USAID
FROM THE AMERICAN PEOPLE

IMPACT EVALUATION FOR THE USAID/ *APRENDER A LER* PROJECT IN MOZAMBIQUE

Baseline Report

Revised June 13, 2013

This report was prepared for USAID/Mozambique by Magda Raupp, Bruce Newman and Luis Revés under Evaluation Services IQC Task Order AID-656-TO-12-00002 awarded to International Business & Technical Consultants, Inc. (IBTCI), with Global Surveys Corporation (GSC Research) as sub-contractor. The authors' views expressed in this plan do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Intentionally blank inside front cover

IMPACT EVALUATION FOR THE USAID/*APRENDER A LER* PROJECT IN MOZAMBIQUE

Baseline Report

Prepared by

International Business & Technical Consultants Inc.

Magda Raupp, Team Leader
Bruce Newman, Statistician/ Data Specialist
Luis Revés, GSC Research, Deputy Team Leader

Revised June 13, 2013

Evaluation Services IQC Task Order AID-656-TO-12-00002
Deliverable 5

CONTENTS

ACRONYMS

APAL	Aprender a Ler
EGRA	Early Grade Reading Assessment
EPT	Éducation pour tous (Education for All)
GSC	Global Surveys Corporation
IBTCI	International Business & Technical Consultants, Inc.
IE	Impact Evaluation
IFP	Instituto de Formação dos Professores
IGA	Institutional Gap Analysis
M&E	Monitoring & Evaluation
MINED	Ministry of Education
RCT	Randomized Control Trial
SDEJT	Service for Education, Youth and Technology
SMA	School Management Assessment
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
UNESCO	United Nations Educational, Scientific, and Cultural Organization
USAID	U.S. Agency for International Development
WEI	World Education Inc.
ZIP	Zonas de Influência Pedagógica

EXECUTIVE SUMMARY

This report presents the findings of baseline data collection conducted from second and third grade students in February/March 2013, at the beginning of the 2013 school year, in 180 schools located in the economic corridors of the Nampula and Zambézia provinces in Mozambique. Evidence-based information resulting from the Baseline Study, which lays the ground for the Impact Evaluation (IE), describes the current situation found in the participating schools and could inform USAID/*Aprender a Ler* (APAL) implementation, and hopefully influence future education policy decisions made by Ministry of Education in Mozambique (MINED). Subsequent data collections are scheduled for September/October of 2013 and 2014.

Assessments of student learning in the primary grades, such as the Early Grade Reading Assessment (EGRA) used in this Baseline Study, offer an opportunity to determine whether children are developing the fundamental skills upon which all other literacy skills build, and, if not, where efforts might be best directed. This is vital information for MINED to help guide the efforts to improve the quality of education in the early grades. The School Management Assessment (SMA) instrument administered in the schools sampled provides a multifaceted view of school and classroom characteristics traditionally associated with student performance.

The findings and conclusions included in this report refer exclusively to a sample of 180 schools located in a set of randomly selected Zonas de Influência Pedagógica (ZIPs) along the economic corridors of the Nampula and the Zambézia provinces and should not be generalized to other schools in other areas of Mozambique. Additionally, data for the Baseline Study were collected at the beginning of the 2013 school year. Therefore, the EGRA scores reflect the performance of the randomly selected 1,798 second and 1,800 third graders at the beginning of second and third grades respectively.

Education Background

Since 1992, following a long colonial period, a 10-year war for independence, and 16 years of civil war, the government of Mozambique has been rebuilding and improving its educational system, and access to primary education has expanded extremely rapidly. Over the period 2000 (SACMEQ II) to 2007 (SACMEQ III), Mozambique has embarked on a rapid expansion of access to primary schooling. In the year 2000, 218,594 students were enrolled in grades 6 and 7 compared to 852,811 in 2011 (Estatistica da Educação: Levantamento Escolar-2011, MINED), an expansion of 390% in little over a decade. This rapid expansion has placed a large burden on an already struggling system resulting in double and triple shift schools, too few qualified teachers, and an overburden on school and district managers, among other critical challenges.

To a certain extent, Mozambique's public education system still suffers from the impact of sustained conflict over many years. Although improvement efforts are underway, school buildings are insufficient and many are in need of significant repair. The inclusion of populations that previously had no access to education has posed major challenges to the education system. To face the influx of new students, schools are frequently forced to function in shifts, which reduces learning opportunities for students. At the same time, the educational system has had to hire teachers not yet prepared for the task of teaching reading to young children or to children who speak a local language as well as hiring directors who do not have the adequate managerial and instructional leadership skills to run a school focused on learning improvement. This challenges the education system but the government is committed to create a more

equitable system and has made significant progress since 1999.[1] The enrollment rate is up, which is a first step to equity, but the inclusion of new student populations into the education system is often accompanied by the late entry and repetition that can be observed in the schools sampled.

In Mozambique primary education is free and compulsory and the official age of entry into school is 6 years. Primary education is subdivided into two levels: lower primary, which consists of five years of schooling (Grades 1 to 5) enrolling a total of 4,373,181 students and upper primary, which comprises two years (Grades 6 and 7) with an enrollment of 852,811 in 2011 (*Estatistica da Educação: Levantamento Escolar-2011*, MINED). Usually, primary schools operate in two shifts. Because of the shortage of school places at this level, some primary schools operate three shifts mainly in urban areas.[2]

The management of the educational system is organized in various levels under the direction of the Ministry of Education. There is a Provincial Directorate of Education for each of the 11 provinces, and this directorate is under the command of a Provincial Director. Below the Provincial Directorate there is the District Directorate headed by a District Director. There are 146 districts in Mozambique. Below the District Directorate there are clusters of schools known in Portuguese as *Zonas de Influência Pedagógica* (ZIPs) formed by one "head" school and a cluster of five, ten or more schools. The "head" school of a ZIP is usually better organized than the surrounding schools and is tasked with reporting to the district and receiving and distributing teaching material and resources that are sent by the MINED. The "head"of the ZIP schools also serve as an example and a magnet for training and improvement of the surrounding schools.

Purpose and Design of the Impact Evaluation

The purpose of the Impact Evaluation (IE) is to address the evaluation questions related to the effectiveness of the intervention implemented by the USAID/APAL and the cost effectiveness and sustainability of two intervention levels. The IE calls for qualitative and quantitative analysis of results obtained by second and third grade students in the 120 schools where USAID/APAL will implement an early grade reading intervention and students and in 60 "Control" schools. The IE uses the Randomized Control Trial (RCT) methodology, with pre-intervention random assignment of 60 schools to receive "Full" treatment, 60 schools assigned to "Medium" treatment, and 60 no-treatment or "Control" schools.[3]

The RCT methodology is best suited for analysis of specific programs or resources that can be manipulated easily within an experiment. The use of RCT is the most effective way to measure impact for three main reasons. First, it allows for direct attribution of the USAID/APAL interventions to improved outcomes because the RCT model controls for other possible determinants of the outcome. Second, the random component of RCT reduces effects of potential unobservable differences between treatment and control groups on the outcome. Third, an RCT provides a rigorous evaluation method to obtain accurate and valid results to inform plans to scale up the most effective and cost-effective intervention. However, it must be noted that the design of the APAL intervention was specific regarding the geographical area where the activities would be implemented: the economic corridors of the Nampula and Zambézia provinces and the number of schools that would participate (180). This forced the IE to select and assign ZIPs and schools to the treatment and control groups within those parameters. Annex A includes maps showing the location of the schools in the sample.

[1] UNESCO. Education for All Global Monitoring Report. http://www.unesco.org/.
[2] The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ). Report on Education 2009-2013
[3] Schools in the "Medium" treatment group will receive teacher in-service training and materials, and those in the "Full" treatment will receive teacher and school management in-service training and materials. Schools in the "Control" group will not receive treatment but will continue, as will the treatment schools, receive MINED interventions.

How Well Are Students Learning to Read?

The EGRA Mozambique[4], which was administered orally to individual students in Portuguese, consisted of seven subtasks: (1) oral vocabulary; (2) oral instructions; (3) concepts about print; (4) letter naming; (5) decoding of individual words; (6) connected text oral reading fluency; and (7) answering reading comprehension questions.

Oral vocabulary and following oral instructions assess children's ability to understand and to respond to instructions given by the teacher. Concepts about print looks at children's familiarity with printed text—recognizing the beginning and the ending of a story, the title, identifying page numbers, etc.—and assesses the child's readiness to handle printed materials. Together with letter knowledge and the ability to read isolated words, these are foundational skills needed for reading with comprehension. The letter naming, reading of individual words, and reading and comprehension subtasks were designed to be timed. However, an examination of the number of letters read (Annex H) shows that over two-thirds of the students attempted to read ten letters or less and over 90% could not read more than two words. In order to not frustrate or embarrass students the EGRA establishes cut off points when examinees are unable to perform the task. For example, if a child cannot read correctly any of the first ten letters the task is terminated. In September/October 2013 and 2014, when we expect performance to improve the time limits will be enforced. The time limit makes it possible to assess the degrees of automaticity achieved in these skill areas. Timed subtasks are usually scored as correct letters per minute or correct words per minute, while untimed tasks are scored as total items correct. Reading fluency required students to read two stories and answer eight reading comprehension questions. Students who were unable to perform a single item on a subtask received a zero score.

Summary of Findings

- The great majority of the 3,598 (1,805 girls and 1,793 boys) second and third graders assessed in the 180 randomly selected schools are not acquiring the foundational skills that will allow them to become fluent readers. This is not very different from the results of the SACMEQ II assessment, which found that only 2.3% of the first graders and 3.9% of second graders had acquired pre-reading and emergent reading skills.[5] Only a very small group of students (less than 2%) demonstrated the desired level of reading fluency.
- In Nampula, where Portuguese is less frequently used than in Zambézia, students take longer to acquire the foundational skills necessary to read. The differences in the mean scores of the two provinces are even more striking due to the high proportion of students in Nampula with little or no ability to follow oral instructions—43% in second and 34% in third grade.
- A review of the school effectiveness literature helps identify school characteristics that limit or impede learning—in grades one and two, the primary focus of learning is the acquisition of reading and basic mathematics skills. Typically, the school characteristics include: inadequate infrastructure; limited school time; insufficient time dedicated to reading; student, teacher, and director absenteeism and tardiness; deficient preparation of teachers and directors; and, absence of a strong of commitment to improving instruction on the part of a good number of school directors.[6] The

[4] The EGRA/Mozambique was adapted from the EGRA developed by the Aga Khan Foundation in Cabo Delgado, Mozambique (2010)

[5] The SACMEQ II Project in Mozambique: A Study of the Conditions of Schooling and the Quality of Education. 2005

[6] Glewwe, Hanushek, Humpage, and Ravina (2011) examined in depth 79 studies (out of 9,000 originally selected) to investigate which education inputs appear to have strong positive impacts on learning. A few variables were found to have statistically significant effects—e.g., availability of desks, tables and chairs and walls or floors; having teachers with greater knowledge of the subject they teach; longer school day; and, providing tutoring. Additionally, and not surprising, teacher absenteeism has a clear negative effect on learning.

information on school characteristics collected by the School Management Assessment (SMA)[7] ties in with school effectiveness research and can be used to explain the low level of reading skills demonstrated by students.

