed of subtasks designed to assess foundational reading skills crucial to children's successful reading and comprehension abilities. These subtasks are based on research regarding a comprehensive approach to reading acquisition across languages, including these five essential components: phonological awareness, alphabetic knowledge/process, vocabulary, fluency, and comprehension (RTI International, 2009). In particular, EGRA measures oral reading fluency, which has been shown to predict later skills in reading and comprehension (Fuchs, Fuchs, Hosp, & Jenkins, 2001). As such, EGRA is conducted orally and one-on-one with pupils, and takes about 15 minutes to administer per child (RTI International, 2009). In its most common application, it is administered to a sample of many children across multiple schools or districts in order to obtain generalizable results and reliable averages.

Increasing literacy is an objective of USAID through basic education globally with efforts to give children basic literacy skills that are a basis for "Improving Lives Through Learning," and national Early Grade Reading Assessments have been conducted in over 40 countries to date with funding from USAID.

2.1.3 WHAT EGRA MEASURES

The EGRA instrument consists of a variety of subtasks designed to assess the foundational reading skills that are crucial to becoming a fluent reader. EGRA is designed to be a method-independent approach to assessment (i.e., the instrument does not reflect a particular method of reading instruction). Instead, EGRA measures the basic skills that a pupil must possess to eventually be able to read fluently and with comprehension—the ultimate goal of reading. EGRA subtasks are based on research regarding a comprehensive approach to reading acquisition across languages. The skills assessed in the EGRA employed in this study were phonics/decoding, fluency, and reading comprehension, as further described in the following paragraphs.

Phonics/decoding builds on the alphabetic principle, beginning with letter–sound correspondences that help children develop automatic recognition of letter–sound patterns in common words. Eventually, phonics is instrumental in the development of instant recognition of most words that are read. This automatic or instant word recognition is manifested by fluent reading of connected text.

Fluency is often defined as the ability to read with speed, accuracy, and understanding. ORF is a common way to assess whether an individual is a fluent reader. Fluency is considered critical for comprehension because rapid, effortless word-identification processes enable readers to focus on the text and its meaning rather than on word identification or decoding words letter by letter (National Institute of Child Health and Human Development, 2000; Perfetti, 1992).

Reading comprehension, considered to be the goal of reading, refers to the ability to actively engage with, and construct meaning from, the texts that are read.

EGRA measures each of the previously mentioned abilities and/or components to assess the foundational reading skills. The skills are tested in individual subtasks and presented in order of increased level of difficulty. Because the first few subtasks are easier, EGRA can therefore measure a range of reading abilities for beginning readers.

2.1.4 THE EGRA INSTRUMENTS FOR ZAMBIA

Table 2 summarizes the subtasks of the EGRA designed for Zambia.

Table 2. Subtasks of the EGRA instruments in Zambia

SUBTASK	SKILL	DESCRIPTION— THE CHILD WAS ASKED TO …
Listening Comprehension (in local language and English)	The ability to comprehend an orally presented story and provide an oral response to question asked.	…orally respond to five questions asked about an orally presented story. *(Untimed subtask)*
Letter Sound Identification (local language)	The ability to produce the sound of a letter that is presented in written form.	… produce the sounds of 100 letters presented in written form. Letters were presented in a grid of 10 rows and 10 columns. *(Timed subtask)*
Invented (Nonword) Reading	The ability to decode unfamiliar words.	… sound out, or decode, unfamiliar words. To ensure that all words would be unfamiliar words, 50 words without meaning but following the spelling and grammatical rules of the language were presented to the child to read. *(Timed subtask)*
Orientation to Print	An understanding of directionality of reading print on a page.	… indicate where one would begin reading printed text on a page and the direction one would read that text. *(Untimed subtask)*
Oral Reading Fluency	The ability to quickly and accurately read connected text on a page.	… quickly and accurately read a passage of narrative text of approximately 60 words in length. *(Timed subtask)*
Reading Comprehension	The ability to orally respond to both literal and inferential questions about the Oral Reading Fluency passage read.	… orally respond to five questions asked about the passage read. *(Untimed subtask)*
English Vocabulary	The ability to identify body parts, objects in the environment, and simple prepositions presented in English.	…point to objects in response to English vocabulary words orally presented. *(Untimed subtask)*
Letter Name Identification (English only)	The ability to produce the name of a letter that is presented in written form.	… produce the names of 100 letters presented in written form. Letters were presented in a grid of 10 rows and 10 columns. *(Timed subtask)*

2.2 OVERVIEW OF COMPLEMENTARY INSTRUMENTS FOR ZAMBIA

2.2.1 PURPOSE

The instruments that complemented the EGRA in Zambia were intended to yield a quick but rigorous and multifaceted picture of school management and pedagogic practice in the country.[3] The instruments were designed to capture “best” indicators of effective schools that, as past research has shown, affect pupil

[3] These complementary instruments were based on a subset of tools from the Snapshot of School Management Effectiveness (SSME) developed by RTI for USAID under the EdData II project.

learning. With the resulting data, school, district, provincial, or national administrators and donors can learn what is currently occurring in their schools and classrooms and assess how to make their schools more effective.

Based on the framework for the analysis of effective schools described in the effective schools literature,[4] the complementary instruments collected information about (1) basic school inputs such as school infrastructure, pedagogical materials, teacher and head teacher characteristics, pupil characteristics, and parental and community involvement; (2) classroom teaching and learning processes, including use of material, instructional content, pupil-teacher interaction, time spent on task, assessment techniques, and administrative oversight; and (3) learning outcomes data, via the application of core portions of the EGRA.

The complementary instruments were administered during a single school day by a multi-person team. Each of the complementary instruments was designed to supply information from a different perspective. The design aimed to balance the need to include a broad mix of variables—to allow potentially impactful characteristics to be identified—with the competing need to create a tool that was as undisruptive to the school day as possible.

The combined complementary instruments were intended to produce a comprehensive picture of a school's learning environment, and when the results from multiple schools in a region are compared, it becomes possible to account for differences in school performance.

2.2.2 SUMMARY AND DESCRIPTION

Table 3 summarizes the complementary instruments designed for Zambia.

Table 3. Complementary instruments in Zambia

INSTRU-MENT LEVEL	INSTRUMENT NAME	WHAT IT MEASURES
School	Head Teacher Questionnaire	School leadership, teacher characteristics and attendance, infrastructure and facilities, school closings
	School Inventory	Infrastructure and facilities, repairs, safety
Classroom	Teacher Questionnaire	Teacher characteristics and practices, pedagogical oversight
Pupil	Pupil Questionnaire	Pupil background, interactions with teacher

[4] This framework for the analysis of school effectiveness is based on research reported by Heneveld & Craig (1996) and by Carasco, Munene, Kasente, & Odada (1996).

2.3 INSTRUMENT ADAPTATION PROCESS FOR ZAMBIA

The standard EGRA, and SSME instruments developed by RTI International and funded by USAID have been used in dozens of countries by numerous organizations. However, this does not mean the instruments are merely translated for the country in which they are to be used. Rather, the base instruments are adapted to the local context and vetted by a body of national experts from the education community in the host country. The grade-appropriate curriculum and textbooks (in this case Zambia grade 2 reading, writing, and mathematics) are analyzed and used to inform changes and adaptations made to the base EGRA instrument. These instruments are truly localized to fit the country in which they are to be used.

A five-day Instrument Adaptation Workshop was held in Lusaka in July 2014 with respected members from the education community across Zambia. Workshop participants included representatives from various branches of the MESVTEE, teacher training institutions, and curriculum development units, among other stakeholder entities. Three experts from RTI led the reading and mathematics portions of the workshop. Two reading and writing experts led the adaptation discussions for the USAID-sponsored EGRA instruments. Together, the workshop participants modified the base EGRA and SSME complementary tools to be used during the Grade 2 NAS.

The instruments were piloted during a study conducted in the first week of October 2014. The pilot study provided an opportunity for the research team to evaluate the instruments to ensure that these were functioning appropriately and that the desired data were being gathered. The pilot sample consisted of 1,494 pupils at approximately 285 schools. During the pilot study, multiple oral reading passages and listening passages were evaluated for the EGRAs in each of the official languages, including the English items. Piloting multiple passages allowed the research team to determine which reading and listening passages were best to use during the full study.

The researchers analyzed the pilot data in part by using a psychometric Rasch analysis to obtain insights into which subtasks were performing reliably and consistently to capture the desired data, and which subtasks and passages could be eliminated without compromising the richness and integrity of the data. In the end, the pilot study provided justification for eliminating some subtasks from the EGRA instrument and for determining which oral reading and listening passages were most appropriate to include in the final instruments. The stakeholders reviewed and approved the final post-pilot instruments. The instruments were rendered into the RTI-developed Tangerine® software. The assessors then used tablets loaded with the Tangerine versions of the instruments to collect the national data.

3 METHODOLOGY

3.1 SAMPLING FRAMEWORK

Table 1 above summarized the study sample, which was developed in collaboration with the Examinations Council of Zambia. It consisted of 486 schools sampled in all 10 national provinces to allow for regional and language-of-instruction representativeness. The sample frame was all schools with grade 2 pupils, including private, GRZ (government), community, and grant-aided school types.

Seventy-three districts were identified and a designated number were randomly selected by stratification of language regions crossed with province, using probability proportional to size systematic sampling. When the language-of-instruction group was a small percentage of the population (for example, in the Lunda and Luvale language regions), oversampling was used in order to obtain a subpopulation sample that would provide estimates with a reasonable level of precision. In all, 48 districts were selected. These districts were then stratified by school type and—where possible—one private, one grant-aided, and eight community or GRZ schools were selected using PPS systematic sampling.

The final stage of sampling occurred during the school visits: One class was selected at random and 10 pupils were systematically selected from that class.

3.2 DESCRIPTIVE STATISTICS

The following five tables ***(Table 4–Table 8)*** further describe the sample population.

Table 4. Number of sampled schools, by language and province

PROVINCE	LANGUAGE							TOTAL SCHOOLS BY PROVINCE
	CHITONGA	CINYANJA	ICIBEMBA	KIIKAONDE	LUNDA	LUVALE	SILOZI	
Central	18	1	30					**49**
Copperbelt			62					**62**
Eastern		50						**50**
Luapula			40					**40**
Lusaka		49						**49**
Muchinga			45					**45**
North Western				20	13	26		**59**
Northern			42					**42**
Southern	50							**50**
Western							40	**40**
Total schools by language	**68**	**100**	**219**	**20**	**13**	**26**	**40**	**486**

Table 5. Number of sampled schools, by rural/urban and province

PROVINCE	NO. OF SCHOOLS		
	RURAL	URBAN	UNKNOWN
Central	35	14	0
Copperbelt	18	40	4
Eastern	45	4	1
Luapula	35	4	1
Lusaka	9	40	0
Muchinga	39	6	0
North Western	51	8	0
Northern	36	6	0
Southern	46	3	1
Western	33	7	0
Total	**347**	**132**	**7**

Table 6. Number of head teacher questionnaires completed, by province

PROVINCE	NO. OF HEAD TEACHER QUESTIONNAIRES
Central	49
Copperbelt	60
Eastern	49
Luapula	40
Lusaka	46
Muchinga	45
North Western	56
Northern	42
Southern	50
Western	39
Total	**476**

Table 7. Number of teacher questionnaires, by province

PROVINCE	NO. OF TEACHER QUESTIONNAIRES
Central	47
Copperbelt	62
Eastern	50
Luapula	40
Lusaka	47
Muchinga	45
North Western	57
Northern	41
Southern	50
Western	40
Total	479

Table 8. Number of school information questionnaires, by province

PROVINCE	NO. OF SCHOOL QUESTIONNAIRES
Central	49
Copperbelt	58
Eastern	50
Luapula	40
Lusaka	47
Muchinga	45
North Western	59
Northern	35
Southern	50
Western	40
Total	**473**

3.3 WEIGHTING

There were three stages of weighting so that the sample of pupil scores could be representative of the overall pupil performance for the desired subpopulations. First, the districts in the sample were weighted so they represented the strata in which they were selected. As noted above, the strata were created by province crossed with language of instruction.

$$\text{Stage 1 weight} = \frac{number\ of\ pupils\ in\ strata\ 1}{number\ of\ pupils\ in\ district \times number\ of\ districts\ slected}$$

Second, the schools in the districts were weighted so they also represented the schools in their respective strata. The strata within district were created by school types found within districts: private schools, grant-aided schools, and community/GRZ schools.

$$\text{Stage 2 weight} = \frac{number\ of\ pupils\ in\ selected\ strata\ 2}{number\ of\ pupils\ in\ school \times number\ of\ schools\ selected\ in\ strata\ 2}$$

Finally, the pupils were weighted so they represented all of the pupils in their respective schools, using the formula:

$$\text{Stage 3 weight} = \frac{number\ of\ pupils\ selected}{number\ of\ pupils\ enrolled\ in\ school}$$

The final calculated weight was at the pupil level and was a product of the three previously mentioned weights.

4 RESULTS AND FINDINGS

4.1 INSTRUMENT RELIABILITY AND VALIDITY

The research team carried out psychometric analyses that were used to determine validity and reliability characteristics of the Zambia instrumentation. This evaluation was conducted before the full data collection (i.e., during the pilot assessment) and detailed information was generated as a result of the psychometric analyses. This information was then used to modify the pilot instrumentation to produce the best assessment tools possible. During the psychometric evaluation, Cronbach's alpha values for the English EGRA indicated acceptable internal consistency on average, $\alpha = 0.85$. For the EGRAs in other languages, each instrument was assessed as shown in ***Table 9.***

Table 9. Internal consistency and reliability check of instruments by language

LANGUAGE	CRONBACH'S ALPHA
Chitonga	0.74
Cinyanja	0.82
Icibemba	0.80
Kiikaonde	0.80
Lunda	0.87
Luvale	0.75
Silozi	0.82

Generally, an alpha score over 0.70 is acceptable and a value over 0.80 is considered very good. In addition, construct validity was assessed by examining the item hierarchy, or the ordering of items within a subtask from easy to difficult that results from an item-level analysis during the Rasch measurement.

Statistics such as these can show how well a set of variables measures an underlying construct, and in the present study, they suggested that the different subtasks of the Zambia EGRA all contributed to measuring early grade pupils' reading knowledge.

4.2 EGRA RESULTS

4.2.1 ORGANIZATION OF THE FINDINGS

This section presents EGRA results for each subtask by language (for local-language subtasks) or by province (for English subtasks). Tables showing mean scores by language and pupil sex are presented here, while more comprehensive tables, which include confidence intervals, are provided in ***Annex A.***

It is important to note that comparisons across languages are not a focus of this report. Although differences may seem apparent, further study would have to be conducted in order to identify explanations of the differences. Also note that there are confounding issues—such as proportion of private schools, and urban/rural

location—that can distort the interpretation of attempted comparisons across languages.

4.2.2 SUMMARY OF SCORES

Listening Comprehension in Local Languages

Table 10 on the following page summarizes pupil performance on the Listening Comprehension subtask, with results disaggregated by language and pupil sex.

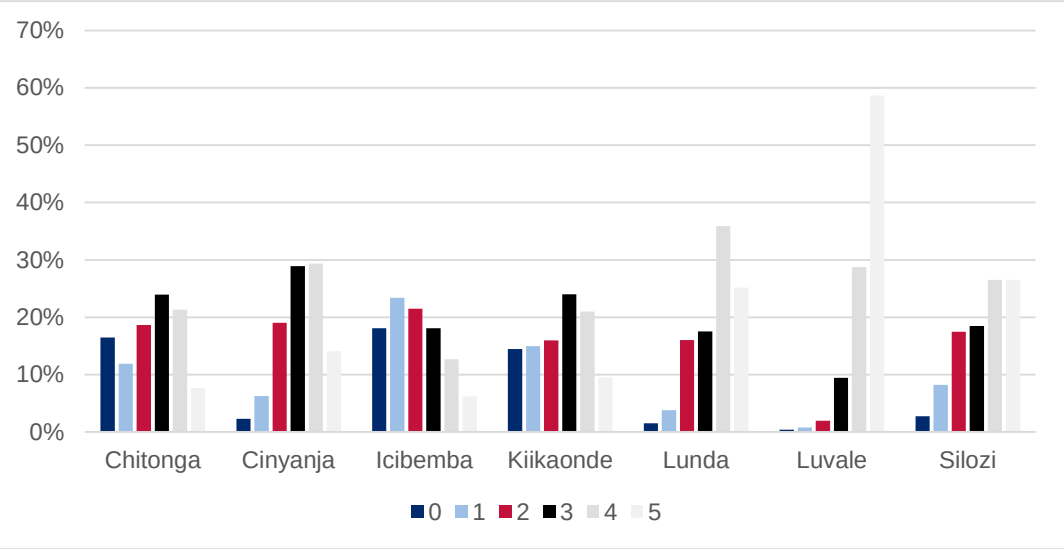
Overall performance on this subtask showed limited ability to correctly respond to comprehension questions based on an oral story presented in the local language of the province in which the pupils lived. Looking at overall percentages out of the total number of questions asked as well as percentages correct out of the number of items attempted, pupils tested in Luvale were able to respond to nearly all questions on average (4.27 items correct overall), with boys and girls performing comparably. For all other languages, pupil scores were lower; pupils tested in Chitonga, Icibemba, and Kiikaonde on average correctly answered fewer than three (less than 55%) of the total questions.

Table 10. Overall performance on Listening Comprehension (local languages), by language and pupil sex

	NO. CORRECT RESPONSES (OUT OF 5 QUESTIONS)			% CORRECT / ATTEMPTED			CORRECT / TOTAL			% ZERO SCORES		
LANGUAGE	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga	2.49	2.42	2.55	50%	48%	51%	50%	48%	51%	15%	17%	14%
Cinyanja	3.27	3.28	3.25	65%	66%	65%	65%	66%	65%	2%	2%	2%
Icibemba	2.04	2.16	1.93	41%	43%	39%	41%	43%	39%	17%	14%	20%
Kiikaonde	2.65	2.78	2.51	53%	56%	50%	53%	56%	50%	10%	11%	9%
Lunda	3.82	3.90	3.76	76%	78%	75%	76%	78%	75%	1%	0%	2%
Luvale	4.27	4.36	4.18	85%	87%	84%	85%	87%	84%	0%	0%	0%
Silozi	3.40	3.32	3.48	68%	66%	70%	68%	66%	70%	2%	2%	2%

The distributions of scores across languages (see ***Figure 1***) show this trend, as well, with pupils tested in Chitonga, Icibemba, and Kiikaonde generating the highest proportion of zero scores, and pupils tested in Luvale generatingi the highest proportion of perfect (5 out of 5 questions correct) scores (59%).

Figure 1. Language distributions of scores on Listening Comprehension subtask (local languages)



This was the first activity in which pupils engaged, and anxiety could have artificially lowered scores. That said, the ability to understand and respond to oral language is a foundational language skill upon which the comprehension of written materials rests, and further attention to these scores is warranted.

Letter Sound Identification and Nonword Reading

Being able to identify the sound(s) associated with each letter is a prerequisite to sounding out, or decoding, unfamiliar words. ***Table 11*** summarizes fluency and zero scores on these two fundamental pre-reading subtasks, with results disaggregated by language and pupil sex.

SAMPLE OF CINYANJA LETTER-SOUND ITEMS:

K m u A d n J K
b A d E f t R h

Table 11. Overall performance on Letter Sound Identification and Nonword Reading (local language), by language and pupil sex

	LETTER SOUND IDENTIFICATION						NONWORD READING					
	NO. CORRECT LETTER SOUNDS / MINUTE			% ZERO SCORES			NO. CORRECT NONWORDS / MINUTE			% ZERO SCORES		
LANGUAGE	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga	3.68	3.32	4.04	57%	57%	56%	1.66	1.56	1.76	85%	86%	85%
Cinyanja	7.21	7.35	7.00	47%	48%	47%	3.20	3.59	2.75	72%	70%	74%
Icibemba	8.77	9.02	8.53	35%	34%	36%	5.60	5.85	5.37	62%	60%	64%
Kiikaonde	5.44	5.84	5.02	57%	61%	53%	2.27	2.29	2.24	76%	73%	79%
Lunda	9.63	12.57	7.06	28%	20%	36%	6.33	9.67	3.41	50%	36%	62%
Luvale	6.75	6.18	7.28	42%	43%	41%	4.28	4.25	4.31	73%	70%	76%
Silozi	8.09	7.40	8.75	44%	46%	42%	5.65	5.10	6.17	56%	59%	53%

Just as pupil performance on Listening Comprehension was relatively low across most languages, so was performance on Letter Sound Identification. Only in Lunda were pupils able to, on average, correctly identify close to 10 letters in the minute allotted, which itself is a rate of one letter identified every six seconds. The trend as reflected in the zero scores also highlights this skill deficiency, with between 28% (Lunda) and 57% (Kiikaonde and Chitonga) of pupils scoring zero on this subtask.

Not surprisingly, performance on the more challenging task of reading nonwords was lower than performance on identifying letter sounds; what is surprising is that it was lower by only a few items. Because learning the alphabet is one of the first early literacy skills mastered, pupils should be expected to perform much higher on a measure of letters than on a more complex task of manipulating those letter sounds to form words. This was not the case, however, as evidenced through a comparison of score distributions on the two subtasks. While the distribution of scores on Letter Sound Identification ***(Figure 2)*** were less heavily weighted on zero scores than were the Nonword Reading scores ***(Figure 3),*** the overall distributions were comparable, with substantial percentages of pupils scoring zero on these tasks.