- The high level of confidence in their teaching skills reported by teachers contradicts the low level of reading skills exhibited by the students.

Summary Conclusions

- If we start from the premise that all children can learn, it becomes necessary to look at the various obstacles built into the education system that reduce learning opportunities for children, specifically, opportunities to learn to read in the early grades.
- Data collected by the SMA show that in 65% of the 92 classes observed in Nampula children were sitting on the bare ground; the same was observed in Zambézia in 63% of the 81 classes observed. In addition, in Nampula 15% of the classes were held under a tree. This undoubtedly limits learning opportunities.[8]
- Due to teacher tardiness, classes frequently started late: 54% of the teachers of the 360 second and third grade classes observed arrived between 10 and 40 minutes late. Considering that the first class of the day is frequently Portuguese, where reading occurs, it is easy to see how reading-specific instructional time is reduced by teacher tardiness.
- In the 180 schools visited, 53% of the directors in Nampula and 68% in Zambézia arrived as much as sixty minutes late. When over half of the directors arrive late, as documented by the SMA (even when our visit had been announced), it is reasonable to say that under these circumstances school management may have difficulty in enforcing standards of punctuality and instructional time or time-on-task.
- Given that only 2% of the students who took the EGRA could read with fluency, there is clearly an absence of focused teaching strategies that give children the opportunity to become readers by the end of third grade. Director, teacher, and student tardiness and absenteeism reduce the number of instructional hours to which the student is exposed each day and make it difficult for the student to acquire the reading competencies needed to progress in school.[9] The first step in the direction of more instructional time per day could be the reduction of the times a teacher or a director is late or absent.
- Increasing the number of instructional hours per day may be necessary but it is not sufficient—as much attention needs to be given to what happens during these instructional hours. Class observations conducted during language instruction in the 180 schools found that in Nampula 37% and in Zambézia 61% of the students were not engaged in any of the activities that were taking place in the classroom. Simply increasing time in school without equal additions to learning and achievement has little value (Hanushek and Woessmann, 2008). Both instructional time and quality of content delivery are critical to student learning.
- The number of classes being conducted "under a tree" (15% in Nampula) and the number of classes conducted with students sitting on the bare ground (62% in Nampula and 63% in Zambézia) suggest that alternative ways to teaching reading, ways that do not assume that children will be seated in a formal classroom, need to be discussed. The number of students in class (57 average) suggests that ways of teaching based on forming groups within a large class should be attempted to offer other

[7] The SMA was adapted by the implementer from the instrument utilized by the Aga Khan Foundation (2010) in Cabo Delgado, Mozambique (EQUIP2)

[8] After examining 79 high quality studies of factors that limit learning, Glewwe et al. found that "…*adequate amounts of desks, tables and chairs raise student test scores, as common sense would suggest.*" [8] and they continue to state that *"…it is also the case that as shown in 83% of research papers examined, that schools that have walls, roofs and floors appear to lead to better outcomes…"*

[9] Abadzi, H. (2007). "Absenteeism and Beyond: Instructional Time Loss and Consequences", World Bank Policy Research Working Paper No. 4376, p. v.

types of instruction that are not based solely in attempting to engage the whole group.

- Interviews conducted with MINED central, provincial and district officers, ZIP coordinators, pedagogical directors, school directors and PTA presidents indicate that there is concern regarding the acquisition of reading skills in the early grades. Challenges identified include the inadequate level of teacher preparation, the insufficient quantity of learning/teaching materials and the limited financial resources available at the schools. The ADE (*Apoio Directo a Escola)* is considered essential since it is the only source of financial resources available to schools. The interviewees, almost unanimously, consider the ZIP strategy (having schools grouped in a cluster around a "head"school) to be the correct strategy to improve education because it facilitates teacher training and the delivery of books and materials. The head schools of the ZIPs already conduct much of the training and take initiatives to improve teaching and learning and facilitate open communication to the district and to the MINED. However, ZIP coordinators point out that in order to fulfill their function they require additional resources
- Out of the 3,598 second and third grade students, 11% reported not living with their mother and 18% not living with the father (Annex H). When asked the reason why they did not live with their mother and/or their father, "orphanage" was reported as the reason by 428 children out of the 817 who reported not living with mother and/or father. The majority of other reasons given (48%) were related to separation, divorce or relocation for work demonstrating that there are multiple causes for the absence of one or more parents. Of the 341 teachers of students assessed, 218 (64%) indicated that they had no students with disabilities. Of the 123 who did report students with disabilities, 80% reported having one or two such students, 12% indicated three and 8% reported four or more, The relatively low levels of these challenges in the classroom, the diverse causes, and the widespread dispersion among schools suggests that while sensitivity towards these issues can certainly be addressed by APAL, it will be difficult to create a package of interventions that effectively address them.

1. THE USAID/*APRENDER A LER* EARLY GRADE READING PROJECT

With funding from the United States Agency for International Development (USAID), and in response to USAID's "All Children Reading" Strategy, World Education is implementing the USAID/*Aprender a Ler* (APAL) – Learn to Read – project from 2012 to 2016. The primary focus of the project is to improve reading outcomes for students in grades 2-3 in over 1,000 urban and rural schools in the Zambezia and Nampula provinces of Mozambique through a two-fold approach: increasing the quality of reading instruction through production, distribution and training in the application of reading instruction materials and training and coaching methodologies; and increasing the quantity of reading instruction through strengthening school management and institutional capacity building at school,district, and provincial levels.

To achieve these goals, USAID/*Aprender a Ler* will 1) train over 5,000 teachers in early grade reading instruction and continuous learning assessment in reading and over 1,000 school-based directors in school management to increase classroom reading time; 2) develop high-quality reading and instructional materials for students and teachers; and 3) design and adapt summative and formative assessment instruments to measure improvements in reading ability, instruction quality, and school management skills.

Working together with local partner Universidade Politecnica and international partner, World Education will ensure strong collaboration with local education institutions at the national, provincial, and district levels to strengthen technical and management capacity. Working closely with local counterparts will contribute towards long-term program sustainability as World Education works to hand over project implementation to the Ministry of Education in Mozambique.

2. THE DESIGN OF THE IMPACT EVALUATION

The IE design requires that USAID/APAL independently apply the treatments to each group. Furthermore, the Control schools must not receive any of the interventions from either treatment group at any time during the 2013 and 2014 academic years. We recognize that these constraints are more difficult and rigorous than is commonly found in intervention projects but the integrity of the IE design depends upon these conditions being respected for the two year period in the 60 Control schools included in the sample.

The purpose of the Baseline Report is not to answer the evaluation questions but rather to describe the pre-implementation scenario and to provide USAID, the implementers, and the stakeholders (primarily MINED) with insights regarding obstacles to improving early grade reading skills in a specific set of schools. The mandate of the IE is not to evaluate the USAID/APAL project or assess its level of implementation but rather to determine the effects of the intervention on a well-defined indicator—early grade reading outcomes.

2.1 The Evaluation Questions

The overall evaluation question to be addressed by the IE is: *To what extent have* USAID/APAL *treatment interventions improved early grade reading outcomes for students in second and third grades in the target schools in the Nampula and Zambézia Provinces?* The four questions that the IE will address related to the impact of the "Medium" and "Full" intervention models are displayed on Table 1.

Table 1: Research questions and data required to answer them

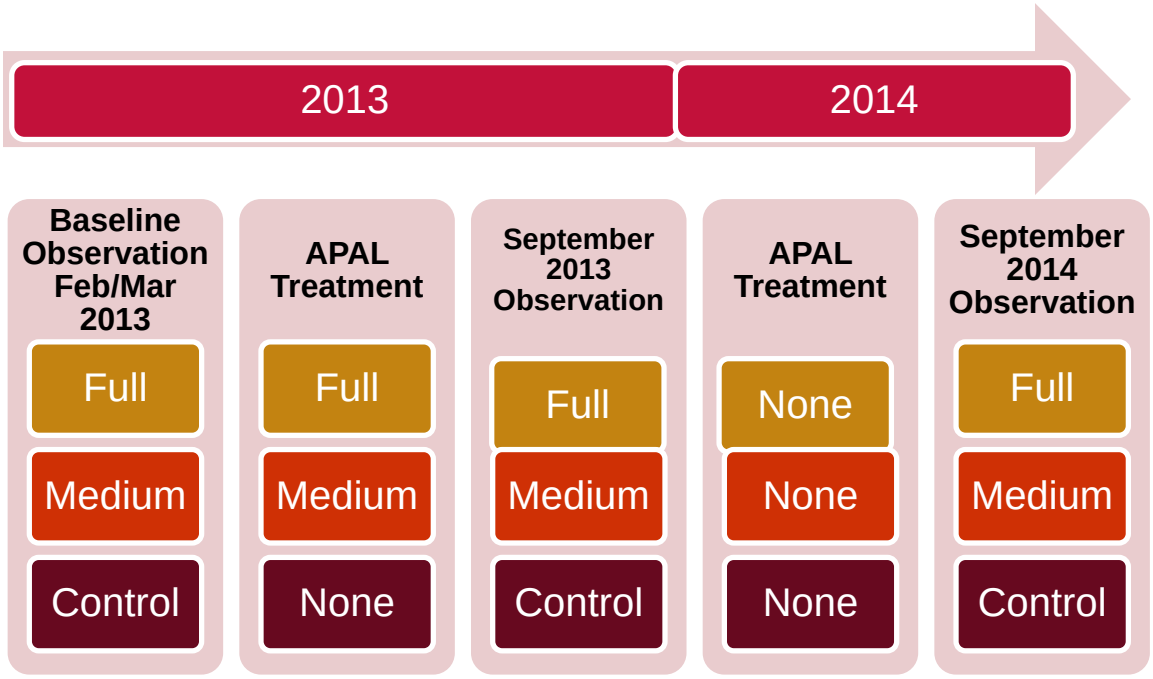
Evaluation Question	Data Type	Data Source
To what extent does the "reading instruction support" treatment intervention cause early grade reading outcomes to improve for students in grades two and three in target schools whose teachers have received training, coaching and support?	Reading scores on 7 subtasks of the EGRA instrument	Random sample of students in randomly selected grade 2 and 3 classes in all schools of randomly selected ZIPs from selected districts in Nampula and Zambézia provinces
To what extent does the treatment intervention of additional "school management" training, coaching and support to school directors cause a significant and additional improvement in early grade reading outcomes when coupled with "reading instruction support" in target schools?	Reading scores on 7 subtasks of the EGRA instrument	Random sample of students in randomly selected grade 2 and 3 classes in all schools of randomly selected ZIPs from selected districts in Nampula and Zambézia provinces
To what extent are the Medium and Full treatment interventions cost-effective?	To be addressed in October 2013 and 2014	*Note: Question to be addressed in October 2013 and 2014*
Of the most cost-effective interventions, which falls within the existing technical and financial management capacity of local education institutional personnel?	Semi-structured interviews	MINED officials at the central, provincial and district levels; directors of head schools of ZIPs; sample of community members

2.2 The RCT Design

The Impact Evaluation for the USAID/APAL calls for analysis of data collected in 180 schools along the economic corridors of Nampula and Zambézia provinces using Randomized Control Trial (RCT) methodology, with 60 schools receiving Full treatment, 60 schools receiving Medium treatment, and 60 no-treatment or Control schools. Working in close collaboration with USAID/Mozambique, WEI implementers of USAID/APAL, and national, provincial, and district MINED officials, IBTCI and its Mozambican partner, Global Surveys Corporation (GSC Research), are conducting the RCT in a sample of schools where USAID/APAL will intervene and implement one of the two treatments. During the initial year, USAID/APAL will be operational in 60 schools in Nampula and 60 schools in Zambézia. All intervention schools for 2013 have been included in the sample.

The general specification of the IE model is based on the evaluation objective of assessing the extent to which USAID/APAL interventions have improved early grade reading outcomes. Thus, the model treats early grade reading outcomes as a function of the Medium or Full interventions. The model will be tested with direct comparisons under three scenarios: with the Medium treatment sample, with the Full treatment sample and with the Control, or no-treatment sample. Figure 1 graphically depicts the Impact Evaluation RCT design.

Figure 1: The Impact Evaluation RCT design



The Randomized Control Trial (RCT) approach to the impact evaluation implies that participating entities will be randomly assigned to either a treatment (intervention) group or to a control group. In the specific case of USAID/APAL, it is desired to evaluate the impact of one of two treatments on the reading performance of students in second and third grade, relative to that of students in Control schools. The RCT allows for direct attribution of the APAL intervention to improved learning outcomes because the model controls for other possible determinants of the outcome. The reading scores obtained pre- and post-intervention will be compared to results obtained by the non-intervention Control group and provide an understanding of how students in the Medium or Full treatment groups would have performed without the benefit of the intervention. The evaluation design called for the establishing of a baseline (February/March of the 2013 school year) in all RCT groups, prior to the initiation of any intervention. Data on reading outcomes and other variables will be collected at the end of the 2013 and 2014 school years (September/October) at all schools from all groups selected for the baseline.

The 2014 end-of-school-year assessment will measure reading outcomes one year after USAID/APAL has finished its direct interventions in the two treatment groups. It is of interest to the IE to determine whether any effects of the intervention on reading outcomes in the two grades increase, decrease or remain the same as at the conclusion of the 2013 direct intervention. These types of comparisons will help to better understand the sustainability of impacts observed, as well as contextualize the benefits of USAID/APAL and therefore impact cost-effectiveness estimations.

Regardless of the benefits of any intervention, if the human, technical, and financial resources are not available to continue to train teachers and to produce the necessary materials for teachers and students, and to support school directors, sustainability is compromised. The IE team conducted qualitative interviews to document the existing financial and technical resources at the province, the district, and at the head school of the ZIP that could potentially be mobilized to absorb the added burden and costs of the intervention. During the second data collection round in September 2013, when the USAID/APAL activities have already been implemented, the IE will return to the same data sources to collect data on

resources available vis-à-vis the specific activities that are part of the intervention. Evaluation question # 3 related to cost-effectiveness will be addressed once the cost of each intervention has been determined.

2.3 The Sampling Strategy

The sampling strategy was developed by the IE data specialist from an examination of MINED 2012 data available of schools and school clusters in the districts targeted within the economic corridors of each province. Because of the ZIP-centric implementation model, ZIPs within these districts were then identified and randomly assigned to treatment or control groups in a manner that ensured a range of similar ZIP "sizes" (second grade total enrollment) in each group. Additional field-applicable methods were then developed (Annex B) to allow random selection of classrooms for each grade (where more than one classroom existed for a given grade in a school) as well as the random selection of individual students among those present on the day of the visit. The use of these techniques avoids selection bias from the field data collection teams themselves or from MINED personnel at the province, district, ZIP or school and classroom levels.

Ideally, the sample sizes in an impact evaluation are based upon power calculations, which allow the specification of the magnitude of changes that can be expected to be detectable with a particular degree of precision. In the USAID/APAL IE, the sample sizes were both predetermined (60 schools for each of the three RCT groups across the two provinces) and limited by the number of intervention schools where USAID/APAL would initiate activities during 2013 (60 for each of the two treatments across the two provinces). These restrictions nonetheless resulted in EGRA sample sizes of approximately 600 students (300 per grade) for each RCT group in each of the provinces, a number judged to be more than ample in EGRA applications in general. The design conserves the fundamental aspect of the RCT methodology: prior random assignment of the RCT groups (Medium, Full or Control) to equivalent ZIPs and schools belonging to each ZIP. Field-determined random selection of classrooms and students ensures that bias is eliminated in choosing the sources of interview, observational or EGRA testing data.

In consultation with MINED, USAID/APAL selected four districts in Nampula province where implementation would be possible during 2013. Table 2 displays selected basic demographic characteristics from the 2007 census for these districts.

Table 2: Demographic characteristics of the districts selected in Nampula

Nampula District/Municipality	Population	Population 5+ Years of Age	% Speak Portuguese	% Cannot Read or Write
Murrupula	140,311	109,927	29.6%	72.4%
Monapo	304,060	242,795	30.9%	71.0%
Rapale	203,733	163,238	38.1%	66.8%
Nampula Cidade	471,717	391,898	83.1%	40.9%
TOTAL	1.119,821	907,858	-	-

Source: 2007 Census, Instituto Nacional de Estadística (INE). Note: According to the INE, the Nampula province illiteracy rate is 60.9%.

Three districts were selected in the province of Zambézia for 2013 and comparable demographic information is displayed in Table 3.

Table 3: Demographic characteristics of the districts selected in Zambézia

Zambézia District/Municipality	Total Population	Population 5+ Years of Age	% Speak Portuguese	% Cannot Read or Write
Cidade de Quelimane	193,343	165,593	93.4%	2.8%
Mocuba	300,628	236,524	61.5%	5.8%
Nicoadala	231,850	188,088	59.8%	58.2%
TOTAL	725,821	590,205	-	-

Source: 2007 Census, Instituto Nacional de Estadística (INE). Note: According to the INE, the Zambézia province illiteracy rate is 62.5%.

2.4 Instrument Development

Why Test Early Grade Reading? National curricula typically specify that students should learn to read in grades 1-2, but many studies have found extensive illiteracy in low-income countries, even in advanced grades. The ability to read and understand a simple text is one of the most fundamental skills a child can learn. Without basic literacy, there is little chance that a child can escape the intergenerational cycle of poverty. Yet, in many developing countries, students enrolled in school for as many as six years are unable to read and understand a simple text. Acquiring literacy skills becomes more difficult as students grow older, increasing the gap between early readers and non-readers. Children who do not learn to read in the early grades are more likely to repeat and eventually drop out. Early assessment of the pre-reading and foundational skills required for fluency allows governments to implement measures to correct deficiencies where they exist.[10]

The primary cause of students not learning to read in the early grades seems to be insufficient amounts of instruction and practice. One USAID-financed study measured both reading fluency and the amount of instructional time spent on reading in samples of 20-30 schools in Nepal, Honduras, Guatemala and Ethiopia.[11] There was much variation among schools but overall fewer than half of the school days were spent in instruction, in-class time-on-task averaged 54 percent of total available time and less than 12 percent of the time was spent actually reading anything, even from the blackboard.

Reading fluency is related to students' performance throughout their school years, so acquiring this skill early is extremely important. When children are learning to read in languages using an alphabet or a syllabary, they must learn the letters and their forms, learn the sounds associated with each letter, and apply this knowledge to decode new words that they can recognize instantly. By the end of this phase, children develop sufficient speed and accuracy in decoding and word recognition that they can read with fluency. Reading with fluency is critical for reading comprehension, as children can concentrate on the meaning of what they read rather than focus on decoding. Recent evidence indicates that learning to read both early and at a sufficient rate, with comprehension, is essential for learning to read well. A substantial body of research documents the fact that children can and do learn to read by the end of grade 2, and indeed need to be able to read to be successful in school.[12]

The tools used to collect data were tailored to Mozambique. USAID/APAL adapted the EGRA and the

[10] Abadzi, Helen. (2009). "Instructional Time Loss in Developing Countries: Concepts, Measurement, and Implications." World Bank Research Observer. 24 (2): 267-290

[11] EQUIP2. (working Paper) "Using Opportunity to Learn and Early Grade Fluency to Measure School Effectiveness in Ethiopia, Guatemala, Honduras, and Nepal."

[12] Abazi, H. (2011). Reading Fluency Measurements in EFA Partner Countries: Outcomes and Improvement Prospects, EFA, FTI Secretariat Update September 2011.

SMA from a Portuguese version previously tested and administered in Cabo Delgado, Mozambique by the Aga Khan Foundation.[13] The SMA, was reviewed in light of the pretesting experience and later again revised by the IE team with input from USAID. The instrument was streamlined, phrasing of questions that led to misunderstandings was clarified and problematic questions were modified or removed. Additional instruments were developed by the APAL/IE team: (i) a Student Interview Protocol, which focused on the characteristics of the students tested and served as a cover page for the EGRA; (ii) a Teacher Interview Protocol; (iii) a Self-efficacy Survey administered to teachers of the students tested; and, (iv) semi-structured interview protocols to gather data from MINED officials, school directors, ZIP coordinators, pedagogical directors and PTA members. The qualitative interviews were designed to provide insights into the potential for the sustainability of the USAID/APAL activities.

Structure and Content of the Final EGRA for Mozambique. The EGRA is an individually administered instrument developed to measure a child's initial reading skills essential to becoming fluent readers who comprehend what they read. Because of its direct links with the skills indispensable for successful reading achievement, EGRA may inform ministries of education, donors, teachers, and parents about primary school students' initial reading skills and assist education systems in setting standards and planning curricula to best meet children's needs in learning to read. The full EGRA battery has eight subtests and is designed to take about 15 minutes per child.[14] The EGRA Mozambique assessed students' knowledge of oral vocabulary, concepts about print, names of letters, decoding skills, oral reading fluency, and comprehension of written text. EGRA Mozambique, as adapted by APAL/WEI, used five subtasks. Two of them—Oral Language and Oral Reading with Understanding—included two skills that were measured and scored separately.

Not all EGRA subtests were used in EGRA Mozambique (e.g., Subtest Invented Language was discarded) and test administration took approximately 15 minutes per child. The EGRA assessment was supplemented by student interviews administered prior to the assessment. The purpose of the interview questions was to clarify the demographic and social contexts that might explain students' reading performance. Assessors were instructed to use the local language for introducing themselves or giving instructions if they noticed that the student did not possess mastery of Portuguese. The actual items tested were in Portuguese, the language of instruction in Mozambique public schools.