SAMPLE OF CHINYANJA NONWORD ITEMS:

Gelu mdzimu nipe kelo atapi umbe ninda ninane

Figure 2. Score distributions of Letter Sound Identification

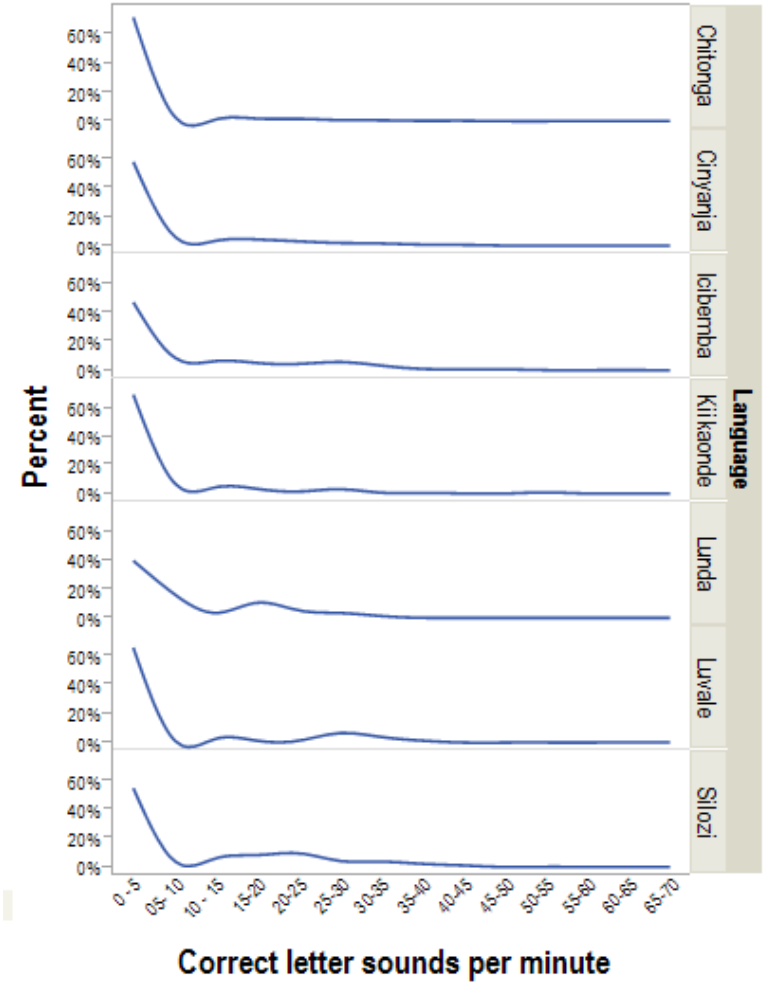
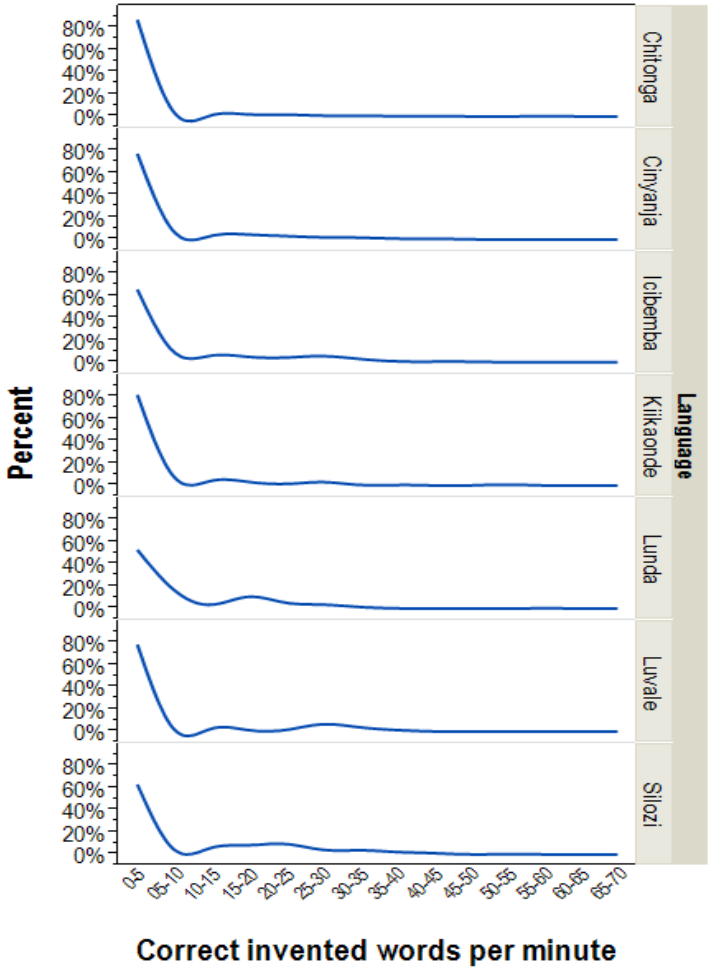


Figure 3. Score distributions of Nonword Reading



Oral Reading Fluency

The end goal of literacy development is to read connected text at a fast and accurate enough level to allow for comprehension of what is read. ***Table 12*** on the following page presents pupil performance on the Oral Reading Fluency subtask, with results disaggregated by language and pupil sex. The table initially reports the average correct number of words per minute by language.

LUVALE READING PASSAGE:

Kwapwile lunga naphwevo vatwaminenga ukhawavo mumusenge. Vawanyinenga kukaluhwa mukuyoya chavo mwomwo kwakulimina kwapwile kwauchi. Kuyoya chavo chapwile chakukwata tuswa nakulya uchi. Echi chapwile nakuneha ushona kuli phwevo chikuma. Lunga ashinganyekele akuya nakutunga kwakamwihi navausoko wenyi. Chiyoyelo chavo chalumukile kaha vathu vevwile kuwaha chikuma.

Reading connected text is often easier than reading isolated words, given the contextual cues that exist in connected text. It is not surprising, therefore, that overall, pupils performed slightly better on this subtask than they did on the Nonword Reading subtask. This result is typical of contexts in which pupils are taught to read familiar words by rote, without going through the process of learning to decode words. However, performance in general was still quite low, with over half of pupils in all languages but Lunda scoring zero on this task. In the languages of Icibemba, Lunda, and Silozi, pupils overall were able to read between 5 and 10 words correctly; in Chitonga, however, 88% of pupils could not read any words correctly and pupils overall read fewer than two words correctly on average.

A comparison of distributions of scores across languages repeats this trend, with ***Figure 4*** below showing a large proportion of pupils scoring zero on this task. ***Table 12*** allows a comparison of the data by language and by pupil sex.

Figure 4. Language distributions of scores on Oral Reading Fluency subtask (local languages)

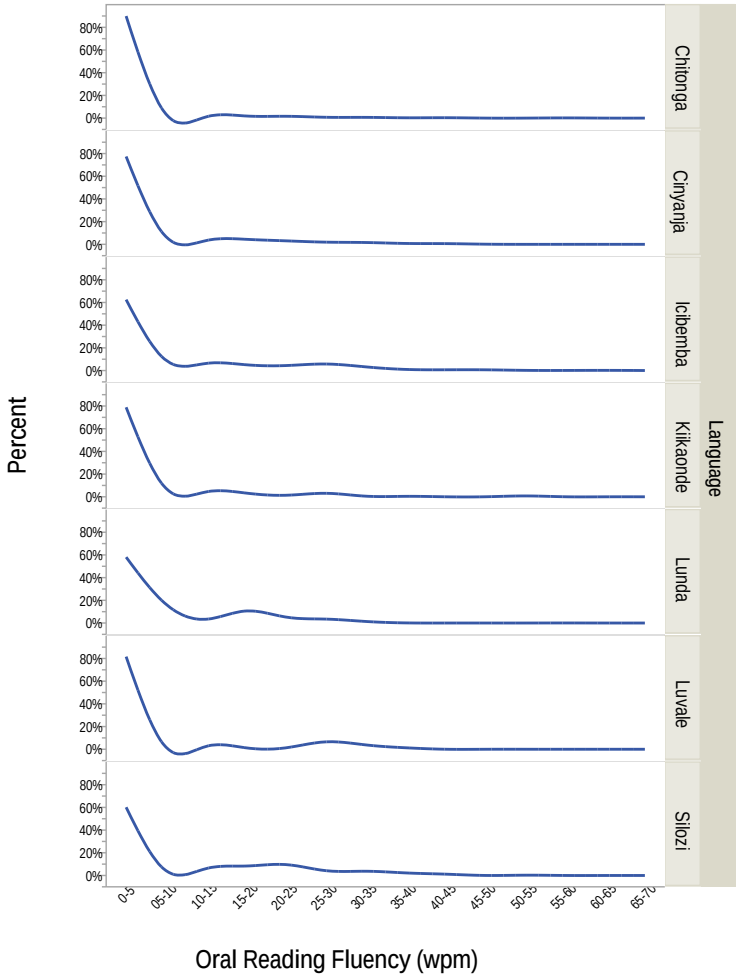


Table 12. Overall performance on Oral Reading Fluency (local languages), by language and pupil sex

	NO. CORRECT WORDS/ MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
LANGUAGE	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga	1.85	1.93	1.77	8%	7%	8%	4%	4%	4%	88%	88%	88%
Cinyanja	4.11	4.35	3.80	21%	23%	19%	10%	11%	9%	64%	65%	64%
Icibemba	7.62	7.85	7.40	34%	36%	33%	16%	17%	16%	52%	49%	55%
Kiikaonde	3.46	2.75	4.19	18%	16%	20%	6%	5%	7%	74%	74%	74%
Lunda	6.28	9.25	3.74	35%	50%	23%	15%	22%	8%	46%	34%	57%
Luvale	4.32	4.26	4.38	16%	16%	17%	10%	10%	10%	80%	82%	79%
Silozi	8.40	7.96	8.83	37%	36%	37%	25%	24%	27%	56%	56%	57%

Table 13 shows oral reading fluency estimates by language, disaggregated by urban/rural. The estimates for urban schools in the Lunda and Luvale languages have been suppressed due to the sample sizes being too small, thus the estimates being unreliable as they come from less than 10 pupils. Chitonga, Kiikaonde, and Silozi had statistically significant higher estimates for the urban schools than the rural schools. While the difference was not statistically significant, Cinyanja had a much higher estimate for rural schools (5.14 WPM) compared with urban (1.92 WPM). This is explained by the fact that 40 schools were sampled in Lusaka district; the migration of people into this area has created a situation where a large percentage of pupils in that district do not have the same mother tongue as the school language of instruction, Cinyanja.

Table 13. Overall performance on Oral Reading Fluency, by language and urban/rural

	ORAL READING FLUENCY, NO. CORRECT WORDS / MINUTE	
	RURAL	URBAN
Chitonga	1.66	6.67*
Cinyanja	5.14	1.92
Icibemba	7.18	7.97
Kiikaonde	2.61	11.46*
Lunda	6.32	#
Luvale	4.51	#
Silozi	7.31	15.34*

suppressed, sample size too small (≤10).

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Orientation to Print

One consideration in interpreting pupil performance on the Oral Reading Fluency subtask is a pupil's awareness of how to actually engage with text on a page—for example, knowledge that one reads in local languages from left to right and must track text from the end of one line to the beginning of another. ***Table 14*** presents pupil performance on the three components of the Orientation to Print subtask, with results disaggregated by language and pupil sex.

Table 14. Overall Performance on Orientation to Print (local languages), by language and pupil sex

	NO. CORRECT RESPONSES (OUT OF 3 QUESTIONS)			% CORRECT / TOTAL			% ZERO SCORES		
LANGUAGE	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga	2.41	2.42	2.40	80%	81%	80%	7%	8%	6%
Cinyanja	2.59	2.63	2.55	86%	88%	85%	8%	8%	9%
Icibemba	2.50	2.56	2.45	83%	85%	82%	9%	8%	11%
Kiikaonde	2.25	2.29	2.21	75%	76%	74%	10%	10%	9%
Lunda	2.94	3.00	2.89	98%	100%	96%	1%	0%	2%
Luvale	2.55	2.68	2.43	85%	89%	81%	8%	4%	12%
Silozi	2.34	2.34	2.33	78%	78%	78%	13%	12%	14%

Overall, pupils performed well on this subtask. All languages averaged at least 75% correct. Performance on this subtask suggests that pupils did understand how to engage with print on a page, and that deficiencies in awareness of orientation to print cannot be considered a major factor in poor Oral Reading Fluency performance.

Reading Comprehension

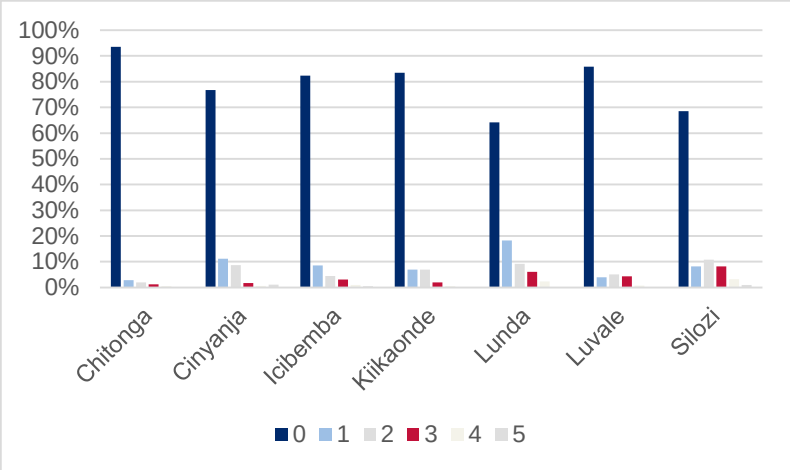
Once having read the Oral Reading Fluency passage, pupils were orally asked five questions about the passage. Pupils were asked only questions that corresponded with the text read, so, for example, pupils who read through only the first 10 words of the passage would be asked just questions that corresponded with those first 10 words of the passage. ***Table 15*** displays pupil performance on the Reading Comprehension subtask, with results disaggregated by language and pupil sex.

Table 15. Overall performance on Reading Comprehension (local languages), by language and pupil sex

	NO. CORRECT RESPONSES (OUT OF 5 QUESTIONS)			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
LANGUAGE	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga	0.13	0.14	0.12	4%	5%	3%	3%	3%	2%	94%	93%	94%
Cinyanja	0.36	0.39	0.33	11%	11%	10%	7%	8%	7%	79%	77%	81%
Icibemba	0.33	0.34	0.32	15%	16%	14%	7%	7%	6%	82%	82%	83%
Kiikaonde	0.21	0.17	0.26	7%	6%	8%	4%	3%	5%	89%	91%	86%
Lunda	0.63	0.89	0.40	14%	20%	9%	13%	18%	8%	62%	49%	73%
Luvale	0.36	0.33	0.39	9%	8%	9%	7%	7%	8%	84%	87%	82%
Silozi	0.74	0.73	0.75	22%	21%	23%	15%	15%	15%	68%	68%	69%

As stated earlier, overall on the Oral Reading Fluency subtask (Figure 4 above), pupils had been able to read between two and nine correct words per minute (WPM). Therefore, they were asked, on average, no more than one comprehension question (each passage was written to have one reading comprehension passage correspond with approximately every 10 words). Even among those pupils who were given and attempted a question, only between 4% (Chitonga) and 22% (Silozi) got it correct. This result shows not only a lack of ability to read connected text quickly and with accuracy, but also the related inability to focus attention from reading individual words to making sense of the meaning of the passage. This result is further illustrated by a comparison of distributions across languages, in ***Figure 5.***

Figure 5. Language distributions of scores on Reading Comprehension subtask (local languages)



English EGRA Subtasks

To explore pupils' proficiency in English, they were given three English-language subtasks. For each of these subtasks, although the items were in English, the instructions were in the local language of the pupil to ensure understanding of the task.

Table 16 illustrates pupil performance on Letter Name Identification, English Vocabulary, and Listening Comprehension subtasks. As indicated earlier, these results were disaggregated by the regions in which pupils were tested.

Interestingly, even while they were tested on letter names in English, pupil performance identifying letter names dramatically exceeded their performance identifying letter sounds in their local languages—in local languages, correct letter sounds identified overall ranged from 3.68 (Chitonga language) to 9.63 (Lunda

language) per minute as seen earlier in Table 11, whereas correct letter names identified in a minute ranged from 7.57 in Muchinga province to 22.78 in Copperbelt province.

Table 16. Overall performance on English-language subtasks, by language and pupil sex

PROVINCE	LETTER NAME IDENTIFICATION (100 LETTERS)		ENGLISH VOCABULARY (20 ITEMS)		LISTENING COMPREHENSION (5 QUESTIONS)	
	NO. CORRECT LETTERS PER MINUTE	% ZERO SCORES	% CORRECT WORDS / TOTAL	% ZERO SCORES	% CORRECT RESPONSES / TOTAL	% ZERO SCORES
Central	17.70	26%	47%	0%	19%	64%
Copperbelt	22.78	27%	56%	0%	33%	42%
Eastern	14.09	34%	34%	1%	9%	82%
Luapula	9.50	33%	36%	1%	7%	81%
Lusaka	21.33	14%	55%	1%	40%	32%
Muchinga	7.57	54%	36%	0%	7%	84%
North Western	13.42	37%	37%	0%	15%	69%
Northern	11.85	33%	43%	0%	11%	74%
Southern	13.32	33%	39%	0%	10%	77%
Western	13.75	38%	39%	0%	12%	74%

This differentiation in performance levels suggests that pupils found the Letter Name Identification activity easier than the Letter Sound Identification activity. It is likely that the academic focus in early grades impacts this trend (e.g., if pupils are taught to learn letter names rather than letter sounds, they more readily identify letter names when tested), although this theory should be explored further before final interpretations are drawn from these data.

The items selected for the English Vocabulary subtask were common body parts (e.g., foot, arm), objects in the environment (e.g., pencil, shoes), and prepositions (e.g., on, behind). Performance on the English Vocabulary subtask was relatively higher than on other subtasks, with few pupils scoring zero (e.g., only 1% of pupils in the Eastern, Luapula, and Lusaka provinces). Even so, only in the two most urban regions (Copperbelt, 56%; Lusaka, 55%) did pupils correctly respond to at least half of the 20 total items. Although many pupils were able to understand at least one-third of these items, overall, the pupils did not demonstrate full familiarity with these English words.

Given this demonstrated lack of understanding of basic English vocabulary, it is not surprising that many pupils were not able to correctly respond to comprehension questions asked of a story related in English. With the exception of Copperbelt and Lusaka provinces—the two provinces in which pupils performed

the best on the English Vocabulary subtask—more than half of pupils in all provinces scored zero on the English Listening Comprehension subtask. In Copperbelt province, only 42% of pupils scored zero while pupils overall correctly responded to 33% of questions; in Lusaka province, even fewer (32%) pupils scored zero while overall they correctly responded to 40% of questions. This differential performance in the Copperbelt and Lusaka regions suggests that English may be more directly taught in those early primary classrooms than in other provinces.

Table 17 shows the English subtask estimates by urban and rural disaggregation. Pupils in urban schools scored better than their rural counterparts, on all subtasks. For correct letters per minute, they scored 26.8 and 12.1 for urban and rural, respectively. Although Table 12 showed different oral reading fluency results for urban/rural for local languages, English proficiency is very dependent on whether the pupils hear and practice English in everyday life, which is far more likely in urban areas, as demonstrated in ***Table 17,*** where the two urban provinces, Copperbelt and Lusaka, performed the best.

Table 17. Overall performance on English-language subtasks, by urban/rural

	NO. CORRECT ITEMS / MINUTE		% CORRECT / ATTEMPTED		% CORRECT / TOTAL		% ZERO SCORES	
	RURAL	URBAN	RURAL	URBAN	RURAL	URBAN	RURAL	URBAN
Letter Name Identification	12.1	26.8***	36%	60%***	12%	27%***	35%	15%***
Vocabulary	7.8 out of 20	11.8 out of 20***	39%	59%***	39%	59%***	0%	1%
Listening Comprehension	0.6 out of 5	2.1 out of 5***	11%	42%***	11%	42%	75%	29%***

*$p < 0.05$, **$p < 0.01$, ***$p > 0.001$.

There were differences by urban and gender subpopulations for the English subtasks, as shown in ***Table 18.*** These highly significant detected differences were in the zero scores for Letter Name Identification and Listening Comprehension, where boys had a significantly lower percentage of zero scores.

Table 18. Overall performance on English-language subtasks, by gender

	NO. CORRECT ITEMS / MINUTE				% CORRECT / ATTEMPTED				% CORRECT / TOTAL				% ZERO SCORES			
	RURAL		URBAN		RURAL		URBAN		RURAL		URBAN		RURAL		URBAN	
	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS	BOYS	GIRLS
Letter Name Identification	12.5	11.7	25.4	28.1	38%	34%*	58%	62%	13%	12%	25%	28%	32%	38%**	16%	15%
Vocabulary	7.9 out of 20	7.7 out of 20	11.6 out of 20	11.9 out of 20	39%	39%	58%	59%	39%	39%	58%	59%	0%	0%	1%	0%
Listening Com-prehension	0.6	0.5	2.1	2.1	12%	11%	41%	43%	12%	11%	41%	43%	73%	78%*	29%	29%

*$p < 0.05$, **$p < 0.01$, *** $p < 0.001$.

The performance on the English subtasks followed a very predictable pattern when explored by school type ***(Table 19).*** Private schools showed very high scores compared with community, GRZ, and grant-aided schools. The correct letters per minute of 43.2 for private schools reflects strong proficiency in this skill, while the mostly rural community schools scored the lowest at 12.7. While it is clear that many pupils in Zambia can correctly identify English letter names and English vocabulary, the English Listening Comprehension subtask really showed that many pupils were struggling to comprehend conversational English and respond to it, with percent zero scores of 72%, 66%, 56%, and 17% for school types community, GRZ, grant-aided, and private, respectively.

Table 19. Overall performance on English-language subtasks, by school type

	LETTER NAME IDENTIFICATION		VOCABULARY		LISTENING COMPREHENSION	
	NO. CORRECT LETTERS PER MINUTE	% ZERO SCORES	% CORRECT WORDS / TOTAL	% ZERO SCORES	% CORRECT RESPONSES / TOTAL	% ZERO SCORES
Community (*n*=640)	12.7	39%	40%	0.7%	13%	72%
GRZ (*n*=2674)	15.0	29%	43%	0%	17%	66%
Grant-aided (from GRZ) (*n*=551)	17.4	19%	47%	0.3%	22%	56%
Private (*n*=445)	43.2	9%	77%	0%	62%	17%
Unknown (*n*=524)	17.4	33%	48%	0.5%	26%	57%

4.3 RESULTS FROM THE COMPLEMENTARY INSTRUMENTS

In order to monitor how well an education system is performing, it is necessary to collect data and report on student learning outcomes. But the response to improve these learning outcomes in Zambia should emanate from a robust theory of change that includes school and teacher characteristics (e.g., training) as well as learning materials (e.g., curricula, textbooks) and teaching practices. Thus, while it is important to track how well children can read, it is equally important to track various nonreading indicators that can peer into possible reasons why children are not learning and that can also test one's underlying theory of change.

4.3.1 LEARNER BACKGROUND AND CHARACTERISTICS

Learner Home Backgrounds

The complementary instruments collected data on the home backgrounds of learners, including their age, socioeconomic status (SES), and local language. As

the latter (local language) has already been presented, this section does not discuss it further. Pupils were asked to identify household items that their family possessed, as well as the means of obtaining water and cooking meals. Responses to these questions were tabulated and used to calculate an SES index. Learners were ranked according to this index and divided into five groups of approximately equal size (quintiles). ***Table 20*** presents the proportion of learners in each SES quintile and school type.[5]

Table 20. Socioeconomic backgrounds of learners, by school type

	SCHOOL TYPE				
SES QUINTILE	COMMUNITY	GRZ	GRANT-AIDED	UNKNOWN	TOTAL
Lowest quintile	3.8%	12.4%	2.3%	0.4%	2.3%
Lowest quintile	3.8%	12.4%	2.3%	0.4%	2.3%
Quintile 2	2.7%	11.4%	2.1%	0.5%	2.2%
Quintile 3	2.6%	12.2%	2.5%	0.7%	2.1%
Quintile 4	2.8%	12.1%	2.1%	0.8%	2.1%
Highest quintile	1.7%	7.2%	2.2%	6.7%	2.0%
Total	**13.6%**	**55.3%**	**11.2%**	**9.2%**	**10.7%**

As seen in the table, most respondents (i.e., 55.3%) were in government schools (GRZ), although significant proportions were also in community, grant-aided, and private schools (i.e., 13.6%, 11.2%, and 9.2%, respectively). Approximately 10.8% of learners were in schools whose type was not identified in the data set due to lack of information in the census data.