The EGRA Mozambique consisted of five subtasks:

- ***Section 1: Oral Vocabulary***. The first part of this subtask includes eight prompts that required students to perform an action (e.g., "*show me your arm*") to determine their level of understanding of basic Portuguese oral vocabulary. A second part of the Oral Vocabulary subtask, with a maximum score of 6, requests that students follow instructions given orally (e.g., "*place the pencil on the paper"*).
- ***Section 2: Concepts about print* (CAP)** measures children's emergent literacy skills by asking them to demonstrate how they get ready to read a book and assessing whether they can recognize the front and back covers of a book, the direction in which to read, what the title of the story is, where page numbers are located, etc. In this subtask, the assessor used a book in order to determine the students' facility in handling printed material. The maximum score was 10.
- ***Section 3: Letter name knowledge***. This was a subtask in which students were shown 10 rows of

[13] Adelman, E., A.M. Moore, S, Manji. (2011) USAID EQUIP2. 2011. Using Opportunity to Learn and Early Grade Fluency to Measure School Effectiveness in Mozambique

[14] The Early Grade Reading Assessment (EGRA) is a 15-minute test administered orally to students in the early grades of primary school. EGRA evaluates students' foundation literacy skills, including pre-reading skills like phonemic awareness and listening comprehension, which have been shown to predict later reading abilities. Research Triangle Institute (RTI), www.rti.org

10 random letters (in both uppercase and lowercase) and asked to name as many letters as they could within one minute, yielding a score of correct letters per minute. The letters were presented to the student in either block or cursive formats (one type on each side of a large plasticized card) as familiarity with the two formats was found to vary during field-testing of the instrument.

- ***Section 4: Word reading***. This was designed to be a timed task that assessed students' fluency in reading words aloud. To facilitate recognition, a large plasticized chart of 30 words of 1 – 3 syllables was presented to the student. The subtask score is the number of words read correctly in one minute.
- ***Section 5: Oral passage reading***. This subtask assessed students' fluency in reading a passage of grade-level text aloud and their ability to understand what they had read. There were two parts to this subtask: The first required that students read two 120 and 135-word stories, respectively, aloud for one minute. The oral reading fluency score was the number of correct words read during that period. After students read a certain number of words (e.g., 13 needed to answer the first question) they were orally asked a question that required them to recall basic facts from the passage. The reading comprehension score was the number of correct answers, with a maximum possible score of 4 per passage or a total of 8 for the two passages.

EGRA administration includes a "stop" rule, which requires assessors to discontinue the administration of a subtask if a student is unable to respond correctly to any of the items in the first line (i.e., the first 10 letters, the first five words, or the first line of the oral reading fluency story). This rule avoids frustrating students who lack the skills to respond. If a sub-task needs to be discontinued, the EGRA assessor marks a box indicating that the subtask was discontinued because the child had no correct answers. Annex C includes a copy of the EGRA.

The School Management Assessment (SMA)[15]

The SMA yields a multifaceted picture of school conditions and management practice. Data collected by the SMA include: school infrastructure, classroom conditions, pedagogical approach; time on task; interactions among students, teachers, administrators, district officials, and parents; record keeping; discipline; and availability of pedagogical materials. By collecting information on the most crucial school effectiveness factors, the SMA is able to produce a rich data set designed to let school, district, provincial, or national administrators or donors learn what is going on in their schools and classrooms. (Annex D includes a copy of the SMA). The SMA instrument includes the following sections:

- **School Characteristics**— Infrastructure, type of building, class size, seating arrangements, facilities, etc.;
- **Class Observation**— The classroom observation required that enumerators spend 40 minutes in each class to conduct 7 observations of one minute each with intervals of five minutes. The objective was to record how many students were involved in various activities proposed by the teacher—e.g., silent reading, reading aloud, handling books, writing—and the number of those who were not participating in those activities. The observer should also note the classroom physical ambiance and conducts an inventory of the materials present in the classroom. A brief manual was developed by the IE team to accompany the SMA and clarify the various issues that could present a problem to enumerators.
- **Teacher Questionnaire**—This section of the SMA collected information on teacher demographics and professional preparation;

[15] The adaptation and field testing of the instrument, the training of the enumerators, the administration of the instrument, and the analysis of the information were under the responsibility of the USAID/APAL implementer (WEI), while the IE team was responsible for providing overall supervision and maintaining the quality of the data collected. Results obtained by the analysis of the SMA, as provided by WEI, are included in this baseline report.

- **School Director Questionnaire**—Included in this section were questions related to the demographics and the professional characteristics of the school director as well as questions related to the director's management skills.

The Self-efficacy Survey. Within social cognitive theory (Bandura, 1982; 1997) self-efficacy is a form of self-evaluation that influences decisions about the amount of effort and persistence put forth when faced with obstacles. Self-efficacy is not a measure of skill. Rather, it reflects what individuals believe they can do with the skills they possess. There is value in measuring teachers' self-perceived competence to deal with the challenges of teaching reading in the initial classes. These data could be helpful in assessing teacher readiness to incorporate new behaviors. It could also inform USAID/APAL regarding how to better educate teachers regarding the different types of strategies that could be integrated into the teaching and learning process. A copy of the Self-efficacy Survey is included in Annex E.

The Protocols for qualitative interviews. Semi-structured protocol interviews were developed to accommodate the different perspectives of MINED officials at the central, provincial, and district levels as well as the perspectives of directors of "head" schools of the ZIPs, ZIP coordinators, pedagogical directors, school directors and PTA members (Annex F).

The main purpose of collecting analyzing qualitative data is to identify common themes or ideas that could be part of a narrative. To provide meaningful findings, the IE developed a set of clear and specific questions focused on four elements related to sustainability: human, technical, financial, and technological resources. These are aspects that need to be examined in order to assess the capability of MINED to sustain USAID/APAL activities that show promise. Findings from interviews can easily be traced back to individual informants and *verbatim* quotations were included to add texture to the narratives.[16]

2.5 Limitations of the Study

A Baseline study can be defined as an analysis of current situation to identify the starting point for an activity, a program or a project. As with all research projects there is the need to tell the reader the factors that limit the study and to what extent the findings can be generalized. These factors could be related to the methodology or to the data collection and analysis. In retrospect, there are a few, specific limitations of this study which should be addressed as a means for improvement or potential strategies for further studies to be conducted in September/October 2013 and 2014.

The first limitation is related to the design of the IE and of the data collected. By obtaining the data only in a set of ZIPs along the economic corridors of Nampula and Zambézia, at a pre-defined number of schools (180), and in schools that could be accessible to the implementer, bias was introduced with respect to proximity to roads. This is not critical to the overall accuracy assessment since ZIPs were randomly assigned to the "Full", "Medium", and "Control" groups and all schools within a ZIP were assigned to a specific treatment group, however, it is important to mention.

Obviously, the characteristics of the schools are not the same where the impacts of roads are not felt. Also, by design, the data collection took place in the first months of the school year. This has at least two implications: (i) no direct correlation can be made between teacher characteristics (as recorded by the teacher EGRA cover sheet) and teacher performance (as recorded by the SMA) and student outcomes as it is possible to do when data are collected at the end of an instructional year and (ii) even

[16] APAL is conducting an Institutional Gap Analysis (IGA) to assess the extent to which the necessary conditions for sustainability of APAL activities exist at the various levels of the MINED. Exchange of information is planned between APAL and the IE.

though we refer to second and third grade reading performance it must be remembered that the participating students were only beginning the second and the third grades and, therefore, their reading ability level is closer to what is expected from first and second graders.

A second limitation is related to the instrumentation and the training of the enumerators. The SMA was not ready for administration prior to the conduct of the training of the enumerators and there was the need to revise and modify the instrument as the training was underway. Possible implications include the lack of familiarity of the enumerators with the instrument, especially with the Class Observation section. Overall, the data collected by the SMA did not correspond to the expectations, especially in what relates to class observation. To integrate all facets of school organization and classroom observation that would help explain student scores at mid- and end-line a more in-depth revision of the instrument will be necessary.

The third and probably most important limitation associated with this study refers to the RCT methodology itself. The RCT methodology presupposes that there is a number of schools in each group, randomly allocated to which group they are in. However, as pointed out by Scriven (2008)[17] "…*any two such groups will always be distinguished by some factors, (e.g., location), or they would not be different groups. And these unavoidable distinguishing factors may be linked in an unexpected way to causally relevant differentiating factors such as local variations in weather or room temperature, or ambient noise level, or facilities management style, which then invalidate the inference to the experimental treatment as being the only possible cause of any outcome differences.* These factors only surface during the course of running an RCT.

In addition, Hawthorne-type effects are significant threats.[18] It may be worth recalling the experiment done in the early days of placebo studies that showed that the placebo effect works just fine *even if* the control group is told they are getting the placebo and are instructed and tested on their knowledge of exactly what this means.[19]

3. DATA COLLECTION

USAID/APAL, with the assistance from the IE team and IBTCI's local partner GSC, trained four supervisors and 24 enumerators to administer the EGRA, interview principals and teachers, conduct inventories of school and classroom resources, and observe reading lessons as part of the SMA survey. After a week-long training workshop in January/February 2013, conducted both in Nampula and in Zambézia, research teams composed of USAID/APAL supervisors and enumerators GSC staff members, and USAID/Mozambique representatives visited a total of 180 public primary schools. Data were collected via direct classroom and school observation and interviews with students, teachers, school directors, MINED district and provincial officers and PTA members.

In each school, one grade 2 and one grade 3 class was randomly selected for participation (if there was more than one class in a school) and 10 students from each selected class were again randomly selected to take the EGRA. Before administering the EGRA, assessors read to each student information about the test and how it would be used. The student was asked to provide oral assent to participate in the assessment.[20] The Student EGRA Cover sheet collected data on the students' age and sex, their home

[17] Scriven, M. (2008). *A Summative Evaluation of RCT Methodology and An Alternative Approach to Causal Research.* Journal of Multidisciplinary Evaluation, Vol 5, No 9, March 8. http//www.jmde.com

[18] The Hawthorne effect has been well established in the empirical literature beyond the original studies.

[19] Clark,R.E. & Sugrue,B.M. (1991) "Research on instructional media, 1978-1988" in G.J. Anglin (ed.) *Instructional technology: past, present, and future* ch.30 pp.327-343 (Libraries unlimited: Englewood, Colorado).

[20] Tests such as EGRA are exempt from application of the Common Rule for the Protection of Human Subjects, 22 CFR 225.101(b).)

environments, repetition history, school attendance, language spoken at home, and other variables.

The 341 teachers[21] were interviewed about their teaching experience and other matters related to the school using the structured protocol developed by the IE team and 333 teachers[22]—172 in Nampula and 161 in Zambézia—completed the Self-efficacy Survey. A brief description of the instrument and its purpose, and a pledge to confidentiality, was given to the teachers who were then asked to give their verbal consent to participate. Teachers were also informed that failure to participate would have no effect on their standing in the school. The instructions were read together by the survey administrator (one of the IE supervisors) and the informant and once the administrator was satisfied that the informant had understood the purpose of the instrument and how to respond to it, the informant was given five minutes to mark his or her responses.

An USAID/APAL enumerator completed the SMA Questionnaire, alternating between grades 2 and 3 on successive schools visited. The school director was interviewed and an enumerator observed the selected grade 2 or grade 3 classes. The focus of the observation was on students and the activities in which they were involved during the seven one—minute observation. The enumerators also took inventory of the school grounds and of the classrooms where the observation was conducted.

IE personnel used the semi-structured interview protocols to interview district, province and central MINED personnel, directors of head schools of ZIPs, and a sample of school directors and PTA members. All interviews were recorded and then sent, electronically, to the GSC main office for later analysis. Annex G includes a complete list of those interviewed, their function and the offices or schools where they work.

Data collection started on February 7, 2013 and was completed on March 13, 2013. A total of 3,598 students and 341teachers (51% female in Nampula; 68% female in Zambézia) participated in the assessments and interview. Table 4 presents the final baseline sample count of districts, ZIPs, schools, directors, teachers and students.

Table 4: Count of data sources by district

Province	Districts included	ZIPs	Number of schools	Total SMA administered	Grade 2 EGRA	Grade 3 EGRA
Nampula	Monapo	5	34	34	338	340
	Nampula Cidade	9	31	31	310	310
	Murrupula	4	22	22	220	220
	Rapale	1	7	7	70	70
Subtotal	**4**	**19**	**94**	**90**	**938**	**940**
Zambézia	Mocuba	7	38	37	380	380
	Nicoadala	7	43	43	430	430
	Quelimane	1	5	5	50	50
Subtotal	**3**	**15**	**86**	**85**	**860**	**860**
TOTAL	**7**	**34**	**180**	**175**	**1,798**	**1,800**

[21] 5% of the 360 teachers of these randomly selected classes were absent on the day of the visit.
[22] 8 teachers chose not to take the survey.

4. DATA PROCESSING

Data quality was checked jointly by APAL and IE supervisors on the field to avoid missing data or missing instruments. Once the data packages were checked by the IE supervisors, they were sent to GSC where data were cleaned and prepared for entry. Information included in each data set of the study (EGRA, Teacher Interview Protocol, Self-efficacy Surveys, Qualitative Interview instruments each had their own data set)[23] was checked for consistent responses. Checks were conducted both within each data set and among the data sets and inconsistent responses were edited only if it was clear which inconsistency was incorrect. All data were entered by March 22, 2013 and ready for analysis, nine days after the last day of the data collection.

The Self-efficacy survey included 20 closed-ended items and was administered face-to-face by IE supervisors, thus reducing the key challenge that surveys must overcome—non-response. The close-ended items obtained standardized responses, which facilitated data processing and production of basic summary statistics.

As the qualitative interview protocols for MINED personnel and school directors were being developed, the IE team explored issues related to data management and analysis of the information since good data collection practices require forethought about how the information is going to be stored and analyzed. Upon completing the data collection, the IE team reviewed the qualitative data carefully and identified statements related to sustainability. The question guide was structured and intentional and the expertise of the interviewers resulted in high quality data allowing complete and accurate transcripts to be produced. Once the transcripts were produced, statements were categorized under six labels: (i) perception of the reading ability of students in early grades; (ii) challenges related to improving reading in the early grades; (iii) availability of financial and human resources; (iv) capacity building necessary to improve early grade reading performance; (v) availability of educational materials to improve early grade reading ability; and (vi) perception of the ZIPs as a strategy to improve school conditions and student performance. The inductive approach utilized allowed findings to emerge from the frequent and significant themes found in the narratives and not let findings be constrained by structured models, frameworks or other preconceived paradigms.

5. FINDINGS

The findings included in this section are not generalizable to districts, ZIPs or schools in Nampula and Zambézia that did not participate in the study. Nor are they representative of the early reading level of all Mozambican students at the beginning of the second and third grades.

5.1 School and Classroom Characteristics

This section discusses the results of the SMA, as reported by USAID/APAL, in order to put the EGRA results into context.

School Infrastructure. School infrastructure serves as an indicator of resource allocations across schools and as an indicator of school management. In 69% of the 360 classes visited in Nampula and 81% of those in Zambézia, students were sitting on the floor on bare ground for lack of benches or chairs. When reflecting about school infrastructure, it should be noted that the 180 schools visited in the two provinces are located in the relatively more prosperous areas of the provinces—accessible schools in

[23] The entering and processing of the SMA data as well as the analyses was conducted by USAID/APAL.

the economic corridors. In Nampula, 15% (14 out of 90) of the classes selected for participation were conducted with the students sitting under a tree for lack of classrooms.

Enrollment, Class Size, Class Composition, and Attendance. The average enrollment in the 360 classes observed was 57 students in Nampula and 54 in Zambézia with an average attendance rate on the day of observation of 70% in Nampula and 64% in Zambézia. In both provinces, the average attendance rate is higher in second than in third grade. Table 5 details the class enrollment and attendance of the 360 classes which form part of the sample.

Table 5: Enrollment and attendance in the classes sampled

Indicators	Second Grade		Third grade	
	Nampula	Zambézia	Nampula	Zambézia
Average class enrollment	55	58	54	54
Average attendance	38	37	37	35
Attendance rate	69.8%	65.3%	68.9%	61.0%*
% girls enrolled in the class	49.6%	49.5%	48.3%	47.8%
% girls present that day	50.5%	50.8%	47.5%	46.4%
Attendance rate girls	71.0%	67.6%	68.6%	59.1%*
Attendance rate boys	69.5%	63.9%	70.1%	63.7%

* Significant p <0.01

Average class sizes in classrooms where the EGRA was administered ranged from 54 in third grade in both provinces to 58 in second grade in Zambézia. Average actual attendance ranged from 35-38 students, yielding classroom attendance rates averaging between 61% (Zambézia third grade, significantly lower than the 69% of their Nampula counterparts) and 70% (Nampula, second grade). Enrollments and attendance of girls and boys were examined to detect gender differences: in second grade girls formed very slightly under one-half of enrollment, and slightly over one-half of those present, also reflected in the somewhat better attendance rates of girls. In third grade, girls' participation in both classroom enrollment and attendance shows signs of erosion, falling to as low as 46% of students present in Zambézia. The attendance rates of girls in both provinces is lower in third grade than that of boys, and Zambézia lags behind Nampula, significantly so for girls (59% in Zambézia, 69% in Nampula).

Length of the school day and of the class time. Even when good teaching techniques are applied, students cannot succeed if they are not given sufficient learning time at school and do not practice specific skills in the classroom and thus time-on-task is an important indicator when determining school effectiveness As it relates to reading instruction it includes such activities as listening to the teacher reading a story, answering questions about the story, working on names of letters, decoding words, reading silently or engaging in group reading activities. Students are off-task if they are interacting socially or are otherwise disengaged.[24] Observation in the randomly selected second and third grade classes showed that 37% of the students in Nampula and 61% in Zambezia were not engaged in any of the activities proposed by the teacher – reading aloud, silent reading or writing.

Late arrival and absenteeism of Teachers and School Directors. In the 175 classrooms where the SMA was administered, over one-half of the teachers arrived late: 51% in Nampula and 58% in Zambézia and average teacher tardiness rates ranged from 5 to 40 minutes. That, and the fact that

[24] At the time when this report was prepared, the results provided to the IE by USAID/APAL did not allow the determination of how much class time was spent practicing reading skills.

students were engaged in cleaning tasks, delayed the start of classes. The impact of late arrival is particularly strong in cases where the length of the school day or shift is short. The late arrival of teachers undermines student learning time and consequently is associated with lower performance (Glewwe, Hanushek, Humpage and Ravina, 2011). The pattern of tardiness becomes even more serious when one considers that, as recorded by the SMA, in the 180 schools visited 53% of the directors in Nampula and 68% of those in Zambézia also arrived late.

Teacher absenteeism has been shown to be a major factor in school ineffectiveness and low student performance (Abadzi, 2007; Fehrler, Michaelowa and Wechtler, 2009). Surveys in several countries show that schools are routinely missing a quarter of their staff, with rural schools faring even worse.[25] In the participating classes, it was observed that six teachers were absent on the day of the visit and between 30 and 60% of the 341 teachers interviewed reported having missed at least one day since the beginning of the school year. It is important to note that we are not referring to all teachers at the school but only to those who teach second and third grade classes which were randomly selected for participation in the baseline study. It is possible that the pattern of absenteeism and tardiness is the same for teachers of other grades and continues throughout the year, thus reducing instructional time for all students.

Availability of Learning/Reading Materials in the Classroom. In the 175 classrooms for which SMA data are available, 76% of the Nampula students and 97% of those in Zambézia had the official MINED reading book, which is distributed for free to every first and second grader. The significant difference between the two provinces could be a result of the fact that the data collection was conducted at the beginning of the school year (February/March) and not all books distributed by the MINED had arrived at the schools. It was observed that 88% of the students had a notebook and 84% had a pencil or a pen. Very few students had any other reading material—6% in Nampula and 11% in Zambézia. Apart from blackboards and chalk, universally found in all classrooms—the great majority of the classrooms had at least one blackboard eraser (79% in Nampula and 90% in Zambézia)—no other material such as glue, scissors, games, paper or similar items was observed. Newspapers, magazines or other printed material were not available.

The walls of the classroom (if there was a structure) were bare—no newspapers, no pictures, no letters or words—and there was no evidence of work prepared by the students on the walls or anywhere else in the classroom. This contrasts sharply with the perception of teachers regarding their ability to make the classroom a place conducive to reading: 87% of the 333 teachers who took the Self-efficacy Survey considered that they were *well prepared* (30%) or *prepared* (57%) to do so.

Observation of the reading lesson and teacher—student interaction. Rather than relying on self-reporting by teachers regarding activities conducted in the classroom and time on-task, there is the need to make direct observations of teachers in action. In spite of being costly and time consuming, class observation is the only way to record what goes on in the classroom. Because of the complexity of the observation method selected, the classroom observations did not provide the crucial insights into how lesson time is spent but the information available suggests that little time is spent on teaching basic reading skills and much time is spent on choral repetition.

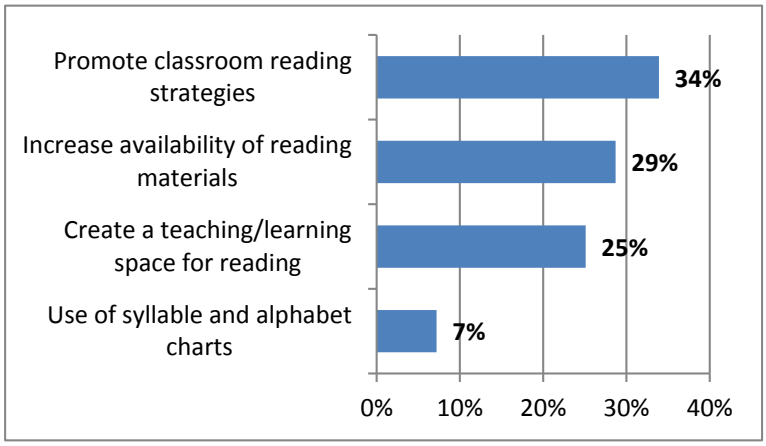
School directors, Pedagogical Oversight and Administrative Support. School principals play a critical role in the quality of education. Their behaviors are believed to be central to the creation and facilitation of an effective teaching and learning environment within a school. The 171 school directors interviewed are predominantly male—80% in both provinces and 60% are 46 years old or more. In

[25] Abadzi, Helen. 2007. *Absenteeism and Beyond: Instructional Time Loss and Consequences*, World Bank Policy Research Working Paper No. 4376, p. v.

Zambézia, a considerable group of directors (52%) has no more than 6 or 7 grades of education plus 3 years of teacher training or 9 or 10 grades plus two years of additional training; in Nampula, 47% have the same professional profile. In contrast, in Nampula, 27% of directors are at the *Bacharel* or *Licenciado* levels[26], compared to only 6% in Zambézia. Forty percent of the school directors have fewer than five years of experience as directors; 80% have been in this function at their current school for less than five years.