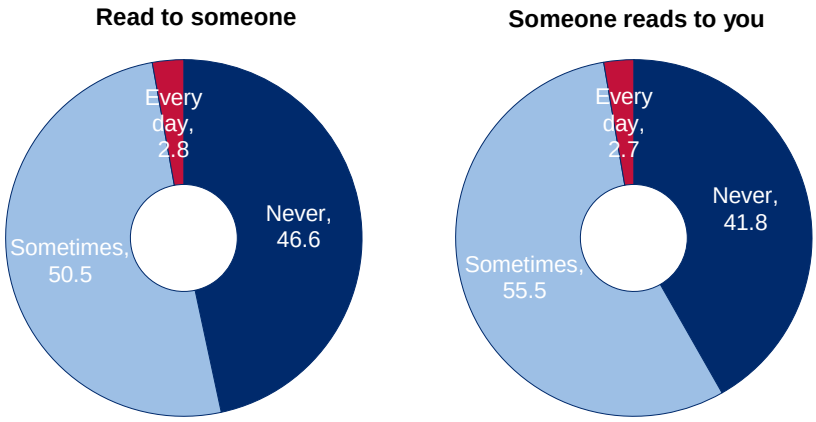
Learners were asked whether they had eaten breakfast before coming to school the morning of the assessment and most of them (54.5%) had not. Moreover, only 31.3% of pupils reported that their school had a feeding program, indicating that at least one in four learners had not eaten anything before beginning school.

Learners tended not to have material (e.g., books, newspapers) at their home that they could read: approximately two in five pupils (i.e., 60.6%) indicated that they did not have reading materials. The lack of home reading material for most learners undoubtedly influenced the frequency with which learners reported reading to others at home and others reading to them (see ***Figure 6***). While

[5] SES quintiles do not contain precisely 20% of the sampled population due to the distribution of pupil responses to household questionnaire items and the subsequent calculations of an SES index.

approximately half of learners asserted that they read to someone at home or someone else read to them sometimes (50.5% and 55.5%, respectively), almost no learners reported reading or being read to on a daily basis (2.8% and 2.7%, respectively). Further, a substantial proportion of learners admitted to never reading or being read to at home (46.6% and 41.8%, respectively).

Figure 6. Frequency of out-of-school reading among grade 2 children

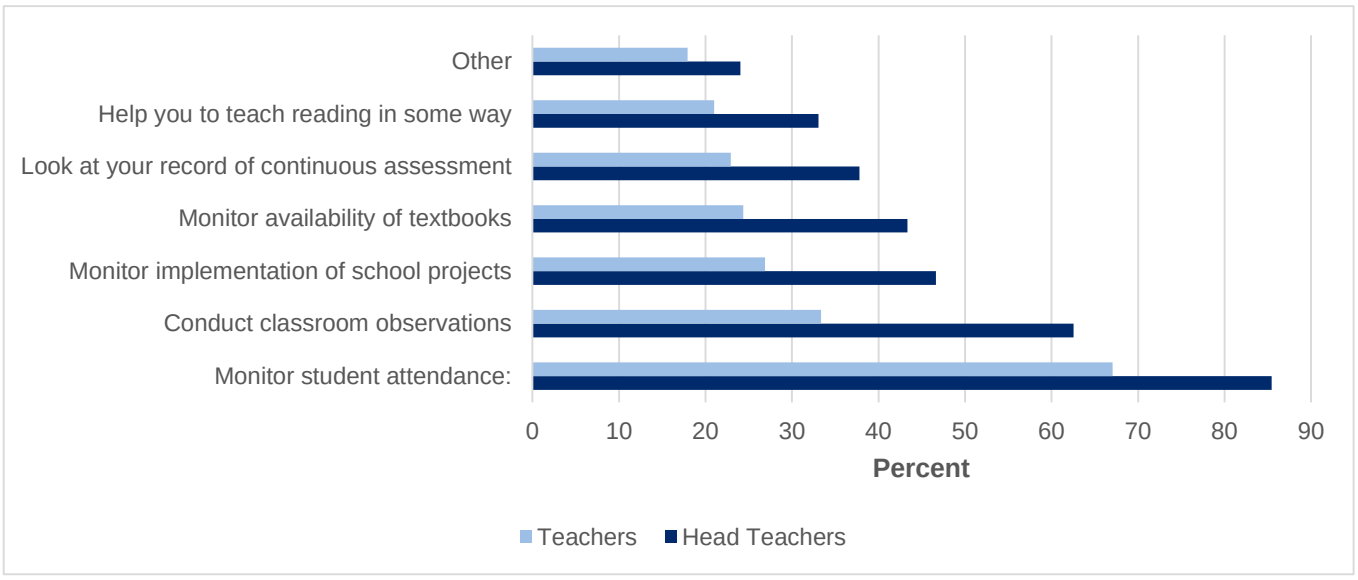


The complementary instruments also collected data on learners' academic backgrounds, including questions on previous classes, preschool attendance, and current patterns of household academic activity. Most learners (90.9%) had been in grade 1 during the previous year, although approximately 8.4% reported having been in grade 2 and were therefore repeating the grade. Preschool attendance was commonly reported among learners (71.5% of learners reported going), but this differed significantly among provinces. Preschool attendance ranged from 91.6% in North Western province to 56.2% in Lusaka; attendance rates were lower than 65% in three other provinces as well.

Maternal and paternal literacy rates, as reported by learners in the pupil interview, were relatively high: 87.2% of learners reported that their fathers could read and 80.0% asserted that their mothers could read. These reported rates of adult literacy differed by province, however. In three provinces (Eastern, Luapula, and North Western), approximately one in three learners asserted that their mother could not read. The lowest reported paternal literacy rate, on the other hand, was 78.9% in Muchinga; in two provinces (Copperbelt and Northern), over 90% of learners reported that their fathers could read.

A key aspect of improved oversight and management in education is the role that the community plays. Communities in Zambia have been trained to (1) monitor attendance of teachers and learners to ensure that teachers report for classes on time, stay, and teach regularly; (2) ensure that learners report for school on time (by checking the attendance registers); (3) conduct class observations—that is, observe lessons; (4) monitor the implementation of school projects (agreed upon during community meetings); (5) monitor the use of teaching and learning materials; (6) monitor and manage human, material, and financial resources; and (7) monitor the school environment to ensure that it is conducive to teaching and learning. To this end, head teachers and teachers were asked whether learners' parents participated in various aspects of school and classroom life, which provides an indication of family involvement at sampled schools. ***Figure 7*** displays the proportion of head teachers and teachers who affirmed that parents participated at school or in the classroom, as well as the activities parents tended to assist with. The most common task parents assisted with were monitoring pupil attendance (85.5% and 67.0% of head teachers and teachers reported help with this) and conducting classroom observations (62.6% and 33.4% of head teachers and teachers reported assistance). Other, less common, activities parents were involved with were monitoring school projects or the availability of textbooks, or looking at assessment records. Of note is that, when asked, approximately 72% of teachers did not find parental or community support helpful, which may be a reflection of the fact that parents and community members appeared mainly to be assisting with administrative, rather than pedagogical, tasks.

Figure 7. Parent participation in schools and classrooms



4.3.2 TEACHER AND CLASSROOM CHARACTERISTICS

The complementary instruments administered along with the EGRA attempted to capture data on the characteristics of teachers and classrooms that related to teaching and the teaching process. These can be grouped into the following categories: teacher characteristics, instructional feedback on learner performance, and teachers' pedagogical moves.

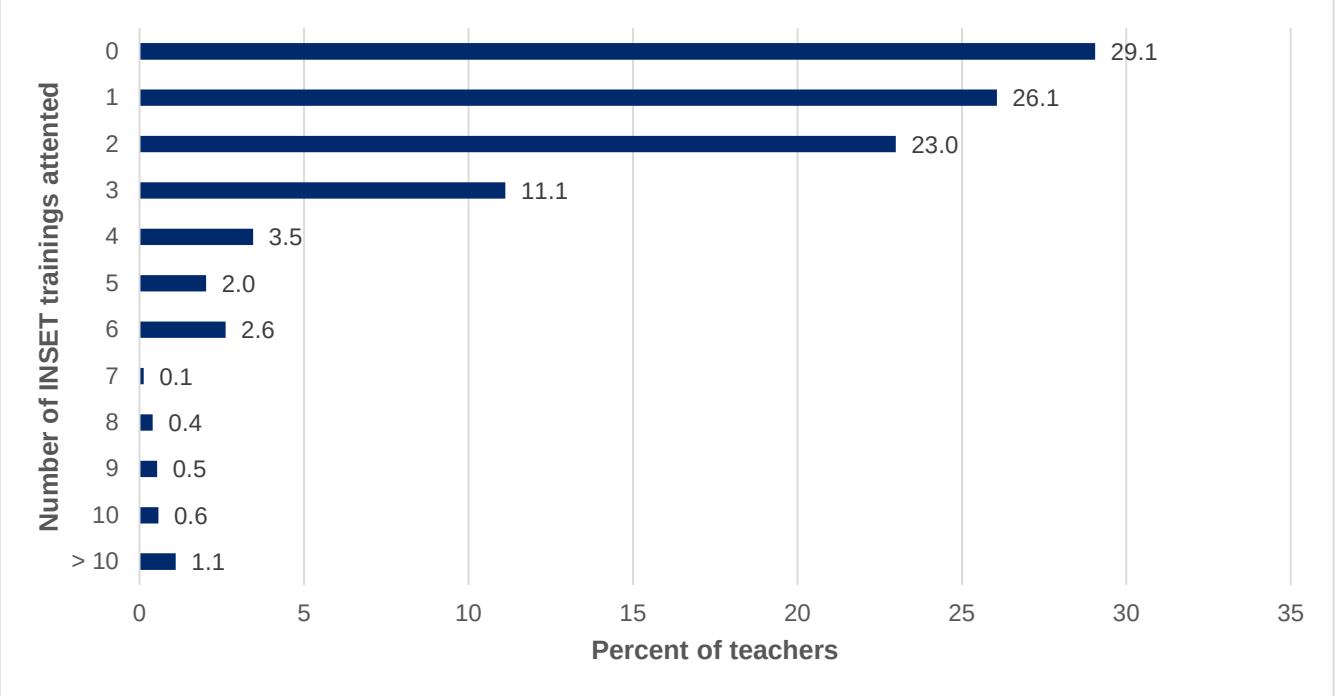
Teacher Characteristics

Most teachers in surveyed schools had completed grade 12 as their highest level of education (86.5%) and were holders of primary teaching certificates (70.0%). However, 17.4% of teachers did indicate having no formal teaching credential. In terms of years of experience, the teacher workforce tended to be rather "green" (inexperienced). ***Figure 8*** presents the distribution of teachers' years of experience in Zambia. As can be seen from the figure, the distribution of experience is positively skewed with most teachers (51%) reporting having five years of experience or less, and more than one-third of teachers (34%) having three years or less. This overall lack of experience in the profile of the teaching workforce carries implications for educational spending (less experienced teachers tend to be paid lower wages), as well as training requirements (less experienced teachers may require more instructionally specific training). To the extent that it is indicative of high rates of teacher turnover, the lack of experience among teachers may be a regular feature of the education system in Zambia rather than a current trend.

Figure 8. Teachers' years of experience

One implication of a "green" teacher workforce is a consequent lack of in-service training. In other words, less experienced teachers may have undertaken less training than their more experienced colleagues, which may be problematic if pre-service training is also insufficient or if in-service training is aligned to national educational priorities and initiatives (new teachers may not be trained in these). Perhaps unsurprisingly, ***Figure 9*** shows that the vast majority of teachers surveyed for this study reported attending very few, if any, in-service education and training (INSET) sessions; approximately 55% of teachers reported having attended one INSET training session or none at all, and less than 22% reported attending more than two training sessions. If the median teacher surveyed for this study had five years of experience (see Figure 8), Figure 9 suggests that this same median teacher had only attended one INSET training session. Needless to say, one training session over the course of five years is insufficient to substantially enhance teachers' attitudes, beliefs, content knowledge, and pedagogical expertise. Nevertheless, it is worth noting that 90.3% of teachers who had attended INSET trainings described them as "very useful."

Figure 9. Number of INSET training sessions teachers have attended

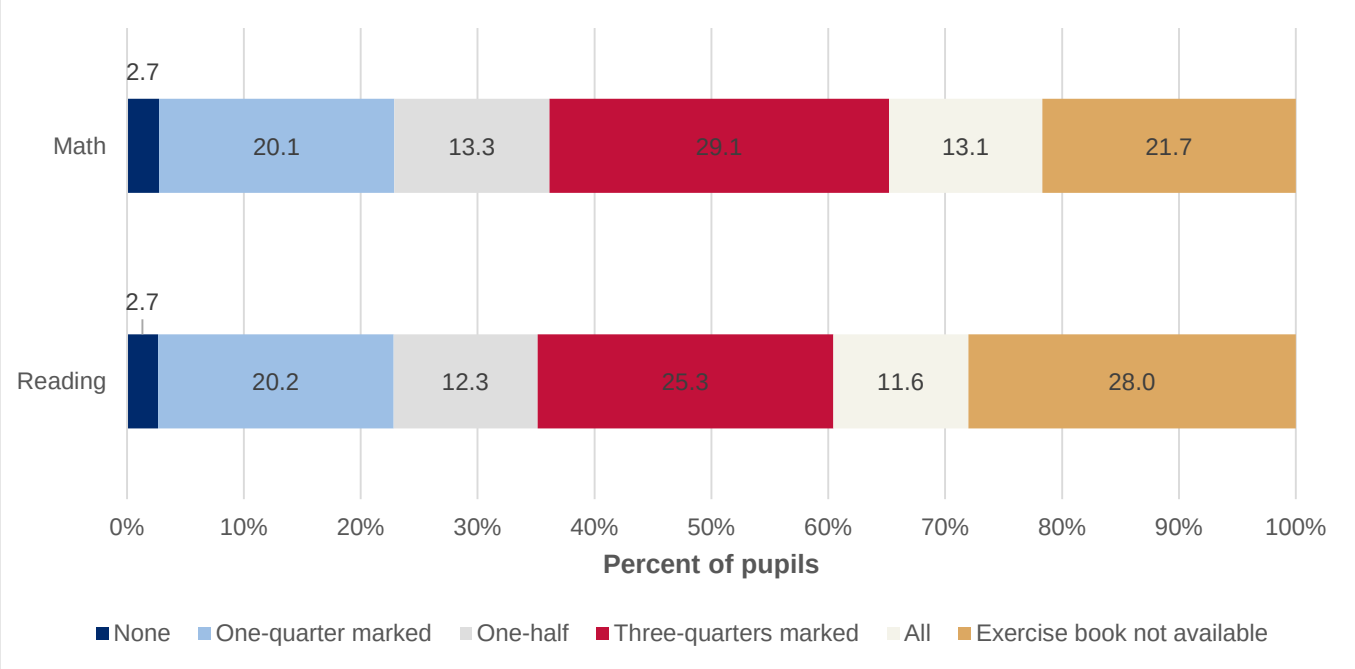


Instructional Feedback to Learners

Although teacher characteristics are helpful for creating a profile of the teaching workforce in Zambia and making certain personnel decisions, they are not inherently indicative of teachers' classroom practices. In other words, teachers' levels of skill, content knowledge, or pedagogical expertise does not inherently increase linearly with years of experience or as teachers collect certifications. As such, the complementary instruments in Zambia also attempted to collect data on teachers' classroom practices and interactions with learners.

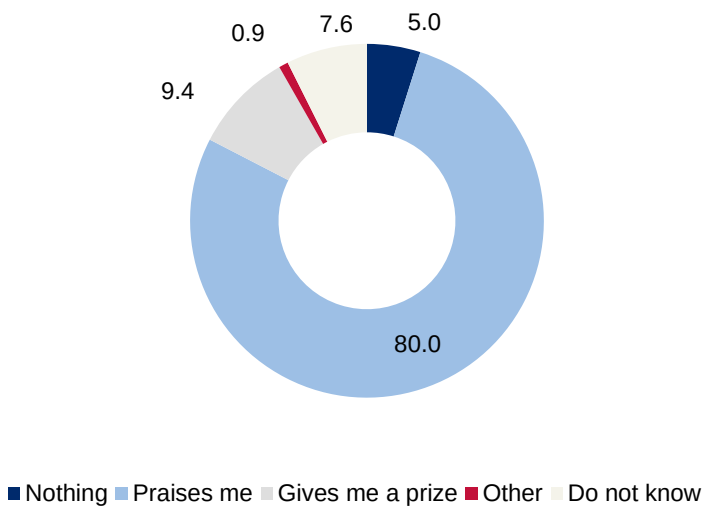
One example of providing instructional feedback to learners is when teachers mark and assess learners' work. ***Figure 10*** depicts the proportion of learners whose reading and mathematics exercise books were marked by their teachers, as well as the estimated proportion of pages marked by the teacher. As shown in the figure, there is a range in terms of how many pages teachers tended to mark: Less than half of pages were marked in approximately one-third of pupil exercise books, while more than half of pages were marked in another third of books. It should be noted that this assessment of teacher marking does not indicate whether learners had completed different amounts of the exercise book and if, therefore, teacher marking patterns reflected the proportion of the exercise book that had been completed. Nevertheless, teachers apparently differed in terms of the amount of feedback they provided their pupils via written corrections (Henderlong & Lepper ,2002).

Figure 10. Teacher marking of learner work



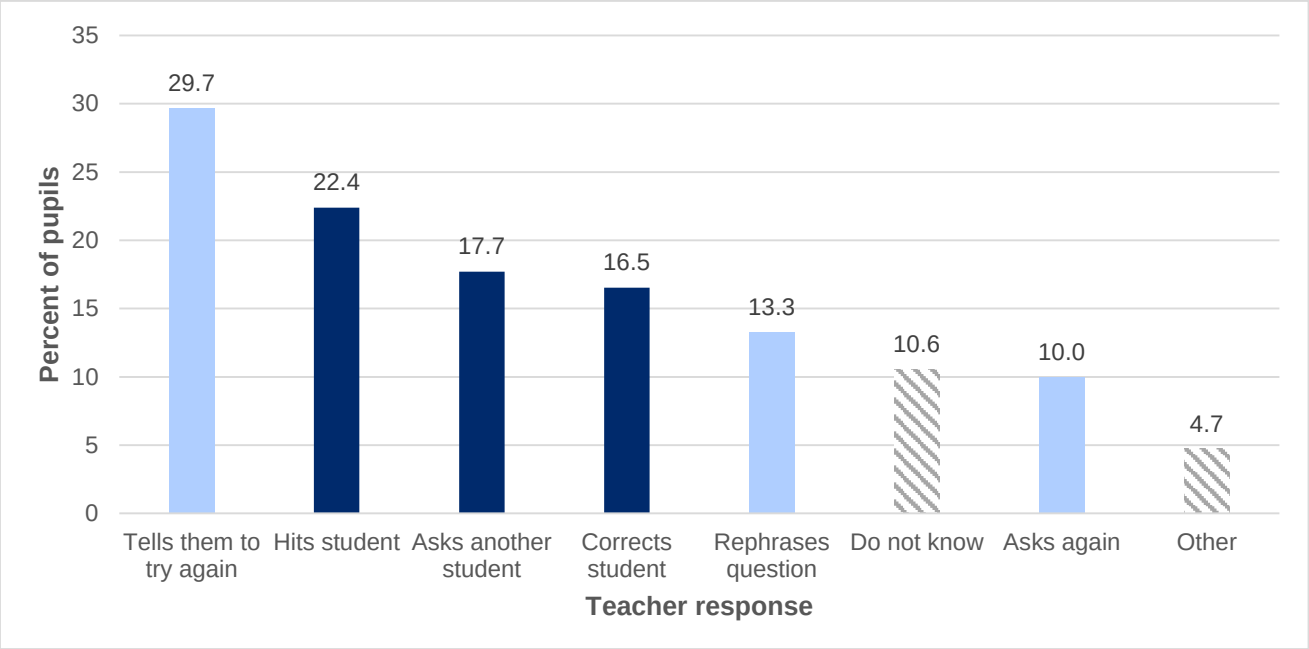
Teachers also provide feedback, either verbally or nonverbally, directly or indirectly, to learners based on their academic performance ***(Figure 11).*** The pupil exit survey asked learners to state their teachers' most common reaction when they exhibited good performance. By far, the most common reaction to good academic performance was for teachers to praise learners (80.0%), followed by giving them a prize (9.4%). While these actions are positive and direct responses to pupil performance (as opposed to passive responses such as doing nothing), it is not clear that they constitute "good" pedagogical actions. Providing learners with a prize in exchange for good academic performance may reinforce a tendency toward extrinsic motivation rather than an intrinsic love for learning. Praise from teachers, too, can vary in terms of its effects on learner beliefs as well as future effort and performance. It has been shown that praise that is based on effort and the processes of learning emphasize the importance of hard work and dedication. "Person" praise, on the other hand (e.g., "You are so smart", or "Great job, you are so good at reading"), can have detrimental effects on learners' views of the role of effort in learning as well as their self-concept as learners.

Figure 11. Teacher reaction to good learner performance



Teachers also provide learners with instructional feedback in response to learners' answers to questions, whether they are correct or incorrect. During these in-class interactions, teachers can give instructional feedback that is active and constructive (i.e., directly responds to the answer given and facilitates learner engagement with the lesson content), or feedback that is passive and destructive (i.e., does not directly respond to the answer given, or displays negative emotion and does not facilitate learner engagement). A question in the pupil exit interview attempted to discern these patterns in teachers' responses to answers that learners offered in class: learners were asked how their teachers normally responded when they offered incorrect answers in class. ***Figure 12*** tabulates learners' reports according to the frequency of teachers' reported behavior. In the figure, dark bars represent passive or destructive teacher responses, while light bars represent active or constructive responses. As shown in the figure, slightly more learners reported that their teachers responded passively or destructively (56.6%) than responded actively and constructively (53%). However, it is worth noting that the most frequent response to incorrect learner answers was the teacher asking learners to try again. While this response is not active (i.e., it does not directly respond to the learner's answer but rather insinuates that it is incorrect), it does encourage learners' reengagement in the exchange by asking them to think again and provide another answer. On the other hand, the relative frequency of passive (i.e., asks another pupil) and destructive (i.e., hits pupil, corrects pupil) teacher responses is not a hallmark of effective teaching practice, and suggests that teachers could make some progress in building a more robust repertoire of active and constructive responses.

Figure 12. Teacher responses to incorrect learner answers



Teachers' Pedagogical Moves

In addition to feedback on learners' work in and out of class, teachers also exhibit "pedagogical moves" (i.e., specific actions they take with an instructional purpose) in the classroom. While teaching and learning was not directly observed for this report, several questions in the teacher and pupil interview attempted to quantify these moves. For example, the teacher interview asked teachers to report the amount of continuous assessment of reading that was being carried out in their classroom. Continuous assessment provides a means to diagnose learners' abilities in reading and to adjust instructional practices based on the results. When asked, approximately 74.1% of teachers reported that they had conducted reading assessment exercises with their pupils in the past month, and most of those who responded affirmatively had assessed more than 30 pupils in this way during the past month. Though it is not clear how, or indeed if, teachers were using the results of this assessment, the act of using assessment to diagnose learners' abilities in reading in itself constitutes "improved" teaching.

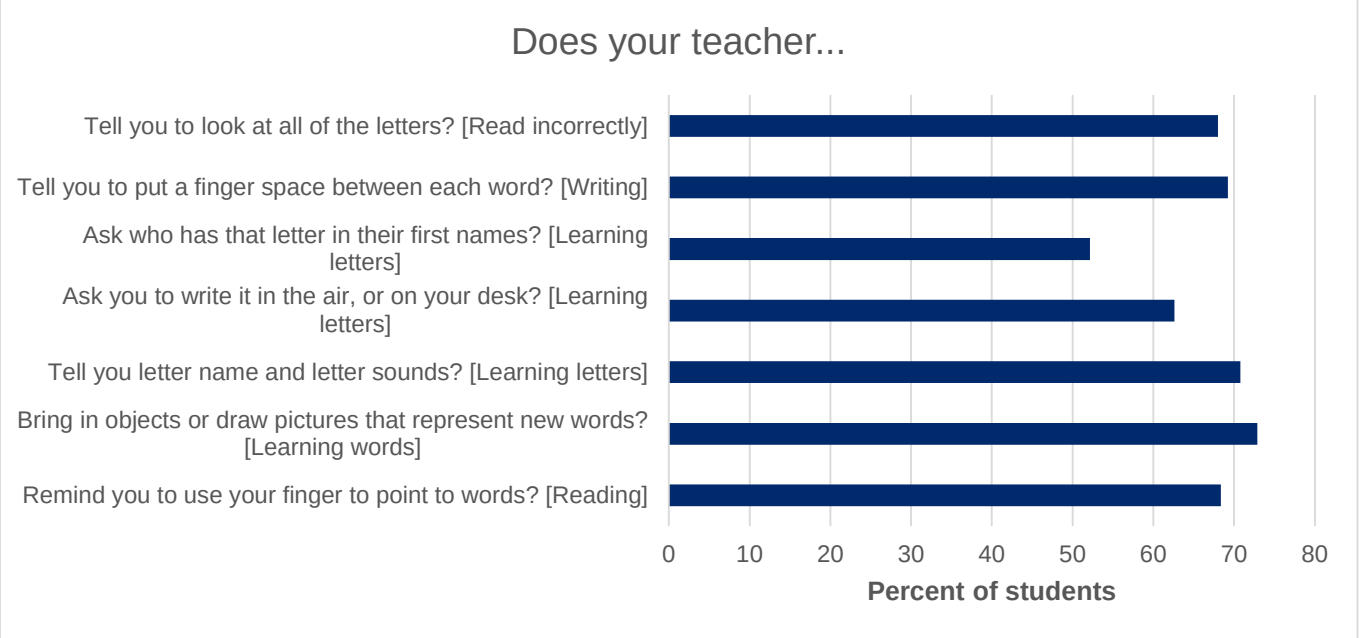
The pupil interview also attempted to discern whether teachers were teaching early grade reading appropriately on a regular basis, by asking pupils about their teachers' instructional actions when they were engaged in specific literacy activities. Learners were asked:

- When you read a word wrong, does your teacher tell you to look at all of the letters?

- When you write, does your teacher tell you to put a finger space between each word?
- When you learn a letter, does your teacher ask what pupils in the class have that letter in their first names?
- When you learn a letter, does your teacher ask you to write it in the air, or on your desk, with your fingers?
- When you learn a letter does your teacher tell you the letter name and letter sounds?
- When you learn new words, does your teacher bring in objects or draw pictures on the chalkboard that represent the words?
- Does your teacher remind you to use your finger to point to words when you read?

Figure 13 presents learners' responses to these questions. Overall, learners indicated that their teachers tended to employ these practices in class, as affirmative responses ranged from 52.2% to 72.9%. While this does not amount to a direct observation of teaching processes, it is nonetheless suggestive of teachers' general practices and whether they tend to use certain pedagogical moves. These results suggest that most teachers in sampled schools tended to use these specific instructional moves more often than not, but some variation in teaching practices exists. Moreover, these data do not describe how often teachers tended to employ these practices.

Figure 13. Teachers' pedagogical prompts to pupils



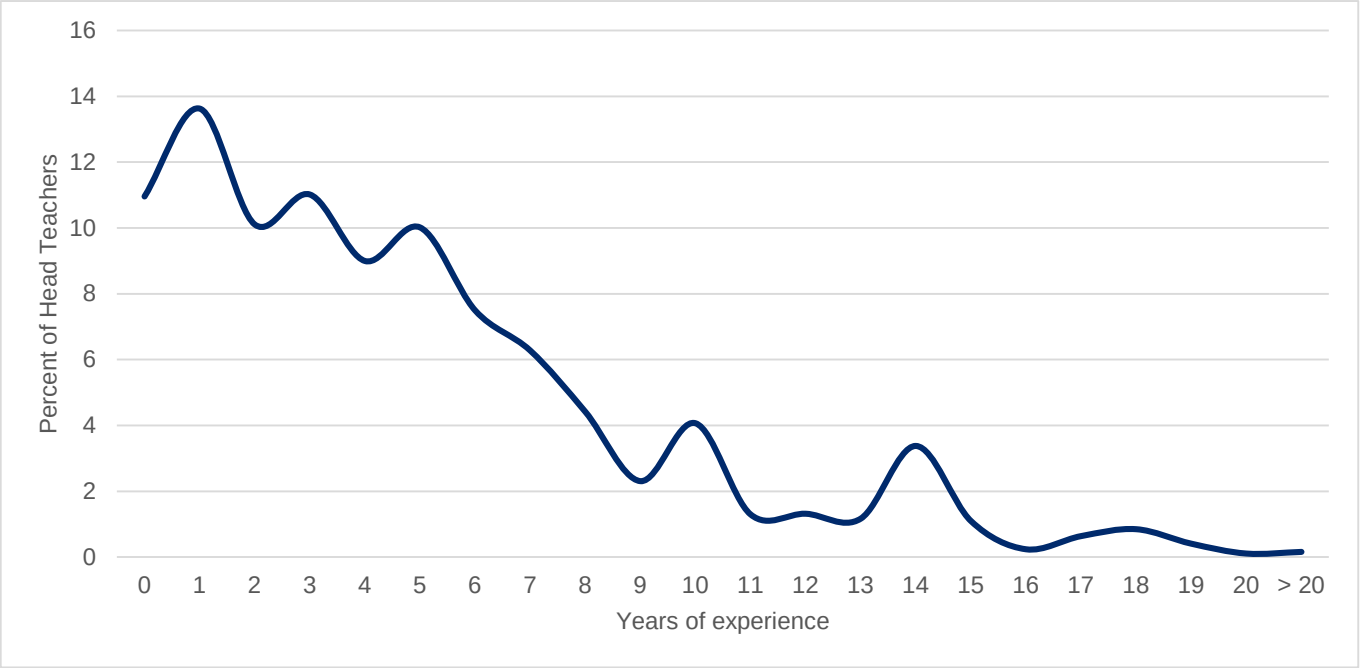
4.3.3 HEAD TEACHER AND SCHOOL CHARACTERISTICS

The complementary instruments collected data on several aspects of head teachers and of schools in Zambia. These aspects can be classified as: head teacher characteristics, school resources, classroom contexts and inputs, and pedagogical oversight.

Head Teacher Characteristics

Similar to teacher characteristics discussed above, the head teacher interview collected data on characteristics that can be used to create a typical profile of head teachers in Zambia. Similar to the classroom teachers, head teachers surveyed for this study tended to report that grade 12 was their highest level of education completed (83.5%), and most reported holding either a primary (49.5%) or secondary (26.1%) teaching credential. Head teachers were also asked to report their years of experience***; Figure 14*** displays the distribution of experience among surveyed head teachers. As was seen in the distribution of experience among teachers (see Figure 8), Figure 14 shows a relatively inexperienced head teacher workforce. Indeed, 34.7% of head teachers counted two years or less of experience, and only 35.3% reported having more than five years of experience. Further, the most frequent response was one year of experience (13.6% of head teachers reported having one year of experience).

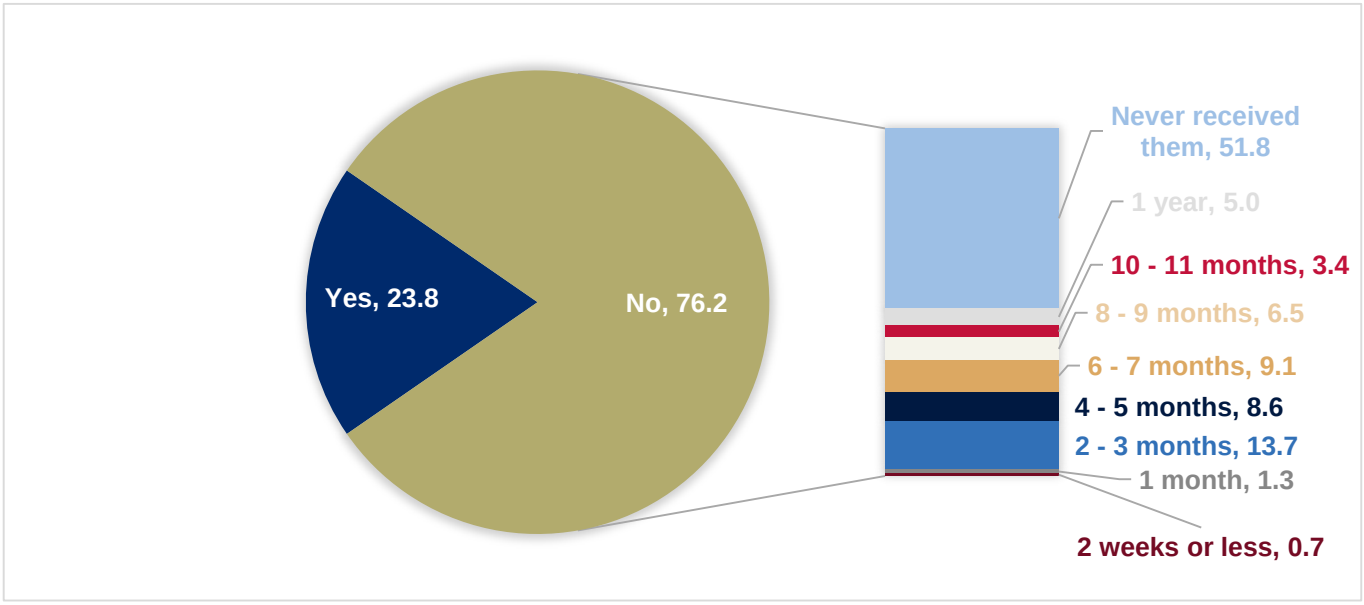
Figure 14. Head teachers' years of experience



School Resources

Teaching and learning also depend in part upon school resources available to teachers and to learners. In this respect, basic school oversight and management should ensure that curricular materials (e.g., textbooks, supplementary materials) are in the hands of every learner. The head teacher questionnaire asked whether the school began the academic year with the appropriate number of books; ***Figure 15*** presents head teachers' responses. It is evident that most (76.2%) schools did not have the appropriate number of books at the beginning of the academic year. Further, more than half of these schools (51.8%) never received the books, and a further 46.2% had to wait at least two months before receiving them. While this phenomenon is relatively common, it does suggest that more than one-third of schools never received an appropriate allocation of textbooks, and the vast majority of schools were made to wait before they were able to use them.

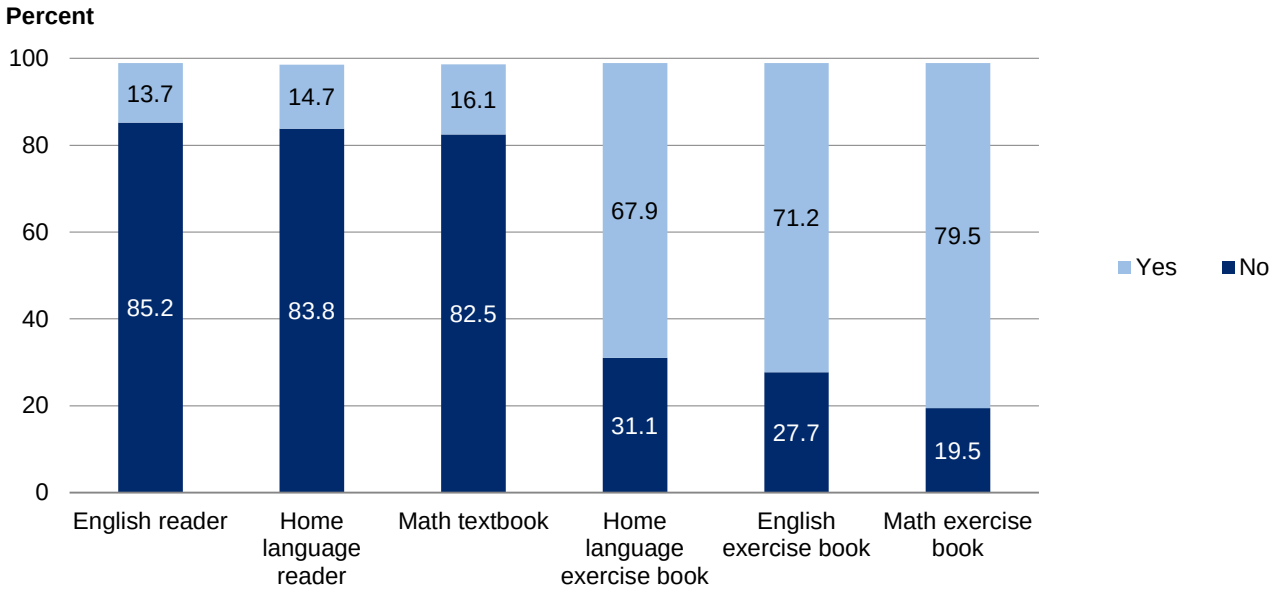
Figure 15. Starting the academic year with the appropriate number of books



Assessors also determined whether learners had certain types of books: textbooks and exercise books in learners' local language, English, and mathematics. ***Figure 16*** displays the proportion of learners who had access to these materials on the day of the interview. As expected, given inefficiencies with book allocation (see Figure 15), most learners (i.e., more than 80%) did not have their own mathematics textbook, English reader, or local language reader. More common was possession of exercise books: more than two-thirds of learners were in possession of local language (67.9%), English (71.2%), and mathematics exercise books (79.5%). A minority of learners (between 20% and 30%), however, did not have access to exercise books. While the mere presence of books or access to them does not necessarily indicate that the textbooks, readers, and exercise books are used well from a curricular or pedagogical point of view, or indeed that they are of sufficient quality to warrant their use in the first place, it does indicate efficient management of instructional materials.

This lack of curricular and instructional materials might be made up for, in part, by the presence (and use) of a school library. However, the majority of schools surveyed for this report (71.7%) did not have a library, and it is unclear whether the 28.3% of schools that did actually used them for instructional purposes.

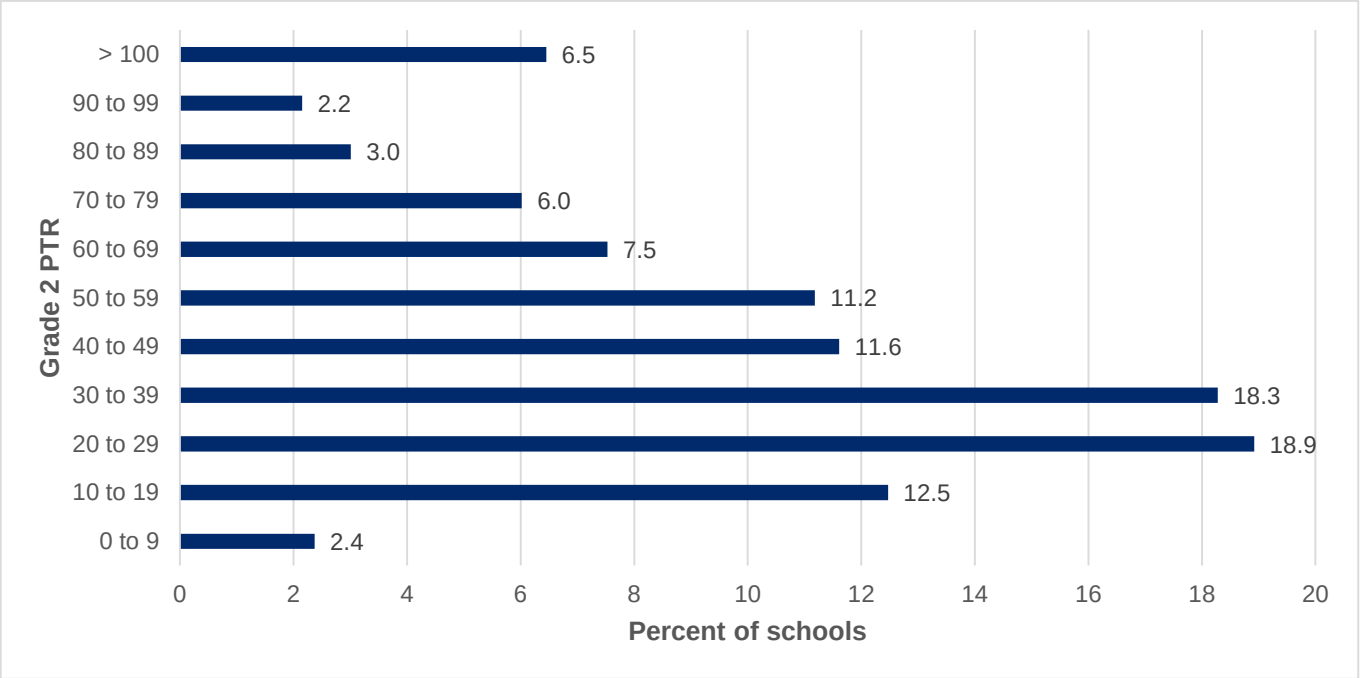
Figure 16. Learner access to textbooks and exercise books



Classroom Contexts and Inputs

Teachers constitute the most significant investment made at the classroom level. While previous sections have discussed teacher characteristics and pedagogical moves, this section focuses on teachers as classroom inputs; that is, the number of teachers in schools and pupil–teacher ratios (PTRs). These indicators, again, do not directly shed light on the processes of teaching (i.e., what teachers do in classrooms and how well they do it), but they do interconnect with teaching practice in that teachers with more learners may be more restricted in terms of instructional approaches. ***Figure 17*** displays the distribution of grade 2 PTRs for schools surveyed for this report, calculated as the number of learners enrolled in grade 2 divided by the number grade 2 teachers. From the figure, it is evident that a slight majority of schools had reasonable PTRs: approximately 52% of schools had PTRs less than the acceptable ratio of 40:1. On the other hand, this finding also implies that nearly one-half of schools had PTRs higher than 40:1; and more than one-third of schools had PTRs in excess of 50:1, including 6.5% with ratios of more than 100:1. This, no doubt, is partly because most schools surveyed for this report (86.8%) had only one grade 2 teacher, and only 5.7% had more than two teachers. At the same time, median grade 2 enrollment was 43 learners, and extended to well over 100 in numerous schools.

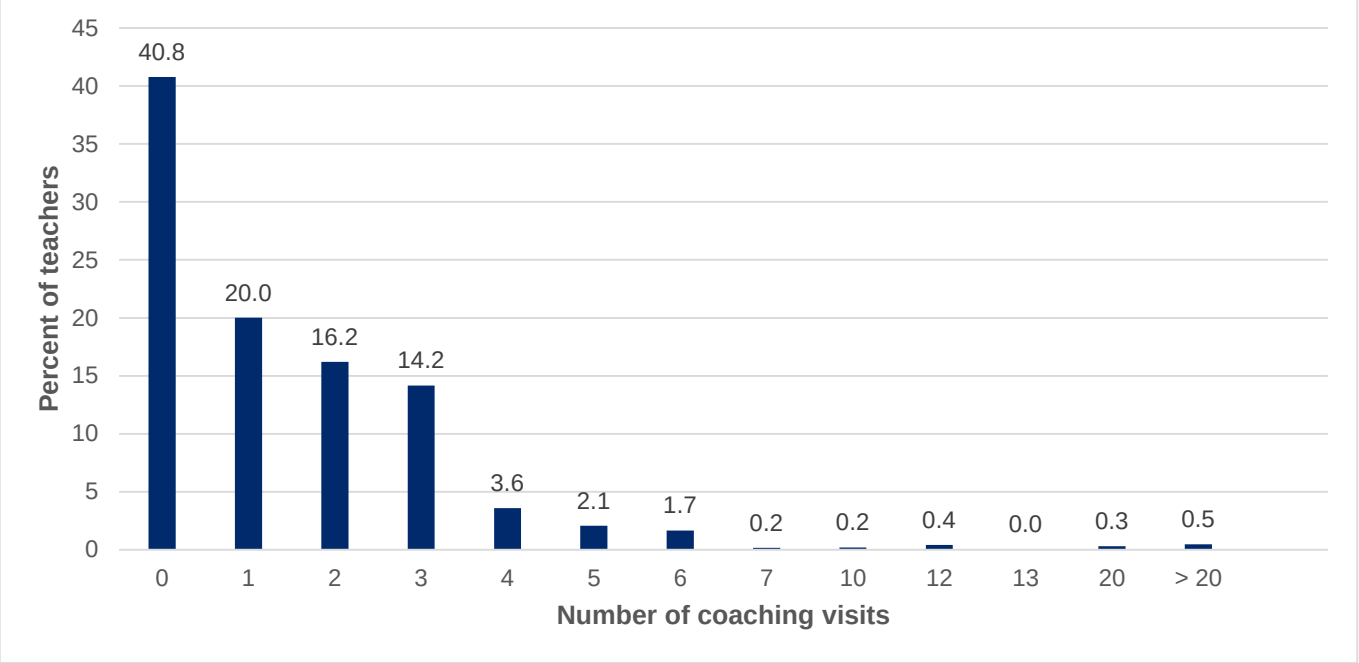
Figure 17. Pupil-teacher ratios in grade 2



Pedagogical Oversight

Another aspect of school-level factors that influence teaching and learning is instructional leadership. If teachers are to enhance their instructional practice and become familiar with new, more ambitious types of pedagogy, they require instructional leaders and, often, coaching. As such, the extent to which teachers receive pedagogical support by instructional coaches is a direct indication of the extent to which systems of education are attempting to advance the skills of their teaching workforce. In this connection, the complementary instruments attempted to discern the number of coaching visits received by teachers, the time allocated to instructional coaching, and teachers' perceptions of the usefulness of the coaching. ***Figure 18*** displays the distribution of reading coaching visits received by grade 2 teachers surveyed for this report. As can be seen in the figure, a significant minority of teachers (40.8%) had not received a single reading coaching visit during the previous academic year, and the average (median) teacher had only received one. Further, only 23% of teachers reported receiving more than two visits during the previous academic year. These findings suggest that, for many teachers, instructional coaching is not a common occurrence; indeed, for 77% of teachers, access to a coach was limited to, at most, two sessions per year. Such infrequency, if not combined with other forms of informal coaching, does not provide teachers with timely feedback on their instructional practice necessary for them to improve.