The information provided in the last paragraph must be weighed against the fact that 20% of directors in Nampula and 36% in Zambézia have not participated in any type of school management training in the past five years. One sensitive point is the fact that 57% of the schools directors also teach a class, which undoubtedly cuts down on their school management time. The lack of the instructional leadership which is expected to be provided by a school director becomes clear by the self-reported information provided by directors regarding measures taken to improve reading at the school. The information displayed on Figure 2 suggests that little assistance is being provided to teachers.

Figure 2: Do school directors promote early literacy?



5.2 Student Characteristics

The EGRA Student Cover Sheet was administered to a total of 3,598 students—1,878 in Nampula and 1,720 in Zambézia—selected for EGRA administration. The higher attendance rates of girls in second grade noted earlier and the lower rates in third grade, especially in Zambézia, led to differences in the composition of the student sample as shown below.

EGRA Administration by Sex	Grade 2	Grade 3
Female	54.0%	46.4%
Male	46.0%	53.6%

The EGRA cover sheet provides an insight into the patterns of enrollment by grade by age in the selected classrooms. The table below breaks out reported student age (at the beginning of the school year) by grade and province. Fully one-half of Nampula second grade students did not know their age (36% in Zambézia): the percentages listed in the body of the table are of students who did. It appears that more second grade Nampula students are younger (35% were aged 7 years or less) than in Zambézia (28%), but what is much more striking is that, overall, 50% of students are 9 years or older in second grade, and 61% of third grade students are 10 years or older at the start of the school year.

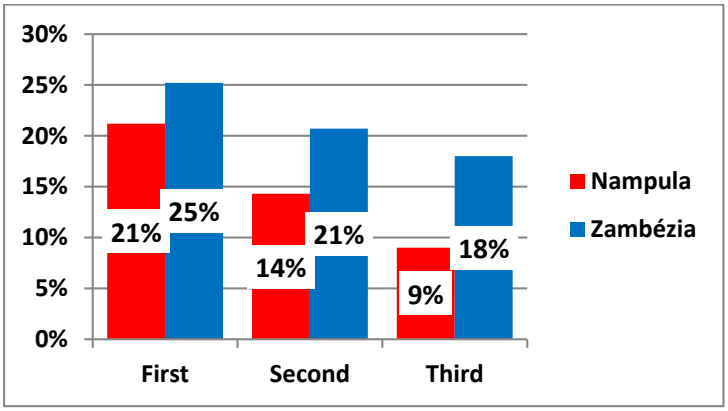
[26] *Bacharel* is equivalent to a B.A. and *Licenciado* is equivalent to a B.A. plus one more year of teaching practice.

Table 6: How old are the students assessed?

Age of Student Assessed	Second Grade		Third Grade		Age of Student Assessed
	Nampula	Zambézia	Nampula	Zambézia	
7 years or younger	35.4%	27.8%	22.8%	20.2%	8 years or younger
8 years	14.7%	21.8%	15.2%	19.8%	9 years
9 years	10.8%	14.2%	16.9%	25.3%	10 years
10 years	18.5%	20.9%	11.5%	14.0%	11 years
11 years or more	20.7%	15.3%	33.5%	20.8%	12 years or more
Don't know	51.6%	36.0%	24.3%	17.8%	Don't know

Assessed students were asked whether they had repeated the current grade or any of the prior grades. Figure 3 shows the self-reported responses. The cumulative effects of higher repetition in each grade in Zambézia could easily explain the 34% of third grade students aged 12 or more, compared to 21% in Nampula. But repetition (and the resulting overage) is a crucial issue because of its association with leaving school early.

Figure 3: Self-reported repetition among students sampled



Students who are over-age for their grade—due to late entry or repetition—have been shown to be at greater risk of leaving school early.[27] The UNESCO UIS reports that in Mozambique 18% of students who are over-aged by 1 - 2 years leave school early while those whose over-age is 3 years or more have a 48% of probability of leaving school early. Repetition and dropout are also key factors affecting the internal efficiency of education systems, and repetition stresses the system capacity because of classroom overloading.

Another variable that could be related to performance on reading assessments is language of instruction versus language(s) spoken at home. Thus, students were asked whether Portuguese was spoken at home and with whom they spoke it. Table 7 below displays the responses given by the students sampled, indicating the relative frequency of use.

[27] Global Education Digest 2012. Opportunities Lost: The impact of grade repetition and early school leaving. UNESCO Institute of Statistics. Montreal, Canada, www.uis.unesco.org

Table 7: How often is Portuguese spoken outside of school and with whom?

Person	Province	Never	Almost never	Occasionally	Almost always
Mother	Nampula	40%	22%	12%	26%
	Zambézia	18%	16%	19%	47%
Father	Nampula	38%	19%	15%	29%
	Zambézia	17%	14%	22%	47%
Siblings	Nampula	40%	20%	12%	28%
	Zambézia	16%	14%	19%	52%
Friends	Nampula	38%	19%	10%	33%
	Zambézia	13%	15%	18%	54%

The use of Portuguese at home is clearly more prevalent in Zambézia and the two provinces are very significantly different in this aspect, more so than any other variable studied. A secondary trend of some interest is that, in both provinces, students tend to speak Portuguese more often with the younger members of their group—siblings and friends—rather than with parents. The low use of Portuguese at home could be seen as one of the factors behind Nampula's poor results on the EGRA, especially in second grade. Figure 4 underscores the dramatic difference between the provinces by showing the percentage of family members and friends with whom Portuguese is "never" or "almost never" spoken.

In addition, students were asked how often they read at home, and also how often someone at home read to them in Portuguese. Between 34% and 40% reported that someone at home did read to them: as the following table shows, siblings were much more often identified as the person that read to them than fathers or mothers, especially in Zambézia. Independently of grade or province, only 28% of students reported reading at home as shown in Table 8.

Figure 4: Non-use of Portuguese by province

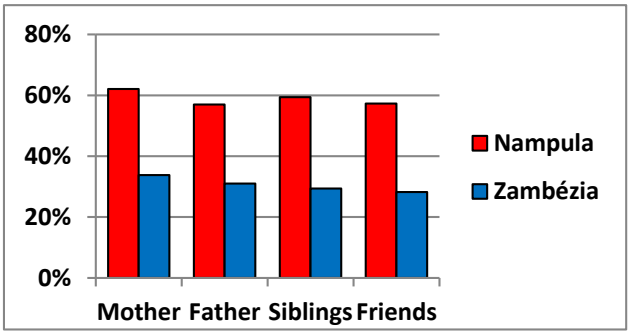


Table 8: Who reads to the child at home?

Who reads to student	Second Grade		Third Grade	
	Nampula	Zambézia	Nampula	Zambézia
Father	31%	27%	25%	25%
Mother	22%	18%	21%	15%
Sibling	34%	41%	39%	48%
Grandparent, Aunt or Uncle	13%	14%	16%	12%

Nearly all children (90-94%) said that they helped at home with tasks they considered to be work (no gender differences were noted); somewhat over 30% reported doing work outside of the house (again with no gender difference, see Annex I), most often helping in agricultural tasks and market activities. Despite the interviews being conducted shortly after the school year had begun, the students self-reported having missed school, significantly more often in Zambézia (around 23%) than in Nampula (13%). Two-thirds of those who had missed school indicated that the absence was due to illness (their own or that of a family member); 16-19% said the cause was due to work at or outside the home. Arriving late at school was again much higher in Zambézia (around 26%) than in Nampula (15%): in this case, work at or outside the home was cited by 52-60% of respondents, distance to the school by 26-30%.

Key socio-demographic and EGRA results were examined by sex (Annex I). No age differences by sex were observed in either province in second grade students, but among the third grade students, boys were significantly older than girls (0.58 years older in Nampula, 0.40 years older in Zambézia). There were no self-reported repetition differences between boys and girls for either Grade 1, 2 or 3, so the cause of this is unclear, although it may have at least part of its roots in the observed tendency for somewhat more girls to start second grade at seven or eight years of age in comparison with boys.

In Nampula, girls more frequently speak Portuguese at home with their mother (29% of girls reported "almost always"; boys 23%) or their father (31% girls; 26% boys). In Zambézia, the opposite was true: boys more often "almost always" spoke Portuguese with their mother (girls, 44%; boys, 50%) or their father (girls, 45%; boys, 49%). No significant differences between the sexes were observed for reading at home, working outside of the home or self-reported missing of school.

Some, but not all, of the student socio-demographic information from the EGRA Cover Sheet showed variations between the RCT treatment and control groups, as detailed in Annex J. These variables (age, self-reported repetition of the current and prior grades, and frequency of Portuguese language use at home with the student's mother and father) were expected to be used to carry out statistical adjustments in the within-year 2013 comparison and the 2013-2014 comparison of EGRA results. There were no statistically significant differences among the groups for age in Nampula in either grade. In Zambézia, there are small but significant differences in age in the Full (8.5 years), Medium (8.8 years) and Control (9.0 years) in second grade, as well as in third grade (9.9, 10.3 and 9.9 years, respectively). Self-reported repetition of first grade was different among the groups in Nampula (Full, 18%; Medium, 23%; Control, 23%) and Zambézia (Full, 22%; Medium, 30%; Control, 24%). Differences in second grade repetition were observed only in Nampula (Full, 14%; Medium, 12%; Control, 17%). Repetition of third grade was not different between the groups in either province.

Language use at home also varied between the treatment and control groups in both provinces. In Nampula, speaking Portuguese with the mother "almost always" was reported to be 30% in students assigned to the Full treatment group, 30% in the Medium treatment group and 18% in the Control group. In Zambézia, these figures were 54%, 39% and 48%, respectively. Portuguese language use with fathers was virtually identical to these results in both provinces. Reading at home, working outside of the home and missing school did not vary by RCT group.

Out of the 3,598 second and third grade students, 11% reported not living with their mother and 18% not living with the father (Annex H). When asked the reason why they did not live with their mother and/or their father, orphanage was reported as the reason by 428 children out of the 817 who reported not living with mother and/or father. The majority of other reasons given (48%) were related to separation, divorce or relocation for work demonstrating that there are multiple causes for the absence of one or more parents.

Of the 341 teachers of students assessed, 218 (64%) indicated that they had no students with disabilities. Of the 123 teacvhers who did report students with disabilities, 70 (57%) indicated that they have one student with a deficiency; 28 (23%) said they had two; 15 (12%) said they had three; and 10 (8%) indicated that they have four or more. The relatively low levels of these challenges in the classroom, the diverse causes, and the widespread dispersion among schools suggests that while sensitivity towards these issues can certainly be addressed by APAL, it will be difficult to create a package of interventions that effectively address them.

5.3 Teacher Characteristics

The teacher interview protocol was applied to a total of 341 teachers of the second and third grades where the EGRA was administered—176 in Nampula and 165 in Zambézia. In Nampula, male and female teachers are split evenly while in Zambézia females accounted for 68% of teachers. Of those who teach the classes sampled, 47% percent are between the ages of 26 and 40 (19% 25 or less) and 50% have five years or less of teaching experience. Twenty-two per cent of all teachers had been at this school for one year or less.

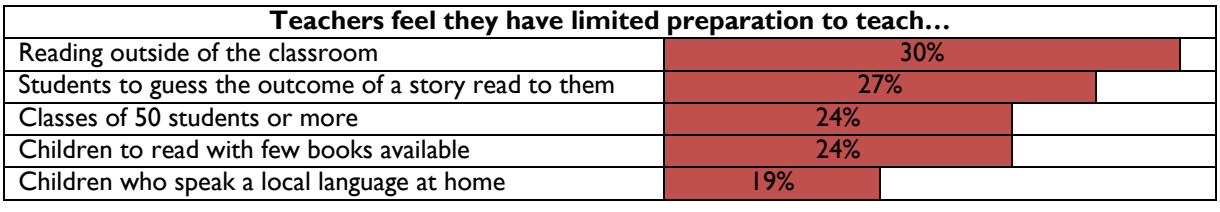
In terms of teacher preparation, 52% are a product of the 10+1 or 10+2 and 12+1 teacher preparation formats; 21% have 6, 7, 8 or 9 grades and one or more years of teacher training; 10% in Nampula and 17% in Zambézia admit having no formal preparation. These data, when viewed in relation to the scores their students obtained on the EGRA, suggest the need for closer scrutiny of the institutions where teacher preparation is being conducted. Compounding the problem, 70% report that they do have the Teacher's Guide and only 19% of the teachers sampled reported having received any in-service training during the prior 12 months. An inadequate level of teacher preparation combined with a lack of in-service training could be one of the factors that explain low student reading performance.

Self-Efficacy Survey Findings

General perceived self-efficacy pertains to optimistic beliefs about being able to cope with a variety of stressors. Brandura (1997) asserts, "*In contrast to other constructs of optimism, perceived self-efficacy explicitly refers to one's competence to deal with challenging encounters*."[28] A total of 330 second and third grade teachers from the 180 schools responded to the self-efficacy survey—172 from Nampula and 158 from Zambézia.

In contrast with the extremely low results obtained by student on the EGRA, teachers rated themselves very highly in terms of their preparation for teaching reading in the early grades. While 98% of the third grade students in Zambézia were not able to answer even one reading comprehension question, 94% of the teachers consider themselves well prepared or prepared to *"teach students to answer questions about a text."* In Nampula, 87% of the second grade students could not name one single letter and yet 93% of the teachers consider that they are *well prepared* or *prepared* "*to teach students to learn and memorize the names of letters."* The results of a chi-square test conducted to determine if there was a difference between teachers' level of perceived self-efficacy in two provinces and in the two grades was not significant. The complete tabulation of the twenty items is included in Annex E together with a copy of the Self-efficacy Survey. Figure 5 shows the few reading- related activities that teachers perceived to be challenging.

Figure 5: Activities that teachers perceive as challenging



[28] Bandura, A. (1997). Self-efficacy: The exercise of self-control. New York: W.H. Freeman and Company

By definition, self-efficacy is the perception of individual ability to perform and complete tasks. Bandura (1997) suggests that if individuals have no basis of the knowledge required to properly assess their ability and the difficulty of the task, their assessment will in the end be flawed. In essence, it is difficult for teachers to objectively evaluate themselves on topics for which they have little knowledge.

The implications for the implementation of the USAID/APAL teacher training activities are twofold. First, it is problematic to provide training to individuals who consider that they are already well prepared to perform a task and thus may not need training. Second, USAID/APAL will have to deal with pre-existing notions of what it means, for example, to *"make the classroom a space conducive to reading"* since 87% of the respondents considered themselves to be *well prepared* or *prepared* to do so. Classroom observation, however, as recorded on the SMA, shows that only in 1% of the Nampula classes observed and 2% of Zambézia classes was student work or any other prompt related to reading displayed on the walls.

5.4 EGRA Findings

The EGRA data were initially examined through descriptive frequencies. Given the significant differences between the two provinces, data were subsequently analyzed separately. The analyses in the following sections describe both the average performance and the distributions of scores for each sub-task for the assessed grade 2 and grade 3 students.

EGRA Results by Subtask

Subtask Oral Vocabulary (Parts of the body, e.g. "Show me your arm")

Table 9 that follows presents the number of correct answers provided to the first of the oral vocabulary subtests, which directed the student to identify eight different parts of their body, such as their arm, foot, head, etc. While the instructions were provided in the student's local language, the actual test items (i.e. "show me your head") were spoken in Portuguese. Few students got all eight items correct, but a very significant proportion of students in Nampula (36% in second grade and 26% in third grade) were unable to identify any or only one of them. Zambézia fared much better in both grades, suggesting that the language spoken at home issue is strongly affecting the oral comprehension of these students in the Portuguese language used at school.

T-tests were used to compare the mean scores between the provinces, but also to compare, within each province, the means of second versus third grade students. These, and the other results presented later for the other subtests, include all students, even if their scores were zero. At the baseline, we are interested in knowing the abilities of all the students in order to better understand the performance challenges that exist. In post-intervention comparisons, in some cases it may be productive to exclude from analyses those students who could not perform at all. But most of the baseline EGRA results, as we will see, suggest that the largest problem faced by USAID/APAL is non-response or non-performance. With the very high percentage of students in this category, all mean scores or subtests are very low. Thus, here we present distribution tables and graphical representations to highlight the problems and not merely focus on mean scores and statistical tests of significance.

Table 9: Understanding of Oral Language (*"Show me…"*)

Correct Answers "Show Me Parts of the Body"	Second Grade		Third Grade	
	Nampula	Zambézia	Nampula	Zambézia
0	16.9%	0.7%	10.8%	0.3%
1	19.1%	5.0%	15.1%	1.4%
2	12.6%	13.5%	14.6%	11.0%
3	16.0%	31.5%	16.1%	31.6%
4	15.0%	26.8%	17.2%	29.0%
5	11.3%	14.1%	13.9%	16.0%
6	5.5%	6.4%	6.2%	8.0%
7	1.7%	1.7%	2.5%	2.2%
8	1.9%	0.2%	3.7%	0.3%
		Sig.		Sig.
	Comparison between Provinces, Same Grade			
N	940	859	938	861
Mean	2.69	3.57	3.15	3.81
Std. Dev.	2.04	1.35	2.08	1.28
t	10.6844		8.0204	
df	1797		1797	
p	< 0.0001		< 0.0001	
			Comparison between Grade 2 and 3	
t			4.8383	3.7834
df			1876	1718
p			< 0.0001	0.0002

Although comparisons of EGRA scores between countries often exclude those subjects with zeros scores, where such high proportions of students are scoreless, as in Mozambique and in the case of Egypt, this is precisely where the intervention can be shown to have an impact. EGRA/Egypt[29] clearly show that where many do not respond at all at baseline, that is where the intervention obtained the largest gain relative to the baseline. Had we used only those students who had scores of greater than zero on many of the subtests, this would leave the IE with so few students that the baseline would have been pointless to have been conducted.[30]

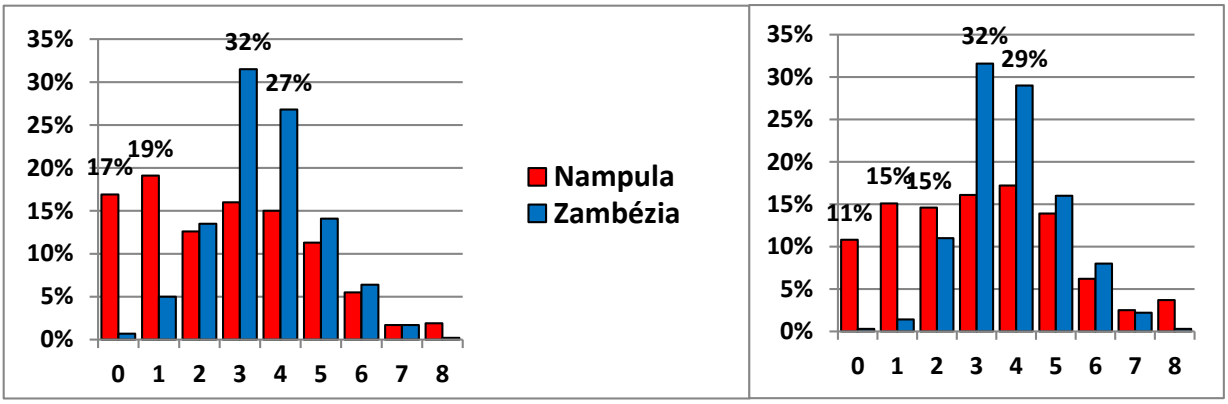
In all cases, the means are significantly different, not unusual given the ample sample sizes. But as the following graphs show, the scores were quite different in their distribution between the two provinces. Note how in both grades, the scores of Zambézia (shown in blue) appear distributed in the classic "normal curve" fashion, while in Nampula (shown in red) the large number of students with low scores distorts the distribution markedly, persisting to Grade 3. While in both provinces scores are low, in Nampula a large proportion of early grade students have not acquired even basic Portuguese comprehension.

[29] http://www.globalpartnership.org

[30] It is our belief that the presentation of these baseline results should not only describe those who have basic reading skills. At this stage of the project, it is also important to know where the problems are and not merely focus on the less than 5% of students where learning to read was observed to be occurring.

Figure 9: Differences in oral comprehension between the two provinces

Second Grade **Third Grade**



Subtask Oral Vocabulary (Following oral instructions, "Place the pencil on the paper")

This subtest measures the oral comprehension ability of the student to respond to simple instructions common to a classroom environment. The subtest consists of six items using a pencil relative (on, behind, in front of, etc.) to other items. The distribution and mean scores are shown in Table 10.

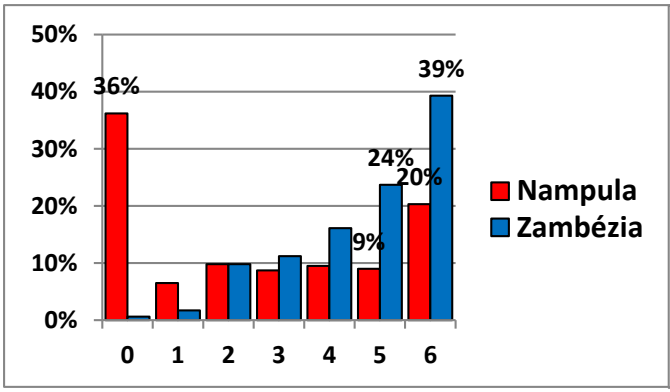
Table 10: Understanding oral language (*"Place the pencil…"*)

Correct Answers "Place the pencil…	**Second Grade**		**Third Grade**	
	Nampula	**Zambézia**	**Nampula**	**Zambézia**
0	36.2%	0.6%	27.3%	0.5%
1	6.5%	1.7%	7.1%	1.0%
2	9.8%	7.3%	5.9%	3.9%
3	8.7%	11.2%	9.9%	8.5%
4	9.5%	16.1%	10.7%	14.2%
5	9.0%	23.7%	9.1%	21.4%
6	20.3%	39.3%	30.1%	50.5%
		Sig.		Sig.
	Comparison between Provinces, Same Grade			
N	940	859	938	861
Mean	2.57	4.69	3.17	5.01
Std. Dev.	2.39	1.41	2.44	1.28
t	22.6445		19.7705	
df	1797		1797	
p	< 0.0001		< 0.0001	
			Comparison between Grade 2 and 3	
t			5.3831	4.9281
df			1876	1718
p			< 0.0001	< 0.0001