Figure 18. Number of reading coaching visits received by teachers per year

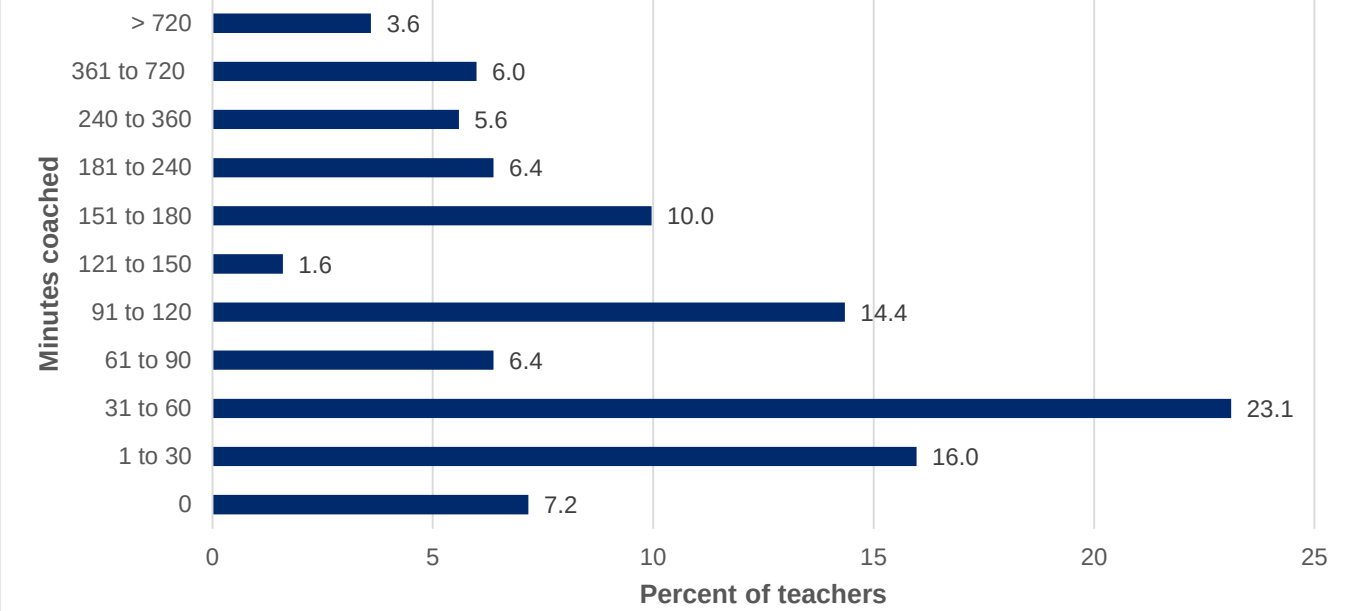


The number of coaching visits, however, may have masked teachers' access to instructional coaching if the length of coaches' visits differed substantially. In fact, average lengths of coaching visits, as reported by teachers, were found to vary extensively, with a reported range of 0–480 minutes. Therefore, total coaching time was calculated for all teachers who reported having at least one coaching visit during the past academic year.[6] ***Figure 19*** presents the results of this calculation. Of those teachers who reported having received at least one coaching visit during the previous year, the average (median) total coaching time was 90 minutes, and nearly one-third of coached teachers received more than 150 minutes of coaching. While this represents a nontrivial amount of time and resources devoted to instructional coaching, it is important to remember that for most teachers who received some coaching (i.e., approximately 75%), these minutes were spread over only a few sessions (three or less) during the year. Moreover, the average coaching time received by teachers drops substantially when one includes those teachers who did not receive any coaching visits (i.e., the 40.8% in Figure 18 who did not receive a single visit). By this calculation, the average (median) amount of coaching minutes received by teachers drops to less than 30 minutes, and nearly 45% of teachers surveyed for this report received no instructional coaching at all. Thus, although instructional coaching is being implemented in Zambia and most teachers are receiving some level of

[6] This calculation involved multiplying the reported number of coaching visits by the average length of each visit, as reported by teachers. Note that some teachers reported receiving coaching visits during the previous academic year, but asserted that the average duration of these visits was zero minutes.

instructional support, both the extent and duration of this support range widely and are allocated to teachers disparately. Further, more than half of teachers surveyed for this report received less than 30 minutes of instructional coaching over the previous academic year, an amount of support that is unlikely to facilitate any degree of instructional change.

Figure 19. Total coaching time received by coached teachers per year



These data do not indicate, however, whether the instructional support received by teachers was of adequate quality, as data were not collected on the interactions between coaches and teachers. Nevertheless, teachers' perceptions of the usefulness of the coaching was elicited by the teacher questionnaire. Teachers who had received at least one instructional coaching visit during the previous year were asked whether they felt the support was "very useful," "somewhat useful," or "not at all useful." All teachers reported that they felt the coaching received to be "very useful" (89.3%) or "somewhat useful" (10.7%); not a single teacher reportedly perceived the coaching to be "not at all useful." Again, this is not a direct measure of the quality of instructional coaching, but these findings at least suggest that teachers who receive instructional support consider the sessions to be a constructive use of time.

4.3.4 FACTORS ASSOCIATED WITH LEARNER PERFORMANCE

Previous sections have reported descriptive statistics and created profiles of learners, teachers, classrooms, head teachers, and schools. These data were collected, reported, and discussed because they are assumed to be related to the processes of teaching and learning, as well as learner achievement on the EGRA, and therefore relevant to school leaders and policy makers. In that regard, these profiles of learners, teachers, classrooms, head teachers, and schools are valuable in and of themselves. However, it is also worth determining whether any of the variables previously discussed were significantly associated with learner performance on the EGRA subtasks.

SES and Learner Performance

The socioeconomic status of learners is often shown to be highly predictive of academic outcomes. Learners in relatively wealthier households are, for example, more likely to have access to additional extracurricular learning resources (e.g., books, tutoring), to attend preschool, and to be healthier, all of which are associated with learning outcomes. This SES difference often translates into higher achievement in academic settings. ***Table 21*** displays average performance on two EGRA subtasks, by learner SES quintile.

Table 21. Performance on reading subtasks, by SES quintile

	READING			
	ORAL READING FLUENCY (LOCAL LANGUAGE)		READING COMPREHENSION	
QUINTILE	AVG. (WPM)	% ZERO SCORES	AVG. SCORE	% ZERO SCORES
Lowest (Q1)	4.84	69.00%	0.28	84.80%
Q2	5.17	64.68%	0.34	80.79%
Q3	5.45	61.98%	0.34	81.54%
Q4	5.09	61.77%	0.32	83.47%
Highest (Q5)	7.74	50.52%	0.42	78.44%

As expected, academic performance as measured by the EGRA increased with learner SES: learners from wealthier households tended to perform better (i.e., average scores were higher and the proportion of zero scores was lower) on both subtasks presented in Table 21. In reading, performance was not substantially better by SES quintile except for learners in the highest (wealthiest) quintile. Average performance on the Oral Reading Fluency and Reading Comprehension subtasks was markedly higher for learners from the wealthiest households, and zero scores tended to be much lower. Reading performance among learners in

other quintiles, however, did not differ greatly. Therefore, it appears as though the household backgrounds of learners (i.e., relative SES) may be associated with reading performance, as measured by the EGRA. However, it should be noted that none of these quintiles can rightly be characterized as performing "well" in reading. Indeed, all quintiles recorded low average scores and high proportions of zero scores.

Regression Analyses of Learner Performance

While Table 21 suggests that SES is predictive of learner achievement in reading, a finding which will be of interest to both school leaders and policy makers, schools traditionally can do little about relative household wealth. They can, however, attempt to mitigate stratification in pupil performance through educational services and interventions provided for all learners. The question for schools, therefore, is whether learner performance is related to actions taken or resources available that are under the ambit of schools, and whether these tend to lessen the "SES effect."

To test these questions, three linear regression models were fitted for three separate outcome variables that corresponded with two subtasks on the EGRA: Oral Reading (local language) and Listening Comprehension (English). An attempt was made initially to identify background characteristics of learners and their households which lay largely outside the purview of schools and school systems and which could serve as "control" variables in subsequent models. Most learner background variables were initially tested for their association with the two selected outcome variables, but only variables that showed a statistically significant relationship with the outcome variables were retained in the final control models. ***Table 22*** presents the final linear regression control models, including beta coefficients and standard errors, for Oral Reading and Listening Comprehension.

Recall that, for each model, only statistically significant variables were retained; the models therefore do not include each variable listed in the table.[7] In the case of categorical variables, such as age categories (i.e. under-age, of-age, and over-age) and school type, the first category serves as the reference category. Thus, the models compare all other categories to the reference categories. In the case of school type, for example, learners who attended government (GRZ) schools, grant-aided schools, and private schools were each compared to learners who attended community schools in terms of their performance on the EGRA. In the table, coefficients can be read as changes in the outcome variables, positive or negative, that are associated with the variables entered in the models. Oral Reading was measured in words per minute and therefore coefficients denote increases or decreases in WPM associated with tested variables. Outcome

[7] Variables not included in the models are denoted by a "--".

variables in the Listening Comprehension model, however, were percentages with a potential range from 0.0 to 1.0; therefore, coefficients can be interpreted as increases or decreases in percentage points on this subtask.

Table 22. Variables associated with learner performance on the EGRA

VARIABLE	SUB-VARIABLE	ORAL READING FLUENCY (LOCAL LANGUAGE) COEFFICIENT	ORAL READING FLUENCY (LOCAL LANGUAGE) STANDARD ERROR	LISTENING COMPREHENSION (ENGLISH) COEFFICIENT	LISTENING COMPREHENSION (ENGLISH) STANDARD ERROR
Urban area		-0.55	1.58	0.07***	0.01
Language of instruction (LOI) = Local		2.53**	0.76	-0.08***	0.01
Has reader		3.89***	1.08	0.08**	0.02
Repeater		-1.74**	0.59	--	--
Age	Under-age			--	--
	Of-age	3.77**	1.13	--	--
	Over-age	4.16***	0.95	--	--
Preschool		--	--	0.08*	0.04
Read to others	Never				
	Sometimes	5.01***	0.57	0.05^{+}	0.03
	Every day	8.89***	1.64	0.05^{+}	0.02
Others read to you	Never	--	--		
	Sometimes	--	--	0.03**	0.01
	Every day	--	--	0.03	0.05
School type	Community	--	--		
	GRZ	--	--	0.04^{+}	0.02
	Grant-aided	--	--	0.06^{+}	0.03
	Private	--	--	0.17***	0.03
SES quintile	Lowest (Q1)	--	--		
	Q2	--	--	0.01	0.01
	Q3	--	--	0.03	0.02
	Q4	--	--	0.04^{+}	0.02
	Highest (Q5)	--	--	0.19***	0.03
Constant		-3.41**	1.17	-0.02	0.03
R^2		0.09		0.32	

Note: Statistical significance denoted by *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and $^{+}p < 0.1$.

The final model fitted for the Oral Reading subtask tested whether variables such as urban location, language of instruction matches that spoken at home, possession of a reader, repeating grades, learners' age, and reading to others at home were associated with learner performance as measured in WPM. Most variables were found to be statistically significant predictors of oral reading in learners' local languages. Being in a class in which the language of instruction matched that which learners spoke at home was associated with reading 2.5 WPM faster, on average. Having a reader in the learner's local language was associated with reading 3.9 WPM faster. As might be expected, learners who had repeated grade 2 were associated with a slower WPM rate; their older peers tended to read approximately 4 WPM faster, on average. Importantly, learners who read to others outside of school, either sometimes or every day, performed significantly better (5.0 and 8.9 WPM faster, respectively) on the Oral Reading subtask than did learners who reported never doing so. While most of the tested variables were found to be significant, it is important to note that this model explained only approximately 9.0% of the variation seen in the Oral Reading

subtask (R^2 = 0.09); other nontested factors may also be significant predictors of oral reading.

More tested variables were found to be significantly associated with performance on the Listening Comprehension subtask, which was conducted in English, and were included in the second model. Living in an urban location was associated with a seven percentage point increase in comprehension score, which is likely to be due in part to the increased frequency of English in urban locations. Unlike oral reading, learners attending schools in which the language of instruction matched their local language was disadvantageous, probably because very few learners (1.9%) spoke English at home. Thus, if the language of instruction matched learners' local language, it was likely to be a language other than English. Having possession of an English reader and preschool attendance were both associated with an eight percentage point increase in comprehension score, and being read to (three percentage points) as well as reading to others outside of school (five percentage points) were associated with increased performance on the Listening Comprehension subtask, as compared to peers who never read outside of school. Interestingly, these relationships held even controlling for SES and school type.

As compared to community schools, attending a government, grant-aided, or private school was associated with a four, six, and 17 percentage point increase in comprehension score. While Table 20 made it clear that very few learners in the lowest four SES quintiles were attending private schools, the private school "effect" on English listening comprehension held even after controlling for SES quintile. This may be because the language of instruction in private schools tends to be English. The largest effect on the Listening Comprehension subtask score was for these learners who came from the wealthiest households (19 percentage points), suggesting that SES was associated with performance on this subtask, but only for learners from the wealthiest households. This model explained approximately 32% of the variation observed in learner listening comprehension scores (R^2 = 0.32), substantially more than the other models.

With the exception of learners' possession of a reader, the variables tested in the models above lie outside the purview of schools and most education-based interventions, and therefore are good candidates for control variables. Possession of a reader, while certainly within the jurisdiction of educational systems, was included in the models of control variables due to the decision in Zambia to provide additional training to teachers rather than provide learners with textbooks which effectively (albeit temporarily) removed this variable from the Ministry's sphere of influence. These are the types of findings that will be of interest to policy makers and school leaders, as they highlight certain characteristics of learners, households, and schools that are associated with learner performance. In addition to these baseline control variables, an important question for the educational system in Zambia is: What actions could schools and teachers take that might influence learning outcomes? This section now turns to this question.

Several variables, collected via the complementary instruments and hypothesized to be influential on learner performance, were classified under five distinct categories, as follows: teacher actions (marking learners' work), teacher characteristics (teacher experience), time reading (independent reading at school), school resources (appropriate number of books, PTRs), and pedagogical oversight (frequency of coaching, total coaching hours). A linear regression model was fitted for each of the three outcome variables and each of the five variable categories, yielding 15 total regression models, in order to test whether certain teacher actions, teacher characteristics, time spent reading, school resources, or aspects of pedagogical oversight were significantly associated with learner performance on the two EGRA subtasks, after holding constant the control variables discussed in Table 22. Regression results are displayed in the tables that follow.

Recall from Table 22 that the baseline model with learner background controls explained approximately 9% of the variation observed in oral reading fluency. ***Table 23*** depicts how this baseline model changed with the addition of variables associated with teacher actions, teacher characteristics, time spent reading, school resources, and pedagogic oversight. The only model that significantly improved upon the baseline model's fit on oral reading was the time reading model, including whether learners read books on their own during school the day before the assessment. Independent reading at school was positively associated with oral reading; learners who reported independent reading the previous day read, on average, 4.5 WPM faster than learners who did not. Interestingly, the inclusion of this variable also reduced the coefficients of reading sometimes or every day to others outside of school by approximately 1 WPM (compare Table 22 and Table 23). Moreover, the coefficient of time reading in school was higher in this model than the coefficients for reading outside of school. This suggests that time allocated to independent reading during school (if pupils are given developmentally appropriate reading material) can somewhat make up for some learners' lack of reading outside of school. This model also significantly increased the proportion of the variation in oral reading explained from 9% to 13% (R^2 = 0.13).

Table 23. Variables associated with learner performance on the EGRA Oral Reading Fluency subtask

VARIABLE	SUB-VARIABLE	TEACHER ACTIONS		TEACHER CHARACTERISTICS		TIME READING		SCHOOL RESOURCES		PEDAGOGICAL OVERSIGHT	
		COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR
Urban area		-0.94	1.49	-0.35	1.41	-0.79	1.47	-0.53	1.25	0.20	1.92
LOI = Local		2.38*	0.75	2.49**	0.77	2.71***	0.70	2.23*	0.93	2.84***	0.64
Local reader		3.61*	1.09	4.02**	1.05	2.78**	0.99	3.02**	0.94	3.72**	1.16
Repeater		-1.74*	0.56	-1.66**	0.62	-1.66**	0.56	-1.51*	0.62	-1.73*	0.77
Age	Under-age										
	Of-age	3.88*	1.14	3.79*	1.17	3.25**	1.17	3.64**	1.09	4.17**	1.37
	Over-age	4.29**	1.04	4.32***	0.98	3.62**	1.02	4.06***	0.95	4.29***	1.22
Read to others	Never										
	Sometimes	4.80**	0.60	5.00***	0.59	3.99***	0.60	4.94***	0.51	4.85***	0.66
	Every day	8.20**	1.49	8.83***	1.69	7.62***	1.65	9.23***	1.83	8.70***	1.95
Teacher marking in exercise book	No pages										
	¼ of pages	-2.38	2.61	--	--	--	--	--	--	--	--
	½ of pages	-1.42	2.88	--	--	--	--	--	--	--	--
	¾ of pages	-0.70	3.05	--	--	--	--	--	--	--	--
	All pages	2.09	2.78	--	--	--	--	--	--	--	--
	Not available	-1.80	2.62	--	--	--	--	--	--	--	--
Teacher experience		--	--	0.01	0.07	--	--	--	--	--	--
School reading		--	--	--	--	4.46***	0.56	--	--	--	--
Started year with textbooks		--	--	--	--	--	--	-1.79	1.27	--	--
PTR		--	--	--	--	--	--	-0.01	0.01	--	--
Frequency of reading coaching	Never	--	--	--	--	--	--	--	--		
	Annually	--	--	--	--	--	--	--	--	0.55	0.79
	>Monthly	--	--	--	--	--	--	--	--	-0.74	0.98
	Monthly	--	--	--	--	--	--	--	--	1.05	1.41
Hours coached		--	--	--	--	--	--	--	--	0.00	0.07
Constant		-2.04	2.75	-3.61*	1.36	-3.59**	1.19	-2.22	1.42	-4.09**	1.44
R^2		0.10		0.09		0.13		0.10		0.09	

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Similar to Table 23, ***Table 24*** shows how the baseline model on Listening Comprehension evolved with the addition of variables associated with teacher actions, teacher characteristics, time spent reading, school resources, and pedagogical oversight. Recall from Table 22 that the baseline model explained approximately 32% of the variation observed in English Listening Comprehension scores. By this measure, none of the additional models built upon the base model; all R^2 values ranged between 0.31 and 0.33. The teacher action model revealed that feedback provided by teachers on learners' work (in the form of corrections in the reading exercise book) was associated with listening comprehension scores. Learners with greater proportions of pages with teacher marks tended to perform significantly better (between 6 and 12 percentage points) on this subtask. In addition, the time reading model showed that time spent reading in school was also marginally associated with listening comprehension scores, although the coefficient was small (i.e. two percentage point increase from those learners who did not read independently). While no variables in the pedagogical oversight model, which included the frequency of coaching visits and total hours teachers were coached, were found to be significant, the variables' inclusion did alter several other coefficients in the model. When controlling for the extent of reading coaching, coefficients for SES quintile 4 and 5 (compared to the lowest quintile), attending private school (compared to a community school), and reading to others outside of school increased. Holding constant the amount of pedagogical support teachers received, attending private school (as opposed to a community school) and coming from one of the wealthiest households (as opposed to one of the poorest) were associated with a 25 and 21 percentage point increase in English listening comprehension scores.

Table 24. Variables associated with learner performance on the EGRA English Listening Comprehension subtask

VARIABLE	SUB-VARIABLE	TEACHER ACTION		TEACHER CHARACTERISTICS		TIME READING		SCHOOL RESOURCES		PEDAGOGICAL OVERSIGHT	
		COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR
Urban		0.07***	0.01	0.07***	0.01	0.07***	0.01	0.07***	0.01	0.05*	0.02
LOI = Local		-0.08***	0.01	-0.08***	0.01	-0.08***	0.01	-0.08***	0.02	-0.09***	0.02
English reader		0.07**	0.02	0.08**	0.02	0.08**	0.02	0.07**	0.03	0.07*	0.03
School type	Community										
	GRZ	0.02	0.02	0.03	0.03	0.04+	0.02	0.04+	0.02	0.02	0.02
	Grant-aided	0.05	0.03	0.06	0.03	0.06+	0.03	0.07*	0.03	0.07	0.05
	Private	0.16***	0.03	0.18***	0.04	0.17***	0.03	0.19***	0.03	0.25***	0.04
Preschool		0.08*	0.03	0.08*	0.04	0.08*	0.04	0.08*	0.03	0.07*	0.03
Read to others	Never										
	Sometimes	0.04+	0.02	0.05+	0.03	0.05	0.03	0.06*	0.02	0.06**	0.02
	Every day	0.04	0.02	0.05+	0.03	0.04	0.02	0.04	0.03	0.10**	0.03
Others read to you	Never										
	Sometimes	0.03**	0.01	0.03*	0.01	0.03	0.01	0.03**	0.01	0.02	0.01
	Every day	0.03	0.05	0.03	0.05	0.03	0.05	0.04	0.04	-0.01	0.04
SES quintile	Lowest (Q1)										
	Q2	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.02
	Q3	0.03	0.02	0.03	0.02	0.02	0.02	0.02	0.01	0.02	0.01
	Q4	0.04+	0.02	0.04+	0.02	0.04+	0.02	0.04	0.03	0.06**	0.02
	Highest (Q5)	0.19***	0.04	0.20***	0.04	0.19***	0.03	0.19***	0.03	0.21***	0.04
Teacher marking in exercise book	No pages			--	--	--	--	--	--	--	--
	¼ of pages	0.06**	0.02	--	--	--	--	--	--	--	--
	½ of pages	0.06*	0.03	--	--	--	--	--	--	--	--
	¾ of pages	0.12*	0.05	--	--	--	--	--	--	--	--
	All pages	0.10**	0.03	--	--	--	--	--	--	--	--
	Not available	0.03	0.03	--	--	--	--	--	--	--	--
Teacher experience		--	--	-0.00	0.00	--	--	--	--	--	--
School reading		--	--	--	--	0.02+	0.01	--	--	--	--
Started year with textbooks		--	--	--	--	--	--	0.08	0.05	--	--
PTR		--	--	--	--	--	--	0.00	0.00	--	--

Table 24. Variables associated with learner performance on the EGRA English Listening Comprehension subtask

VARIABLE	SUB-VARIABLE	TEACHER ACTION		TEACHER CHARACTERISTICS		TIME READING		SCHOOL RESOURCES		PEDAGOGICAL OVERSIGHT	
		COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR	COEFFICIENT	STANDARD ERROR
Frequency of reading coaching	Never	--	--	--	--	--	--	--	--		
	Once	--	--	--	--	--	--	--	--	-0.02	0.03
	>Once / month	--	--	--	--	--	--	--	--	-0.03	0.03
	Once / month	--	--	--	--	--	--	--	--	-0.03	0.03
Hours coached		--	--	--	--	--	--	--	--	0.00	0.00
Constant		-0.07^{+}	0.04	-0.16	0.03	-0.02	0.03	-0.04	0.04	0.04	0.03
R^2		0.33		0.32		0.32		0.33		0.31	

*p < 0.05, **p < 0.01, ***p < 0.001, ^{+}p < 0.1.

Overall, these models demonstrated that a number of learner background variables were associated with reading outcomes, as measured by the EGRA. One positive result in this regard was that relative household wealth and private school attendance were not ubiquitously significant predictors of learner performance; these variables were significant only for Listening Comprehension scores but not for Oral Reading Fluency scores. Interestingly, the private school advantage appeared to remain even after controlling for learners' SES quintile. At the same time, these findings suggest that there are also a number of actions that policy makers, school leaders, and teachers can take that are associated with learner performance. At the policy level, focusing on the timely provision of textbooks appears to be important, while attempting to reduce PTRs does not. At the school and classroom levels, allowing time for independent reading during school hours (with developmentally appropriate materials and instructional support) and providing feedback to learners on their classwork arose as important instructional attributes, while teachers' years of experience did not. Instructional coaching did not significantly predict learner performance on these two subtasks, but this could be because of how it was measured (in the frequency of visits and total hours coached). Perhaps it would be important to gather data on the type of coaching received, the quality of interactions between coaches and teachers, teachers' perceptions of the benefits of coaching, and whether teachers tend to put into practice what they have learned from coaches.

5 CONCLUSIONS AND RECOMMENDATIONS

5.1 RECOMMENDATIONS FROM THE GRADE 2 NAS

The results of the grade 2 NAS in Zambia strongly suggest that the teaching of reading in at least some languages is focused on memorization of word reading, without adequate attention placed on learning how to decode words or on comprehension strategies, whether of oral or written language. While developing automaticity with reading familiar words is a necessary skill in learning to read, it is not a sufficient one, and it appears that across languages pupils are not well prepared to move from individual word reading to more fluent reading of connected text.

Although the results of the reading assessment are clear—children are not performing at the level that we would like them to in either foundational pre-literacy skills or actual reading of connected text with meaning—the explanation may lie as much in *how* they are taught as in *what* they are taught. The results of this study suggest that pupils are memorizing what they learn, rather than engaging with understanding. Teachers need to be supported in adopting pedagogically more effective approaches for teaching reading in the early grades.

5.2 RECOMMENDATIONS FROM THE COMPLEMENTARY INSTRUMENTS

The data obtained by the complementary instruments for this study were used to paint a portrait of teachers, head teachers, and schools in Zambia. In terms of teacher and head teacher characteristics, both of these workforces were found to be rather inexperienced, and teachers had not received a significant amount of in-service training. Instructionally specific training was also relatively uncommon, as most teachers received visits from reading coaches at most semiannually, and coaching resources (time, visits) were not equitably distributed between teachers or schools. Ratios of learners to teachers appeared to be rather low for most schools (though some were quite high), but many schools and learners lacked access to reading and mathematics textbooks.

In terms of teaching practice, teachers surveyed for this report appear to have been providing feedback, both written and verbal, to learners in their classrooms. Learners reported a high frequency of praise for good academic performance and

most teachers took time to mark learners' classwork. Although direct observations were not conducted for this study, learners reported that their teachers frequently employed several pedagogical practices that were appropriate for learning letters, learning words, and practicing reading in the early grades. However, teachers did not also employ active and constructive feedback when learners' responded incorrectly to a question; slightly more learners reported passive or destructive teacher responses (e.g., hitting, calling on someone else).

Several background characteristics of learners were found to be statistically significant predictors of learner performance on the reading assessment when regression analyses were fitted onto pupil outcomes on the Oral Reading and Listening Comprehension subtasks. Positively, household wealth and private school attendance were not ubiquitously significant predictors of learner performance, and there appear to be a number of actions that policy makers, school leaders, and teachers can take that are associated with learner performance. For example, focusing on the timely provision of textbooks, allowing time for independent reading during school hours (with developmentally appropriate materials and instructional support), and providing feedback to learners on their classwork appear to have been important. On the other hand, attempting to reduce PTRs or increase teachers' experience did not. Instructional coaching did not significantly predict learner performance on these two subtasks, but this may have been because of how it was measured (in the frequency of visits and total hours coached). Future studies could gather data on the type of coaching received, the quality of interactions between coaches and teachers, teachers' perceptions of the benefits of coaching, and whether teachers tend to put into practice what they have learned from coaches.

5.3 RECOMMENDATIONS FROM THE EXERCISE TO CALCULATE AND COMPARE "BASELINE" AND MIDTERM DATA

As part of this analysis report, RTI was requested to produce a statistical comparison of the 2014 grade 2 NAS results to those from several previous early grade reading evaluations carried out in Zambia. ***Annexes B and C*** explain the equating process used in responding to this requirement, the inherent limitations of the exercise, the results of the comparisons that were feasible based on the existing data sets.

To summarize, RTI was able to approximate a comparison of the 2011/2012 assessments with the 2014 EGRA. As anticipated, however, significant differences in performance generally were not detectable, given that—among other differences in the samples—the 2014 EGRA included private schools.

5.4 POLICY DIALOGUE WORKSHOP AND RECOMMENDATIONS

On July 1—3, 2015, the MESVTEE and ECZ hosted a policy dialogue workshop in Lusaka. The purpose of this workshop was to review the findings of the Grade 2 NAS, examine the implications arising from those findings, make recommendations for this report, and set benchmarks and targets for reading in Zambia.

The workshop was opened and addressed by the Directors of ECZ and Standards and Curriculum and the USAID Education Director.

During the first day of the workshop, the results of the study were presented to an audience of approximately 125 people representing the following a wide range of organizations, among others:

Ministry of Education, Science, Vocational Training, and Early Education

Examination Council of Zambia

Curriculum Development Centre

Zambia National Education Coalition

Provincial and District Education Officers

Local and international nongovernmental organizations working in the field of early grade education

After the presentation of the results on day one, the attention of the workshop switched to generating recommendations for this report. The second and third days of the workshop were dedicated to setting benchmarks for student performance.

In creating the recommendations that follow, the participants were organized into groups each group made recommendations on each of the key issues emerging from the report. The recommendations from each of the groups were presented in plenary. The emerging themes of those recommendations are discussed in the remainder of this section below. [8]

5.4.1 TEACHERS

The workshop participants recommended that experienced, well qualified teachers should be assigned to the lower grades in order to establish a solid foundational

[8] Following the National event in Lusaka, five additional one-day dissemination events were held regionally in Livingstone, Solwezi, Chipata, Mongu, and Kasama. These were attended by a total of 509 people. Recommendations from these events are included in those presented in this report.

understanding of reading. These teachers should have specialized expertise in teaching reading. The participants also recommended that in order to attract quality teachers to early-grade instruction, incentives and salary structures should be aligned to teachers' qualifications. Furthermore, because students need to be assessed regularly teachers should receive training in how to use assessment results to improve teaching methods and lesson content. In the same way that students need to be assessed, it was recommended that both the internal and external monitoring of teachers should be intensified

5.4.2 TEACHER TRAINING

The workshop participants recommended that college courses in education should train teachers on research based approaches to the teaching of reading. This could include producing a clear set of teaching strategies and lesson routines for teaching reading that teachers can use in their classrooms. The participants also noted that when newly qualified teachers start their careers, schools should provide them teachers with experienced teachers as mentors and also provide them with support from the school-leadership. Furthermore, it was also recommended that a national Continuing Professional Development (CPD) program should be established at all levels of the education system to incentivize best-practices and continued growth in teaching approaches and pedagogies. Finally, both internal and external teacher monitoring should be intensified

5.4.3 MATERIALS

The workshop participants recommended that schools and teachers should be provided with quality learning materials in a timely manner so that the materials may be of use, and that there should be an adequate supply of materials so that learners may benefit from them. It was felt that decentralizing the procurement of teaching and learning materials could help with their timely and sufficient distribution. It was felt that including the materials for all of the terms into a single booklet would also streamline lesson-planning and pupil learning. Finally, it was recommended that teachers be trained and encouraged to develop and use their own stories and other learning aids to enhance student learning.

5.4.4 PRACTICE AND THE ROLE OF PARENTS

The workshop participants recommended that parents should be encouraged to take an interest in the education of their children. In particular it was recommended that parents should be made aware of the importance of parents reading to their children at home and of asking their children to read to them. In addition, it was felt that communities as a whole could be encourages to participate in children's learning by supporting an increase in reading time at

school and at home. Finally, it was recommended that schools should assign homework regularly, so that parents would expect their children to have homework in which they could show and interest and with which they could help their children.

5.5 BENCHMARKS AND TARGETS

As part of the policy dialogue workshop hosted by the MESVTEE and ECZ on July 1—3, 2015, participants set benchmarks and targets for reading.

The benchmarks reported in ***Table 25*** were developed based on the results of Zambian students in the 2014 Grade 2 NAS and informed by a range of international benchmarks, the participants' experience with and knowledge of the Zambian context, and technical support provided by the researchers who led the Grade 2 NAS.

The 5-year targets reported in Table 25 were set by the participants, with support from the technical experts based on the assumptions that the recommendations recorded in this report are implemented and that a concerted effort by all role players is exerted with respect to the implementation of these recommendations. The targets attempt to temper the desire for dramatic improvements with the realization that change in education is generally slow and that the recommendations made in this report—in particular, regarding the development of teaching and learning approaches and the associated teacher training—all take time.

It should also be noted that the participants in the benchmarking workshop made two key decisions (that are reflected in Table 25). First it was agreed that instead of developing benchmarks and targets for the individual languages, a single set of benchmarks would be developed for all languages. This decision was informed by the fact that the performance across languages in the Grade 2 NAS was more similar than not. Second, it was agreed to develop indicators for what are referred to in the table as "emergent readers" because the number of "readers" at baseline is so low that to see a significant shift in only five years is unrealistic. By contrast, a noticeable shift in the proportion of "emergent readers" is more reasonable in the same timeframe.

Table 25. National benchmarks and targets for reading in Zambia.

		READING		
		CORRECT NON-WORDS PER MINUTE	CORRECT WORDS PER MINUTE (ORF)	COMPREHENSION
BENCHMARKS	EMERGENT READERS	15	20	40%
	READERS	30	45	80%
ZERO SCORE	BASELINE (2014 STUDY DATA)	68%	65%	80%
	PROPOSED 5 YEAR TARGET	27%	26%	32%
EMERGENT READERS	BASELINE (2014 STUDY DATA)	12%	11%	7%
	PROPOSED 5 YEAR TARGET	36%	33%	21%
READERS	BASELINE (2014 STUDY DATA)	2%	1%	2%
	PROPOSED 5 YEAR TARGET	8%	4%	8%

REFERENCES

Carasco, J. F., Munene, J. C., Kasente, D. H., & Odada, M. (1996). *Factors influencing effectiveness in primary schools: A baseline study.* Report for the Uganda Improving Educational Quality (IEQ) Project - phase 1 research. Kampala: Uganda National Examinations Board. Retrieved from http://pdf.usaid.gov/pdf_docs/pnabz292.pdf

Fuchs, L., Fuchs, D., Hosp, M. K., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*(3), 239–256. http://dx.doi.org/10.1207/s1532799xssr0503_3

Gove, A., & Cvelich, P. (2011). *Early reading: Igniting education for all. A report by the Early Grade Learning Community of Practice.* Revised edition. Research Triangle Park, NC: Research Triangle Institute. Retrieved from http://www.rti.org/pubs/early-reading-report-revised.pdf

Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: a review and synthesis. Psychological bulletin, 128(5), 774–795.

Heneveld, W., & Craig, H. (1996). *Schools count: World Bank project designs and the quality of African primary education.* Washington, DC: World Bank. Retrieved from https://www.openknowledge.worldbank.org/handle/10986/9974

National Institute of Child Health and Human Development [US]. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* National Institutes of Health Publication No. 00-4769. Washington, DC: U.S. Government Printing Office.

Perfetti, C.A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Erlbaum.

RTI International & International Rescue Committee. (2011). *Guidance notes for planning and implementing EGRA.* Washington, DC: RTI and IRC. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=318

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

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360–407. http://dx.doi.org/10.1598/rrq.21.4.1

World Bank: Independent Evaluation Group. (2006). *From schooling access to learning outcomes—An unfinished agenda: An evaluation of World Bank support to primary education.* Washington, DC: World Bank.

ANNEX A: EGRA RESULTS BY LOCAL LANGUAGE SUBTASK, INCLUDING STANDARD ERROR VALUES AND 95% CONFIDENCE INTERVALS

Table A-1. Listening Comprehension

	NO. CORRECT RESPONSES / MINUTE			% CORRECT / ATTEMP-TED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga (*n*=680)	2.49 (2.15, 2.82)	2.42 (2.02, 2.82)	2.55 (2.26, 2.84)	49.7% (43%, 56%)	48.4% (40.4%, 56.4%)	51% (45.3%, 56.8%)	49.7% (43%, 56.5%)	48.4% (40.4%, 56.4%)	51% (45.3%, 56.8%)	15.4% (7.2%, 23.5%)	16.5% (7.2%, 25.9%)	14.2% (6.4%, 22%)
Standard error	0.15	0.17	0.12	2.9%	3.5%	2.5%	2.9%	3.5%	2.5%	3.5%	4.1%	3.4%
Cinyanja (*n*=1007)	3.27 (2.84, 3.7)	3.28 (2.91, 3.65)	3.25 (2.75, 3.75)	65.4% (57%, 74%)	65.6% (58.3%, 73%)	65% (55%, 75%)	65.4% (56.8%, 74%)	65.6% (58.3%, 73%)	65% (55%, 75%)	1.8% (-0.6%, 4.2%)	1.9% (-0.4%, 4.2%)	1.6% (-1%, 4.2%)
Standard error	0.19	0.16	0.22	3.8%	3.2%	4.4%	3.8%	3.2%	4.4%	#N/A	1.0%	1.1%
Icibemba (*n*=2176)	2.04 (1.8, 2.28)	2.16 (1.9, 2.43)	1.93 (1.68, 2.17)	0.41	43.3% (38%, 48.6%)	38.5% (33.6%, 43.4%)	40.8% (36%, 45.7%)	43.3% (38%, 48.6%)	38.5% (33.6%, 43.4%)	16.8% (12.3%, 21.3%)	14% (9.4%, 18.6%)	19.5% (14.3%, 24.8%)
Standard error	0.12	0.13	0.12	2.3%	2.5%	2.4%	2.3%	2.5%	2.4%	2.2%	2.2%	2.5%
Kiikaonde (*n*=200)	2.65 (0.85, 4.45)	2.78 (1.18, 4.39)	2.51 (0.5, 4.53)	53% (17%, 89%)	55.6% (23.5%, 87.7%)	50.3% (10%, 90.5%)	53% (16.9%, 89%)	55.6% (23.5%, 87.7%)	50.3% (10%, 90.5%)	9.9% (-21.8%, 41.7%)	11.1% (-27.5%, 49.7%)	8.7% (-16.3%, 33.8%)
Standard error	0.42	0.37	0.47	8.4%	7.5%	9.4%	8.4%	7.5%	9.4%	7.4%	9.0%	5.8%
Lunda (*n*=131)	3.82 (3.75, 3.89)	3.9 (3.55, 4.24)	3.76 (3.6, 3.91)	76.5% (75%, 78%)	48.4% (40.4%, 56.4%)	51% (45.3%, 56.8%)	76.5% (75%, 77.9%)	77.9% (71%, 84.8%)	75.2% (72.1%, 78.3%)	1.3% (0.3%, 2.3%)	0.2% (0%, 0.3%)	2.3% (0.8%, 3.8%)
Standard error	0.02	0.08	0.04	0.3%	1.6%	0.7%	0.3%	1.6%	0.7%	0.2%	0.0%	0.3%

Table A-1. Listening Comprehension

	NO. CORRECT RESPONSES / MINUTE			% CORRECT / ATTEMP-TED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Luvale (n=254)	4.27 (4.04, 4.5)	4.36 (4.07, 4.65)	4.18 (3.99, 4.37)	85.4% (81%, 90%)	48.4% (40.4%, 56.4%)	51% (45.3%, 56.8%)	85.4% (80.7%, 90.1%)	87.3% (81.5%, 93.1%)	83.6% (79.9%, 87.4%)	0.1% (-0.4%, 0.7%)	0% (0%, 0%)	0.3% (-0.7%, 1.3%)
Standard error	0.07	0.09	0.06	1.5%	1.8%	1.2%	1.5%	1.8%	1.2%	0.2%	0.0%	0.3%
Silozi (n=400)	3.4 (2.47, 4.32)	3.32 (2.6, 4.03)	3.48 (2.35, 4.61)	0.68	48.4% (40.4%, 56.4%)	51% (45.3%, 56.8%)	68% (49.5%, 86.4%)	66.3% (52.1%, 80.6%)	69.5% (46.9%, 92.1%)	2.1% (-2.9%, 7.1%)	2% (-1.9%, 6%)	2.2% (-4%, 8.3%)
Standard error	0.29	0.22	0.36	5.8%	4.5%	7.1%	5.8%	4.5%	7.1%	1.6%	1.2%	1.9%

Table A-2. Letter Sound Identification

	NO. CORRECT LETTER SOUNDS / MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga (*n*=680)	3.68 (2.9, 4.46)	3.32 (2.55, 4.1)	4.04 (3.19, 4.89)	13.6% (10%, 17%)	13% (8.6%, 17.4%)	14.1% (11%, 17.2%)	3.7% (2.9%, 4.5%)	3.3% (2.5%, 4.1%)	4% (3.2%, 4.9%)	56.8% (46.7%, 66.8%)	57.4% (45%, 69.8%)	56.2% (48.2%, 64.1%)
Standard error	0.34	0.34	0.37	1.6%	1.9%	1.3%	0.3%	0.3%	0.4%	4.4%	5.4%	3.4%
Cinyanja (*n*=1005)	7.21 (5.01, 9.4)	7.35 (5.24, 9.47)	7 (4.3, 9.7)	27.6% (15%, 40%)	27.5% (16.9%, 38.2%)	27.5% (12.1%, 42.9%)	7.4% (5.3%, 9.4%)	7.4% (5.3%, 9.5%)	7.3% (5%, 9.5%)	47.2% (22.1%, 72.2%)	47.9% (27.1%, 68.6%)	46.6% (16.9%, 76.2%)
Standard error	0.97	0.94	1.19	5.6%	4.7%	6.8%	0.9%	0.9%	1.0%	#N/A	9.2%	13.1%
Icibemba (*n*=2174)	8.77 (7.04, 10.51)	9.02 (7.2, 10.85)	8.53 (6.66, 10.4)	0.34	34.4% (29.2%, 39.7%)	32.8% (27.2%, 38.3%)	8.8% (7%, 10.5%)	9% (7.2%, 10.9%)	8.6% (6.7%, 10.4%)	35.1% (27.8%, 42.3%)	33.8% (26.2%, 41.4%)	36.3% (28.7%, 43.8%)
Standard error	0.83	0.87	0.89	2.5%	2.5%	2.7%	0.8%	0.9%	0.9%	3.5%	3.6%	3.6%
Kiikaonde (*n*=200)	5.44 (-5.76, 16.64)	5.84 (-6.56, 18.25)	5.02 (-6.33, 16.38)	19.1% (-18%, 57%)	20.8% (-23.8%, 65.5%)	17.3% (-15.8%, 50.4%)	5.4% (-5.8%, 16.6%)	5.8% (-6.6%, 18.2%)	5% (-6.3%, 16.4%)	57% (7.4%, 106.7%)	60.7% (3.1%, 118.4%)	53.2% (5.5%, 100.9%)
Standard error	2.60	2.88	2.64	8.7%	10.4%	7.7%	2.6%	2.9%	2.6%	11.5%	13.4%	11.1%
Lunda (*n*=131)	9.63 (-1.08, 20.34)	12.57 (-0.87, 26.01)	7.06 (-0.65, 14.77)	37.7% (15%, 61%)	13% (8.6%, 17.4%)	14.1% (11%, 17.2%)	9.6% (-1.1%, 20.3%)	12.6% (-0.9%, 26%)	7.1% (-0.7%, 14.8%)	28.5% (7.3%, 49.7%)	19.8% (2.6%, 37.1%)	36.1% (13.1%, 59%)

Table A-2. Letter Sound Identification

	NO. CORRECT LETTER SOUNDS / MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Standard error	2.49	3.12	1.79	5.3%	6.9%	3.5%	2.5%	3.1%	1.8%	4.9%	4.0%	5.3%
Luvale (*n*=254)	6.75 (4.42, 9.08)	6.18 (4.02, 8.33)	7.28 (4.64, 9.92)	24% (17%, 31%)	13% (8.6%, 17.4%)	14.1% (11%, 17.2%)	6.7% (4.4%, 9.1%)	6.2% (4%, 8.3%)	7.3% (4.6%, 9.9%)	42.1% (30.1%, 54.1%)	42.9% (31.2%, 54.5%)	41.4% (28.4%, 54.3%)
Standard error	0.73	0.68	0.83	2.1%	2.2%	2.1%	0.7%	0.7%	0.8%	3.8%	3.7%	4.1%
Silozi (*n*=400)	8.09 (-1.26, 17.44)	7.4 (-1.87, 16.66)	8.75 (-0.7, 18.21)	0.29	13% (8.6%, 17.4%)	14.1% (11%, 17.2%)	8% (-1.5%, 17.4%)	7.3% (-2%, 16.7%)	8.6% (-1%, 18.1%)	44.2% (9.6%, 78.7%)	46.5% (7.5%, 85.4%)	41.9% (11.2%, 72.6%)
Standard error	2.94	2.91	2.97	9.6%	10.1%	9.1%	3%	2.9%	3.0%	10.9%	12.2%	9.7%

Table A-3. Nonword Reading

	NO. CORRECT NONWORDS / MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga (*n*=680)	1.66 (0.85, 2.48)	1.56 (0.92, 2.21)	1.76 (0.66, 2.87)	7.8% (4%, 12%)	7.3% (4.5%, 10.2%)	8.3% (3%, 13.5%)	3.3% (1.7%, 5%)	3.1% (1.8%, 4.4%)	3.5% (1.3%, 5.7%)	85.4% (80.3%, 90.5%)	86.1% (82.4%, 89.8%)	84.6% (76.7%, 92.5%)
Standard error	0.35	0.28	0.48	1.7%	1.2%	2.3%	0.7%	0.6%	1.0%	2.2%	1.6%	3.4%
Cinyanja (*n*=1007)	3.2 (2.32, 4.07)	3.59 (2.64, 4.55)	2.75 (1.98, 3.53)	18.8% (14%, 23%)	21.4% (16.6%, 26.2%)	16% (11.3%, 20.7%)	6.4% (4.6%, 8.1%)	7.2% (5.3%, 9.1%)	5.5% (3.9%, 7%)	71.7% (63.7%, 79.7%)	69.8% (62.8%, 76.9%)	73.7% (64.1%, 83.4%)
Standard error	0.39	0.42	0.34	2.1%	2.1%	2.1%	0.8%	0.8%	0.7%	#N/A	3.1%	4.3%
Icibemba (*n*=2174)	5.6 (4.44, 6.77)	5.85 (4.66, 7.04)	5.37 (4.02, 6.71)	0.29	30.4% (24.7%, 36.1%)	27.8% (21.1%, 34.4%)	11.3% (8.9%, 13.6%)	11.8% (9.4%, 14.2%)	10.7% (8.1%, 13.4%)	61.8% (54.9%, 68.8%)	59.9% (53.3%, 66.6%)	63.7% (55.7%, 71.7%)
Standard error	0.56	0.57	0.64	2.8%	2.7%	3.2%	1.1%	1.1%	1.3%	3.3%	3.2%	3.8%
Kiikaonde (*n*=200)	2.27 (-2.88, 7.41)	2.29 (-3.01, 7.58)	2.24 (-2.79, 7.28)	13.7% (-16%, 44%)	14.2% (-16.2%, 44.6%)	13.1% (-16.8%, 43.1%)	4.5% (-5.8%, 14.8%)	4.6% (-6%, 15.2%)	4.5% (-5.6%, 14.6%)	76.2% (24%, 128.4%)	73.3% (12.5%, 134.1%)	79.2% (33.2%, 125.1%)
Standard error	1.20	1.23	1.17	7%	7.1%	7.0%	2.4%	2.5%	2.3%	12.1%	14.1%	10.7%
Lunda (*n*=131)	6.33 (3.32, 9.34)	9.67 (4.88, 14.46)	3.41 (2.64, 4.17)	34.5% (17%, 52%)	7.3% (4.5%, 10.2%)	8.3% (3%, 13.5%)	12.7% (6.6%, 18.7%)	19.3% (9.8%, 28.9%)	6.8% (5.3%, 8.3%)	49.7% (26.2%, 73.2%)	36.2% (4.7%, 67.8%)	61.5% (47.8%, 75.2%)
Standard error	0.70	1.11	0.18	4.1%	5.3%	2.2%	1.4%	2.2%	0.4%	5.5%	7.3%	3.2%

Table A-3. Nonword Reading

	NO. CORRECT NONWORDS / MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Luvale (*n*=254)	4.28 (3.68, 4.88)	4.25 (3.43, 5.06)	4.31 (3.65, 4.96)	17.9% (17%, 19%)	7.3% (4.5%, 10.2%)	8.3% (3%, 13.5%)	8.6% (7.4%, 9.8%)	8.5% (6.9%, 10.1%)	8.6% (7.3%, 9.9%)	72.9% (69.7%, 76%)	69.9% (66.9%, 72.9%)	75.7% (72.2%, 79.1%)
Standard error	0.19	0.26	0.21	0.4%	0.7%	0.7%	0.4%	0.5%	0.4%	1%	0.9%	1.1%
Silozi (*n*=400)	5.65 (3.08, 8.22)	5.1 (1.89, 8.32)	6.17 (3.75, 8.59)	0.32	7.3% (4.5%, 10.2%)	8.3% (3%, 13.5%)	11.2% (6%, 16.4%)	10.1% (3.6%, 16.7%)	12.2% (7.5%, 16.9%)	56% (28.1%, 83.9%)	59.1% (23.8%, 94.3%)	53.1% (32%, 74.2%)
Standard error	0.81	1.01	0.76	7.1%	8.3%	5.9%	1.6%	2.1%	1.5%	8.8%	11.1%	6.6%

Table A-4. Oral Reading Fluency

	NO. CORRECT WORDS / MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga (*n*=680)	1.85 (0.64, 3.06)	1.93 (0.81, 3.05)	1.77 (0.36, 3.18)	7.8% (3%, 13%)	7.4% (3.8%, 11%)	8.3% (1.8%, 14.8%)	4.3% (1.6%, 7.1%)	4.5% (2%, 7%)	4.2% (0.9%, 7.6%)	88.2% (82.2%, 94.1%)	88.3% (83.3%, 93.4%)	88% (79.9%, 96.1%)
Standard error	0.52	0.49	0.61	2%	1.6%	2.8%	1.2%	1.1%	1.5%	2.6%	2.2%	3.5%
Cinyanja (*n*=1007)	4.11 (2.94, 5.27)	4.35 (3.25, 5.46)	3.8 (2.61, 5)	21.2% (16%, 26%)	22.7% (18.2%, 27.3%)	19.5% (13.5%, 25.4%)	10.2% (7.3%, 13.1%)	10.9% (8.1%, 13.6%)	9.4% (6.4%, 12.3%)	64.2% (54.5%, 73.9%)	64.8% (59.9%, 69.7%)	63.8% (48.7%, 78.9%)
Standard error	0.52	0.49	0.53	2.3%	2.0%	2.6%	1.3%	1.2%	1.3%	#N/A	2.2%	6.7%
Icibemba (*n*=2169)	7.62 (5.98, 9.26)	7.85 (6.1, 9.6)	7.4 (5.51, 9.29)	0.34	35.8% (29.1%, 42.4%)	33% (26.2%, 39.8%)	16.4% (13.2%, 19.6%)	17% (13.4%, 20.6%)	15.8% (12.1%, 19.6%)	52% (45.5%, 58.5%)	49% (41.8%, 56.1%)	54.8% (46.7%, 62.9%)
Standard error	0.78	0.84	0.90	2.9%	3.2%	3.2%	1.5%	1.7%	1.8%	3.1%	3.4%	3.9%
Kiikaonde (*n*=200)	3.46 (-4.44, 11.36)	2.75 (-4.07, 9.58)	4.19 (-5.16, 13.53)	18.3% (-21%, 58%)	16.3% (-18.5%, 51.1%)	20.3% (-24.3%, 64.8%)	6.2% (-7.9%, 20.3%)	4.9% (-7.3%, 17.1%)	7.5% (-9.2%, 24.2%)	73.5% (16.3%, 130.8%)	73.5% (17.2%, 129.9%)	73.5% (15.4%, 131.7%)
Standard error	1.84	1.59	2.17	9.2%	8.1%	10.4%	3.3%	2.8%	3.9%	13.3%	13.1%	13.5%
Lunda (*n*=130)	6.28 (4.49, 8.06)	9.25 (6.48, 12.01)	3.74 (3.44, 4.05)	35.4% (19%, 52%)	7.4% (3.8%, 11%)	8.3% (1.8%, 14.8%)	14.5% (11.4%, 17.7%)	21.9% (17.5%, 26.2%)	8.1% (7.5%, 8.8%)	45.8% (25.2%, 66.4%)	33.6% (4.3%, 62.8%)	56.5% (46%, 67.1%)
Standard error	0.41	0.64	0.07	3.9%	5.5%	1.7%	0.7%	1.0%	0.2%	4.8%	6.8%	2.4%
Luvale (*n*=254)	4.32 (3.35, 5.3)	4.26 (2.76, 5.76)	4.38 (3.15, 5.61)	16.3% (15%, 18%)	7.4% (3.8%, 11%)	8.3% (1.8%, 14.8%)	10.1% (7.8%, 12.3%)	9.9% (6.4%, 13.4%)	10.2% (7.3%, 13%)	80.2% (77.3%, 83.1%)	81.9% (74.7%, 89.1%)	78.7% (75.7%, 81.6%)
Standard error	0.31	0.47	0.39	0.5%	1.7%	1.0%	0.7%	1.1%	0.9%	0.9%	2.3%	0.9%

Table A-4. Oral Reading Fluency

	NO. CORRECT WORDS / MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Silozi (*n*=400)	8.4 (3.86, 12.94)	7.96 (1.83, 14.08)	8.83 (5.2, 12.45)	0.37	7.4% (3.8%, 11%)	8.3% (1.8%, 14.8%)	25.1% (10.6%, 39.7%)	23.6% (5.1%, 42.2%)	26.5% (14.5%, 38.6%)	56.3% (29.9%, 82.8%)	55.8% (25.1%, 86.5%)	56.8% (33.6%, 80%)
Standard error	1.43	1.92	1.14	7.5%	8.7%	6.6%	4.6%	5.8%	3.8%	8.3%	9.7%	7.3%

Table A-5. Orientation to Print

	NO. CORRECT RESPONSES / MINUTE			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga (*n*=680)	2.41 (2.23, 2.59)	2.42 (2.26, 2.58)	2.4 (2.18, 2.61)	80.3% (74%, 86%)	80.6% (75.3%, 85.9%)	79.9% (72.8%, 87.1%)	7.3% (4.7%, 9.9%)	8.1% (4.7%, 11.5%)	6.5% (3.4%, 9.5%)
Standard error	0.08	0.07	0.09	2.6%	2.3%	3.1%	1.1%	1.5%	1.3%
Cinyanja (*n*=1007)	2.59 (2.14, 3.04)	2.63 (2.23, 3.03)	2.55 (2.04, 3.06)	86.4% (71%, 101%)	87.7% (74.3%, 101.1%)	84.9% (68.1%, 101.8%)	8.3% (-0.6%, 17.1%)	8.1% (0.3%, 15.9%)	8.5% (-1.8%, 18.9%)
Standard error	0.20	0.18	0.22	6.7%	5.9%	7.5%	Not available	3.5%	4.6%
Icibemba (*n*=2174)	2.5 (2.42, 2.58)	2.56 (2.47, 2.65)	2.45 (2.33, 2.57)	0.83	85.2% (82.3%, 88.2%)	81.7% (77.7%, 85.8%)	9.4% (7.3%, 11.4%)	8.1% (5.9%, 10.3%)	10.6% (7.2%, 13.9%)
Standard error	0.04	0.04	0.06	1.3%	1.4%	1.9%	1%	1.1%	1.6%
Kiikaonde (*n*=200)	2.25 (1.67, 2.83)	2.29 (1.51, 3.07)	2.21 (1.83, 2.6)	75% (56%, 94%)	76.3% (50.3%, 102.3%)	73.7% (60.9%, 86.5%)	9.7% (-14.7%, 34.2%)	10.3% (-24.7%, 45.3%)	9.1% (-9.9%, 28.1%)
Standard error	0.13	0.18	0.09	4.5%	6.0%	3.0%	5.7%	8.1%	4.4%
Lunda (*n*=131)	2.94 (2.85, 3.03)	3 (3, 3)	2.89 (2.71, 3.08)	98.1% (95%, 101%)	80.6% (75.3%, 85.9%)	79.9% (72.8%, 87.1%)	1.3% (0.3%, 2.3%)	0% (0%, 0%)	2.4% (0.9%, 4%)
Standard error	0.02	0.00	0.04	0.7%	0.0%	1.4%	0.2%	0.0%	0.4%
Luvale (*n*=254)	2.55 (2.53, 2.57)	2.68 (2.55, 2.8)	2.43 (2.33, 2.52)	84.9% (84%, 86%)	80.6% (75.3%, 85.9%)	79.9% (72.8%, 87.1%)	7.9% (6.8%, 9%)	4% (1.2%, 6.8%)	11.6% (6.9%, 16.2%)
Standard error	0.01	0.04	0.03	0.2%	1.3%	1.0%	0.4%	0.9%	1.5%
Silozi (*n*=400)	2.34 (1.75, 2.92)	2.34 (1.8, 2.88)	2.33 (1.69, 2.98)	0.78	80.6% (75.3%, 85.9%)	79.9% (72.8%, 87.1%)	12.9% (2.3%, 23.5%)	11.9% (1.7%, 22.1%)	13.8% (2.4%, 25.2%)
Standard error	0.18	0.17	0.20	6.1%	5.7%	6.8%	3.3%	3.2%	3.6%

Table A-6. Reading Comprehension

	NO. CORRECT RESPONSES / MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Chitonga (*n*=680)	0.13 (0.03, 0.23)	0.14 (0.05, 0.23)	0.12 (0, 0.24)	4.1% (2%, 6%)	4.7% (2.4%, 7%)	3.4% (0.7%, 6.2%)	2.6% (0.6%, 4.5%)	2.8% (1%, 4.5%)	2.4% (0%, 4.8%)	93.6% (89.3%, 98%)	93.3% (90.1%, 96.6%)	93.9% (88.4%, 99.5%)
Standard error	0.04	0.04	0.05	1%	1.0%	1.2%	0.8%	0.8%	1.0%	1.9%	1.4%	2.4%
Cinyanja (*n*=1009)	0.36 (0.28, 0.44)	0.39 (0.32, 0.46)	0.33 (0.24, 0.41)	10.7% (9%, 13%)	10.9% (9%, 12.9%)	10.3% (7.7%, 12.9%)	7.2% (5.7%, 8.8%)	7.8% (6.4%, 9.2%)	6.5% (4.8%, 8.3%)	78.7% (76%, 81.4%)	76.6% (74.4%, 78.7%)	81% (78.1%, 84%)
Standard error	0.03	0.03	0.04	1%	0.9%	1.1%	0.7%	0.6%	0.8%	Not available	1.0%	1.3%
Icibemba (*n*=2181)	0.33 (0.23, 0.43)	0.34 (0.22, 0.46)	0.32 (0.22, 0.42)	0.15	15.6% (12.6%, 18.5%)	14.1% (10.4%, 17.9%)	6.6% (4.7%, 8.5%)	6.9% (4.5%, 9.2%)	6.4% (4.3%, 8.4%)	82.4% (77.9%, 86.8%)	81.6% (76.3%, 87%)	83% (78.1%, 88%)
Standard error	0.05	0.06	0.05	1.5%	1.4%	1.8%	0.9%	1.1%	1.0%	2.1%	2.5%	2.4%
Kiikaonde (*n*=200)	0.21 (-0.49, 0.92)	0.17 (-0.53, 0.86)	0.26 (-0.49, 1.01)	7.1% (-11%, 25%)	5.8% (-10.9%, 22.5%)	8.4% (-11.2%, 27.9%)	4.3% (-9.9%, 18.4%)	3.3% (-10.5%, 17.2%)	5.2% (-9.7%, 20.2%)	88.7% (50.5%, 126.8%)	91.2% (54.6%, 127.8%)	86.1% (45.1%, 127%)
Standard error	0.16	0.16	0.17	4.1%	3.9%	4.5%	3.3%	3.2%	3.5%	8.9%	8.5%	9.5%
Lunda (*n*=131)	0.63 (0.55, 0.7)	0.89 (0.79, 0.99)	0.4 (0.38, 0.41)	14.2% (11%, 17%)	4.7% (2.4%, 7%)	3.4% (0.7%, 6.2%)	12.5% (11%, 14%)	17.8% (15.8%, 19.7%)	7.9% (7.6%, 8.2%)	61.7% (42.8%, 80.6%)	49.2% (20.9%, 77.6%)	72.7% (64.6%, 80.7%)
Standard error	0.02	0.02	0.00	0.7%	0.7%	0.4%	0.4%	0.5%	0.1%	4.4%	6.6%	1.9%
Luvale (*n*=254)	0.36 (0.27, 0.45)	0.33 (0.21, 0.45)	0.39 (0.28, 0.49)	8.7% (7%, 11%)	4.7% (2.4%, 7%)	3.4% (0.7%, 6.2%)	7.2% (5.3%, 9%)	6.6% (4.2%, 9.1%)	7.7% (5.5%, 9.9%)	84.5% (81.9%, 87%)	86.8% (82.6%, 90.9%)	82.3% (77.5%, 87.2%)
Standard error	0.03	0.04	0.03	0.7%	0.7%	0.8%	0.6%	0.8%	0.7%	0.8%	1.3%	1.5%

Table A-6. Reading Comprehension

	NO. CORRECT RESPONSES / MINUTE			% CORRECT / ATTEMPTED			% CORRECT / TOTAL			% ZERO SCORES		
	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS	OVERALL	BOYS	GIRLS
Silozi (*n*=400)	0.74 (0.34, 1.14)	0.73 (0.06, 1.39)	0.75 (0.41, 1.1)	0.22	4.7% (2.4%, 7%)	3.4% (0.7%, 6.2%)	14.8% (6.8%, 22.8%)	14.5% (1.3%, 27.8%)	15% (8.1%, 21.9%)	68.5% (47.1%, 89.9%)	67.7% (36.3%, 99.1%)	69.3% (54.5%, 84%)
Standard error	0.13	0.21	0.11	3.5%	5.6%	2.3%	2.5%	4.2%	2.2%	6.7%	9.9%	4.6%

ANNEX B. BASELINE COMPARISON REPORT

OVERVIEW

The purpose of this report is to compare baseline student scores with the student scores from the grade 2 NAS, which serves as a midterm evaluation. For that purpose, the baseline student scores were created from reading evaluations completed under three USAID projects in Zambia: Read to Succeed, Time to Learn, and the Education Data for Decision Making (EdData II) Early Grade Reading Assessment (EGRA) in Bemba. These were conducted in 2012 and 2013.

Establishment of baseline values was a challenge because of the multiple types of locations from which schools were previously sampled. For example, the schools that were sampled for the Read to Succeed project were very much urban, whereas the schools from the 2013 EdData II assessment were sampled from provinces—mostly in rural areas—with Bemba as their primarily language. If a study sampled only urban or only rural schools for a specific language, we were not able to make that estimate "representative" of that demographic.

Additionally, in order to have a more homogenous sample, we were constrained in using only non-private schools (i.e., government, community, and grant-maintained schools). By comparison, the grade 2 NAS sample included private schools. The three previous studies combined totaled 306 sampled schools.

Thus, the baseline estimates for the following languages were restricted by the characteristics of the previous studies.

- **Bemba baseline estimates:** The 2013 Bemba data were "representative" of both urban and rural schools. For this exercise, we used a total of 13 urban and 82 rural schools for weight estimations.
- **Cinyanja baseline estimates:** Similar to the Bemba language, the 2013 sample for the Cinyanja language was composed of both urban and rural areas. We used 19 urban schools and 84 rural schools for weight estimations.
- **Kiikaonde baseline estimates:** The available Kiikaonde estimates are "representative" of rural schools only. We used 42 schools—41 GRZ and 1 community—to estimate the weights.

- **Chitonga baseline estimates:** Similar to Kiikaonde, the scores available for the Tonga language are "representative" of rural schools only; we used 10 rural schools for weight estimations.
- **Silozi baseline estimates:** For the Silozi language, we had sample reading assessment results only from the GRZ schools. We used 46 rural and 6 urban schools to estimate the weights.

CALCULATION OF EQUATED MIDTERM SCORES

The equated midterm pupil estimates are the grade 2 NAS scores from the main section of this report multiplied by the equating ratio as described in detail in ***Annex C.*** Note that only the subtasks related to the passage reading were equated—that is, the oral reading fluency and reading comprehension scores. ***Table B-1*** shows the multiplier ratios that were applied to student scores to create the equated student scores.

Table B-1. Equating ratios to be applied to midterm Oral Reading Fluency and Reading Comprehension scores

		LOCAL LANGUAGE				
		CHITONGA	CINYANJA	ICIBEMBA	KIIKAONDE	SILOZI
Ratio multiplier for midterm scores	Oral Reading Fluency score	1.073	0.837	0.770	0.728	1.114
	Reading Comprehension score	0.912	0.461*	0.930	0.880	0.727

SUMMARY OF SCORES

Table B-2 shows the baseline estimates for Oral Reading Fluency by language next to their equated midterm estimates. Although there are some differences to be noted where the confidence intervals for the paired estimates do not overlap, to conclude this being a significant change in the oral reading fluency rate would be challenging due to the limitations created by the sample from which the baseline was created. For example, the estimates for Cinyanja were 1.11 (0.71, 1.5) and 3.44 (2.46, 4.41) words per minute for the baseline and midterm, respectively. Although these estimates do not have overlapping 95% confidence intervals, recall that the schools in the baseline sample were limited non-private schools whereas the midterm also included private schools, which constitute a higher percentage of the schools in Lusaka where Cinyanja is the language of instruction. Additionally, both sets of scores were low; students at both baseline and midterm on average sight-recognized only a few words.

Table B-2. Oral Reading Fluency: Pupil estimates for baseline and equated midterm, by language

	ORAL READING FLUENCY (WORDS PER MINUTE)			
LANGUAGE	BASELINE		EQUATED MIDTERM	
Chitonga	0.46	(-0.44, 1.36)	1.98	(0.69, 3.28)
Standard error	0.4		0.56	
Cinyanja	1.11	(0.71, 1.5)	3.44	(2.46, 4.41)
Standard error	0.2		0.43	
Icibemba	1.5	(0.93, 2.07)	5.87	(4.61, 7.13)
Standard error	0.29		0.60	
Kiikaonde	0.89	(0.39, 1.38)	2.52	(-3.24, 8.27)
Standard error	0.25		1.34	
Silozi	1.31	(0.58, 2.05)	9.36	(4.3, 14.41)
Standard error	0.37		1.59	

After the students attempted to read the passage in an EGRA, they then answered reading comprehension questions based on how far in the passage they managed to read. These results are shown in ***Table B-3.*** Because the average reading fluency rates were low (as shown in Table B-2), students were not able to read many words and struggled to connect the words into meaningful sentences. As a result, the reading comprehension scores (number of answers correct out of 5) were also low.

Table B-3. Reading Comprehension: Pupil estimates for baseline and equated midterm, by language

	READING COMPREHENSION SCORE (NO. CORRECT OUT OF 5)			
LANGUAGE	BASELINE		EQUATED MIDTERM	
Chitonga	0.006	(-0.01, 0.02)	0.12	(0.03, 0.21)
Standard error	0.006		0.04	
Cinyanja	0.027	(0.003, 0.05)	0.17	(0.13, 0.2)
Standard error	0.012		0.02	
Icibemba	0.021	(0.01, 0.03)	0.31	(0.22, 0.4)
Standard error	0.004		0.04	
Kiikaonde	0.013	(0.01, 0.02)	0.19	(-0.43, 0.81)
Standard error	0.004		0.14	
Silozi	0.019	(0.01, 0.03)	0.54	(0.25, 0.83)
Standard error	0.37		1.59	

Table B-4 shows the per minute pupil estimates for the Nonword Reading subtask (correct invented words). The overall average for correct nonwords per minute in Kiikaonde was 7.51 at baseline and 2.27 for the midterm. The confidence intervals for these two estimates (5.01, 10.02) and (-2.88, 7.41) overlap, indicating that due to the variability in the samples, it would not be prudent to conclude that one estimate was statistically significantly higher than the other.

Table B-4. Nonword Reading: Correct invented words pupil estimates for baseline and midterm, by language

	CORRECT INVENTED WORDS PER MINUTE			
LANGUAGE	BASELINE		MIDTERM	
Chitonga	0.42	(-0.31, 1.14)	1.66	(0.85, 2.48)
Standard error	0.31		0.35	
Cinyanja	1.56	(0.992, 2.13)	3.20	(2.32, 4.07)
Standard error	0.29		0.39	
Icibemba	2.05	(1.22, 2.87)	5.60	(4.44, 6.77)
Standard error	0.41		0.56	
Kiikaonde	7.51	(5.01, 10.02)	2.27	(-2.88, 7.41)
Standard error	1.2		1.20	
Silozi	9.71	(7.91, 11.52)	5.65	(3.08, 8.22)
Standard error	0.85		0.81	

The estimates for the Letter Sounds subtask are shown in ***Table B-5.*** The only language for which the estimates do not overlap is Icibemba, which reported 4.84 (3.82, 5.86) and 8.77 (7.04, 10.51) words per minute for the baseline and midterm, respectively.

Table B-5. Letter Sounds: Pupil score estimates for baseline and midterm, by language

	CORRECT LETTER SOUNDS PER MINUTE			
LANGUAGE	BASELINE		MIDTERM	
Chitonga	0.61	(0.11, 1.11)	3.68	(0.85, 2.48)
Standard error	0.22		0.34	
Cinyanja	5.26	(4.4, 6.13)	7.21	(5.01, 9.4)
Standard error	0.44		0.97	
Icibemba	4.84	(3.82, 5.86)	8.77	(7.04, 10.51)
Standard error	0.51		0.83	
Kiikaonde	6.37	(5.26, 7.47)	5.44	(-5.76, 16.64)
Standard error	0.54		2.60	
Silozi	8.79	(7.06, 10.53)	8.09	(-1.26, 17.44)
Standard error	0.85		2.94	

ANNEX C. REPORT ON THE EQUATING PILOT STUDY

SUMMARY

The ratios shown in ***Table C-1*** were calculated to be applied to the midterm Oral Reading Fluency and Reading Comprehension scores to equate them to the baseline scores, as characterized in ***Annex B.*** The equated summary variables were created by multiplying the midterm scores by the appropriate ratio detailed in Table C-1.

Note that this exercise equated mean Oral Reading Fluency and Reading Comprehension subtask scores only, rather than scores for a full set of EGRA subtasks.

Table C-1. Equating ratios to be applied to midterm oral reading fluency and reading comprehension scores

		LOCAL LANGUAGE				
	SUBTASK	CHITONGA	CINYANJA	ICIBEMBA	KIIKAONDE	SILOZI
Ratio multiplier for midterm scores	Oral Reading Fluency	1.073	0.837	0.770	0.728	1.114
	Reading Comprehension Score	0.912	0.461*	0.930	0.880	0.727

* This is a quite an extreme ratio multiplier. In other words, the two Reading Comprehension scores for Cinyanja did not equate well. A note to this effect appears in Annex B.

BACKGROUND ON EQUATING

Equating is a statistical procedure used to convert scores from multiple forms, or variations, of a test to the same common measurement scale. It is common—for reasons such as test security or instrument validation—for evaluators to prepare and administer more than one version. This conversion process adjusts for any differences in difficulty levels among test forms so that a score on one form can be matched to an equivalent value on another form. In other words, equating

assumes that even the most carefully crafted test forms will have some inherent differences, and makes it possible to estimate the score that a person taking one test form would have received had they taken a different test form (Holland & Dorans, 2006; Kolen & Brennan, 2004).

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

This annex specifically examines equating for the local-language Oral Reading Fluency and Reading Comprehension scores, the two student subtask scores assessed from the Early Grade Reading Assessment (EGRA) reading passage. The initial passages were developed in 2012 in five Zambian languages: Chitonga, Cinyanja, Icibemba, Kiikaonde, and Silozi.

Note that for purposes of this analysis, *baseline* refers to the instruments used in Zambia in 2012 and 2013, while *midterm* refers to the grade 2 National Assessment Survey instruments administered in 2014 and analyzed in this report.

For the grade 2 National Assessment Survey, the equating recommendations were applied only to the student scores for Reading Comprehension and Oral Reading Fluency when they were compared with the re-weighted baseline, as discussed in Annex B.

PILOT PROCEDURE

The pilot was conducted October 2–3, 2014. The aims of the pilot, in addition to the equating, were to assess the quality of data and performance of the assessors along with item-level and variable frequency analysis of the new instruments.

All students assessed were asked to read both the baseline and midterm passages (given in a randomized order) and then answer a number of reading comprehension questions based on how far through the passage they got. Because equating is possible only with nonzero scores, it was decided that the sample should be grade 4 students and selected grade 2 students who could read. Based on the results from the previous early grade reading evaluations in Zambia, it was expected that had the sample consisted of randomly selected grade 2 students, it was unlikely that there would have been enough readers in the pilot sample to sufficiently equate the passages, if necessary.

Details on the final pilot sample appear in ***Table C-2.*** It was reported by assessors that identifying grade 2 students who could read was challenging. Indeed, even after teachers had specifically pointed out grade 2 students who could read, it was found that many of them could not. This situation justified the decision to sample grade 4 students to ensure the pilot sample would contain enough nonzero scores.

Table C-2. Number of students sampled for pilot assessment, by language and grade

	LANGUAGE AND GRADE									
	CHITONGA		CINYANJA		ICIBEMBA		KIIKAONDE		SILOZI	
	G2	G4	G2	G4	G2	G4	G2	G4	G2	G4
Number of students sampled:	5	220	18	168	13	412	6	188	6	110
Number of students scoring zero (both passages):	5	81	5	52	3	88	2	80	0	32

ANALYSIS PROCESS

Because the students were assessed for their reading fluency and comprehension in both passages, this was a matched-pairs study in which each student was his/her own match. Any scores that were unreasonable (for example, a student scoring 1 word per minute [WPM] on the baseline passage and 200 WPM on the midterm passage) were set to "missing because of error in assessment." For small samples such as this one, assessment errors can heavily influence mean scores.

The equating technique proposed to be applied to the passages was a *means ratio method.* For this method, the means of the baseline and midterm scores were calculated and a ratio of the two scores was created by dividing the baseline mean by the midterm mean. Then, all the midterm means were multiplied by the ratio. This step equated the midterm scores to the baseline scores. For this process to be confirmed as valid, however, the following assumptions needed to be assessed:

- The scatterplot of the baseline and midterm reading fluencies had to be roughly linear, with an intercept of zero.
- The distribution of scores in the midterm pilot had to be similar to the actual midterm scores even after zero scores were removed, which would take into account differences in the attributes of the sampled students.

RESULTS BY LOCAL LANGUAGE

CHITONGA

Figure C-1 shows a roughly linear pattern with the matched pilot data for the Oral Reading Fluency subtask for Chitonga. While there was more variability as oral reading fluency rates exceeded 40 words per minute, the straight line (i.e., the regression line forced through the origin) fits the data to sufficiently validate the equating procedure. ***Figure C-2*** shows the distribution of student scores once the zero scores were removed, in order to judge whether the students in the pilot and study samples were from the same population.

Figure C-1. Scatterplot of baseline oral reading fluency (ORF 2) versus midterm oral reading fluency (ORF 1) for Chitonga passages

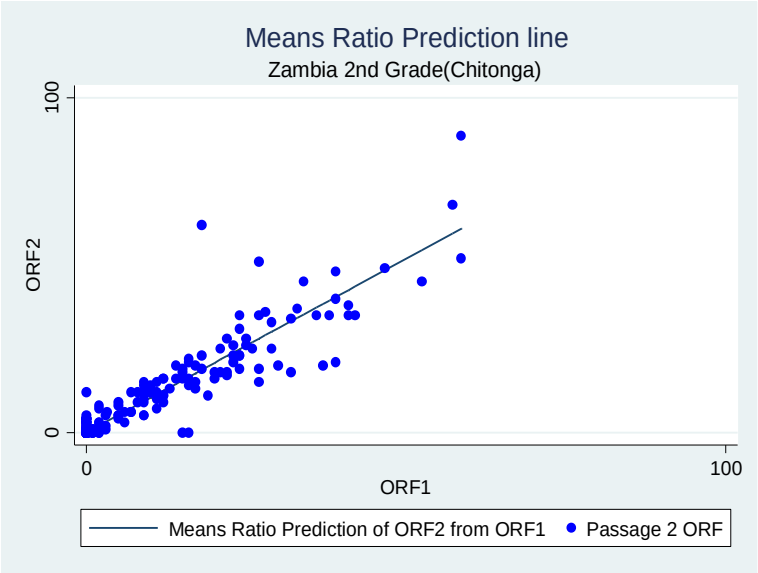
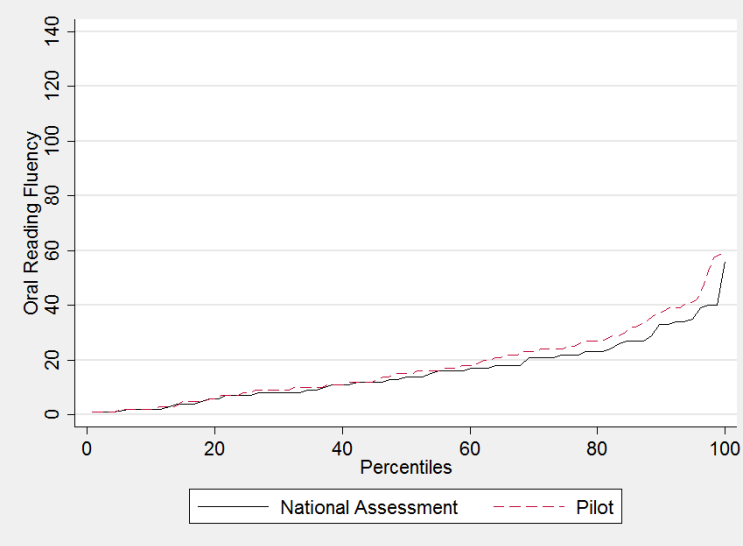


Figure C-2. Cumulative distribution plots of pilot versus actual reading fluency scores for Chitonga passages



CINYANJA

The relationship between the reading fluency obtained by the pupils in the pilot sample for Cinyanja is explored with the scatterplot in ***Figure C-3.*** The pattern indicates a reasonable linear association between the variables. ***Figure C-4*** shows that while the pilot data scores were stronger (40th percentile was approximately 13 WPM for the NAS and 17 WPM for the pilot), they are sufficiently close, particularly considering the low average scores for the grade 2 NAS and the small number of students who scored over 20 WPM.

Figure C-3. Scatterplot of baseline oral reading fluency (ORF 2) versus midterm oral reading fluency (ORF 1) for Cinyanja passages

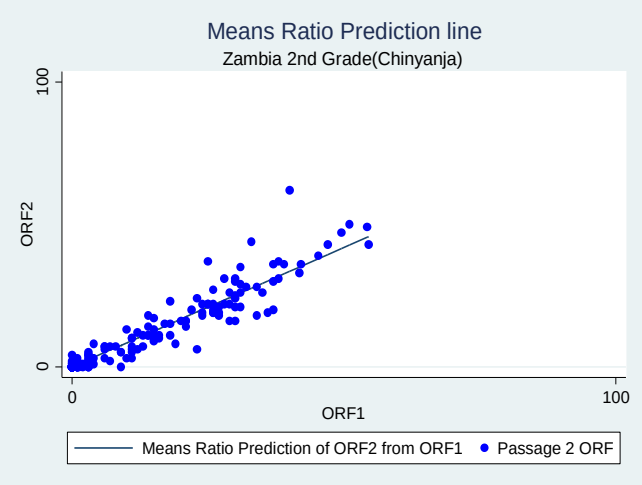
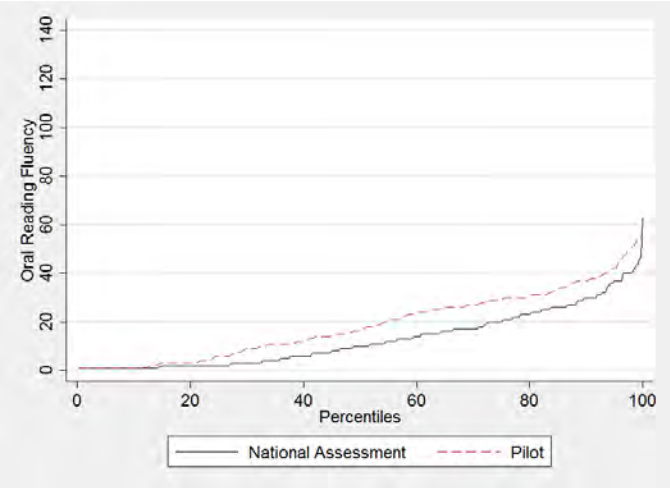


Figure C-4. Cumulative distribution plots of pilot versus actual reading fluency scores for Cinyanja passages



ICIBEMBA

The two Oral Reading Fluency subtask scores for Icibemba are shown in ***Figures C-5 and C-6***. Ignoring the clear challenges with the administration of the midterm reading assessment, where some students scored high in one passage but zero in another, again the linear association looks acceptable. The cumulative distributions in Figure C-6 are similar.

Figure C-5. Scatterplot of baseline oral reading fluency (ORF 2) versus midterm oral reading fluency (ORF 1) for Icibemba passages

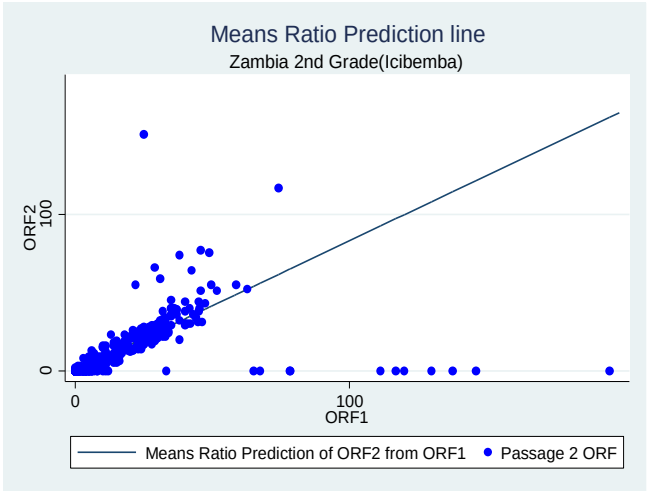
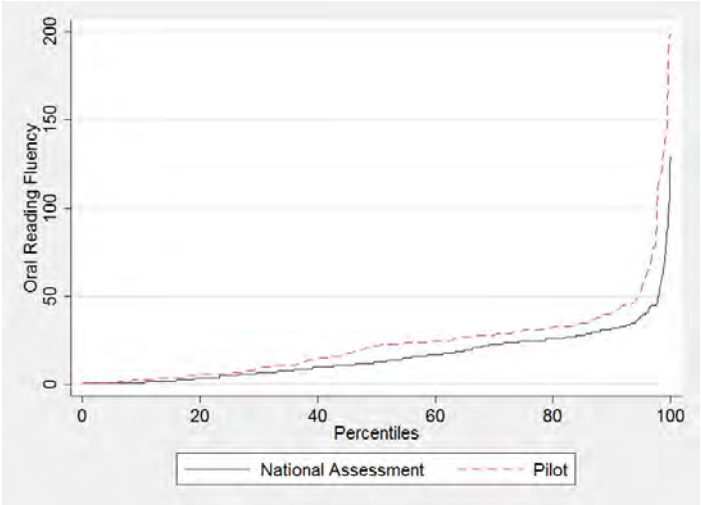


Figure C-6. Cumulative distribution plots of pilot versus actual reading fluency scores for Icibemba passages



KIIKAONDE

The fluency scores for the two passages in Kiikaonde also had a linear association, as shown in ***Figure C-7.*** The cumulative distributions in ***Figure C-8*** seem to become more dissimilar as the oral reading fluency rate increases, but are very close when the scores are 20 WPM or less. Considering the majority of students scored in this range, these distributions are acceptable.

Figure C-7. Scatterplot of baseline oral reading fluency (ORF 2) versus midterm oral reading fluency (ORF 1) for Kiikaonde passages

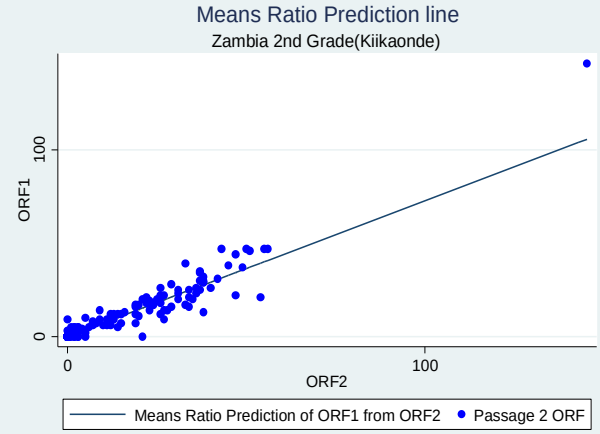
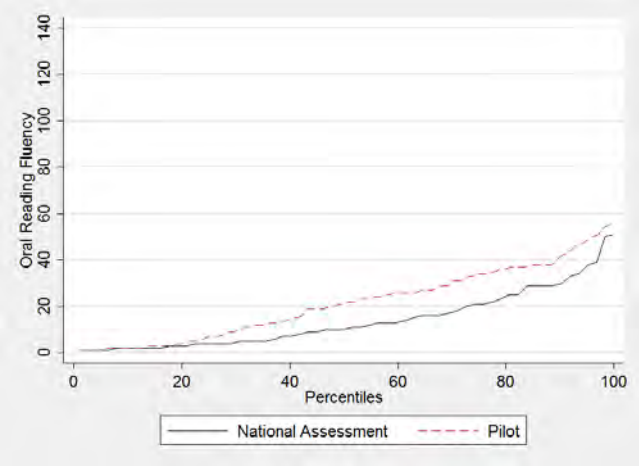


Figure C-8. Cumulative distribution plots of pilot versus actual reading fluency scores for Kiikaonde passages



SILOZI

While the Silozi reading fluency scores had more students in the range greater than 50 wpm, the association between the two passages appears linear, as shown in ***Figure C-9.*** The distribution of student scores in Silozi for oral reading fluency is shown in ***Figure C-10.*** Like in other languages, the distribution of scores varies more as the scores increase, but again because most of the Oral Reading Fluency subtask scores were low, we focus on the left-hand side of the figure where the distributions are suitably similar.

Figure C-9. Scatterplot of baseline oral reading fluency (ORF 2) versus midterm oral reading fluency (ORF 1) for Silozi passages

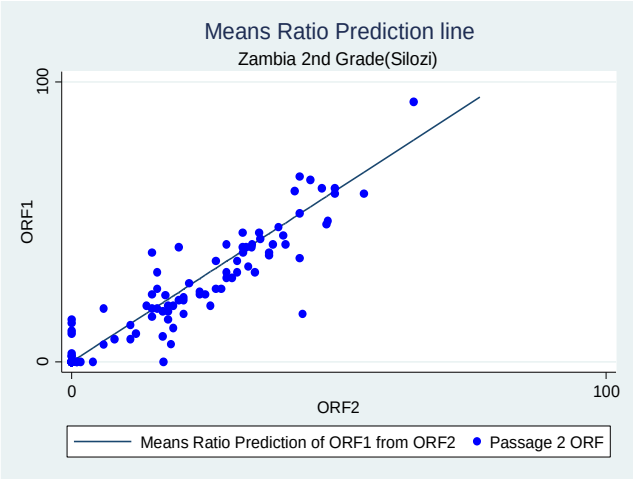


Figure C-10. Cumulative distribution plots of pilot versus actual reading fluency scores for Silozi passages

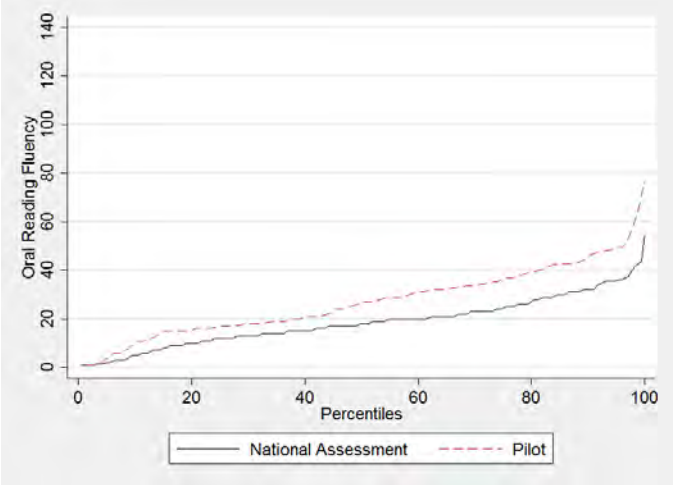


Table C-3 shows the mean scores calculated for the Oral Reading Fluency and Reading Comprehension subtasks, by language and passage. These scores were used to calculate the equating ratios.

Table C-3. Mean, standard error, and sample size for midterm and baseline passages, by language

LANGUAGE		READING COMPREHENSION SCORE		ORAL READING FLUENCY (WORDS PER MINUTE)	
		MIDTERM PASSAGE	BASELINE PASSAGE	MIDTERM PASSAGE	BASELINE PASSAGE
Chitonga	Number of students	98	104	223	223
	Mean	1.6	1.4	9.1	9.8
	Standard error	0.1	0.1	0.9	1.0
Cinyanja	Number of students	112	92	184	184
	Mean	1.8	0.8	12.6	10.5
	Standard error	0.1	0.1	1.1	1.0
Icibemba	Number of students	256	226	416	415
	Mean	1.4	1.3	17.9	13.8
	Standard error	0.1	0.1	1.2	1.0
Kiikaonde	Number of students	85	86	190	189
	Mean	1.5	1.3	13.3	9.7
	Standard error	0.2	0.1	1.5	1.1
Silozi	Number of students	74	77.0	116	115
	Mean	2.8	2.0	18.1	20.8
	Standard error	0.2	0.1	1.6	1.9

REFERENCES TO ANNEX C

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

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices.* New York: Springer Science+Business Media.