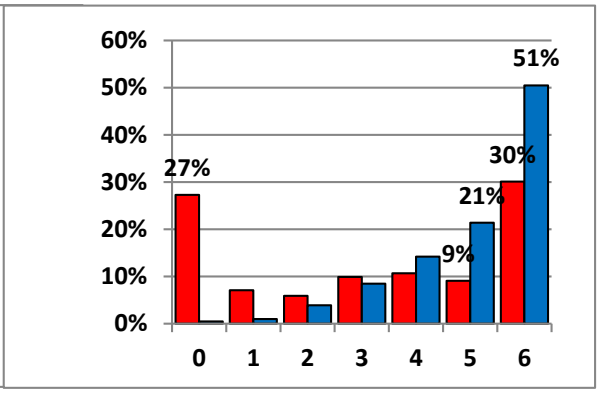
The differences in the mean scores of the two provinces are even more striking, due to the high proportion of students in Nampula with little or no (43% in second and 34% in third grade) ability to follow oral instructions (compare to 2% in Zambézia). On the other end of the distribution, note that 63% of second graders in Zambézia score 5 or 6, compared to just 29% in Nampula. These patterns are repeated in third grade, as the following graphs clearly demonstrate:

Figure 10: Understanding of Oral Language in the two provinces

Second Grade



Third Grade



Subtask Concepts about printed materials

Oral vocabulary aside, learning to read also requires learning the use of printed materials. This subtest of the EGRA was carried out in Portuguese but, if necessary, instructions were given in the local language. The 10 items ranged from identifying the cover of a book, direction of reading, to following the words as spoken. The following table shows that, again, students in Nampula lag behind their counterparts from Zambézia, with fully 22% of Nampula second graders unable to respond correctly to a single item (5% in Zambézia). Mean scores within province did increase significantly between second and third grades, suggesting that learning does occur in these grades and that the EGRA is capable of measuring gains made during the school year. Table 11 shows the distribution and mean scores of the subtask.

Table 11: Demonstrated familiarity with printed materials

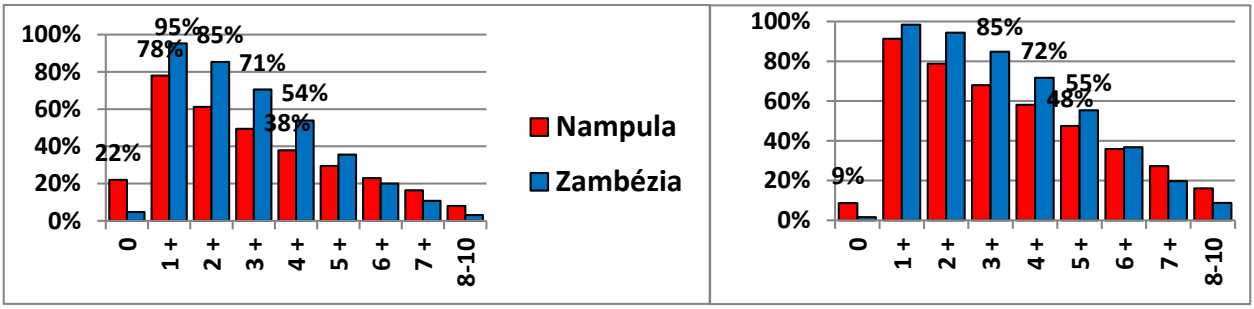
Correct Responses Concepts about Print	Second Grade		Third Grade	
	Nampula	Zambézia	Nampula	Zambézia
0 correct	22%	5%	9%	2%
1 or more	78%	95%	91%	98%
2 or more	61%	85%	79%	94%
3 or more	49%	71%	68%	85%
4 or more	38%	54%	58%	72%
5 or more	29%	36%	48%	55%
6 or more	23%	20%	36%	37%
7 or more	16%	11%	27%	20%
8 - 10 correct	8%	3%	16%	9%
		Sig.		Sig.

	Comparison between Provinces, Same Grade			
N	940	859	938	861
Mean	3.20	3.93	4.46	4.98
Std. Dev.	2.83	2.12	2.85	2.15
T	6.1466		`4.3391	
Df	1797		1797	
P	< 0.0001		< 0.0001	
			Comparison between Grade 2 and 3	
T			9.6132	10.1979
Df			1876	1718
p			< 0.0001	< 0.0001

Figures 11 illustrate the differences between provinces: in second grade, 54% of students from Zambézia had mastered four or more of the skills measured by the subtest, while only 38% from Nampula had. Third grade shows a relative improvement: 55% from Zambézia and 48% from Nampula had now mastered five or more of the ten items. Still, this means that half of all students remain poorly equipped to use print materials. An item analysis reveals that some skills, such as selecting and reading a letter or word from the text, are only developed in a very small percentage of the students.

Figure 11: Mastery of subtask Concepts about Printed Materials

Second Grade **Third Grade**



Subtask Letter/Sound Recognition

To learn to read, students must gain mastery at recognizing the symbols (letters) that words are composed of. It is not enough to learn to read, students must gain mastery at recognizing the symbols (letters) that words are composed of. It is not enough to "recite the alphabet" without associating the written letters with the sounds they produce by themselves or in combinations. This subtest of the EGRA is designed to measure whether students can identify letters presented in random order, in both lowercase and uppercase. This is a timed subtest, in which students read as many letters as they can in one minute. The actual letters and their frequency on the list are based on the frequency of their use in Mozambican Portuguese appropriate to the third grade level. During school visits, it was determined that some teachers emphasized block letters, others cursive, others used both. The official Portuguese language textbook for these grades uses both. We thus prepared a large, plasticized two-sided card with identical letters in both formats and asked the student to choose the one he or she was most familiar with.

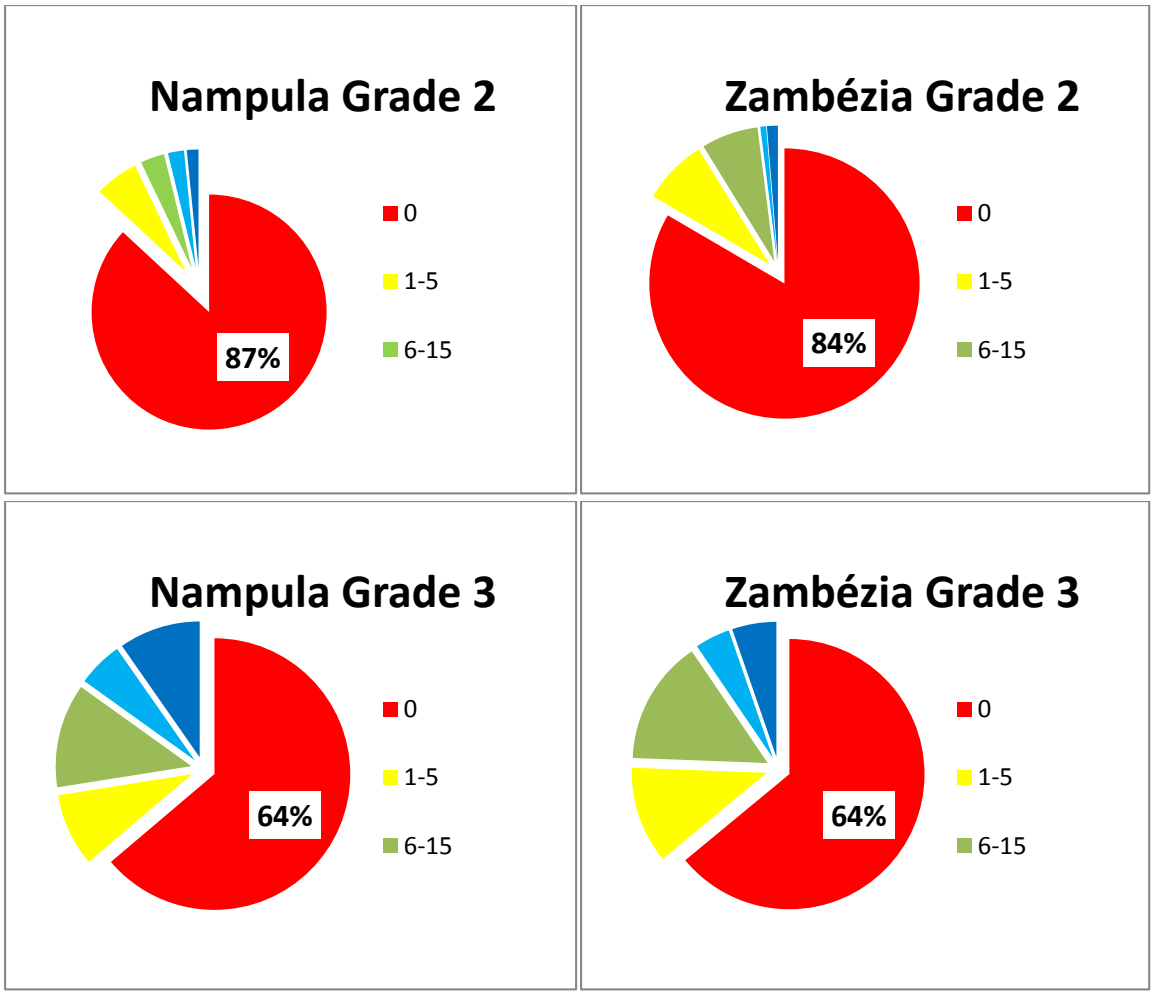
The results are very disappointing, in both provinces and in both grades. The following table shows that fewer than one in five second grade students were able to correctly identify a single letter in the one-minute period; only about one in three could in third grade. The greater similarity between the provinces suggests that the greater oral comprehension noted earlier in Zambézia does not automatically transfer into the skills measured by this subtest: letter recognition is not being effectively taught and thus a key building block of reading acquisition has not been established. The gains between grades in mean scores, while not achieving satisfactory results in an absolute sense, do show that the second grade is addressing some of the deficits in this area. But not nearly enough, and again, far too many students are advancing in grade level without acquiring the necessary skills to develop reading ability. Table 12 displays the results for this subtask.

Table 12: Ability to identify and read letters

Letter Recognition: Letters Read Correctly in One Minute	Second Grade		Third Grade	
	Nampula	Zambézia	Nampula	Zambézia
0	86.9%	83.5%	63.8%	64.0%
1-5	6.0%	7.8%	8.7%	11.6%
6-15	3.3%	6.8%	12.4%	14.9%
16-25	2.2%	0.7%	5.4%	4.2%
26 or more	1.6%	1.3%	9.7%	5.3%
		Sig.		Sig.
	Comparison between Provinces, Same Grade			
N	940	859	938	861
Mean	1.66	1.51	6.70	4.88
Std. Dev.	7.05	5.47	13.88	11.12
t	0.5009		3.0521	
df	1797		1797	
p	0.6165 NS		0.0023	
			Comparison between Grade 2 and 3	
t			9.9237	7.9719
df			1876	1718
p			< 0.0001	< 0.0001

The following graphs highlight the problem: some students are developing these skills, but the critical problem is the magnitude of the student population, in both grades, who have not. Far too many students could not face the challenge of identifying the random letter, but simply recited what they had learned of the alphabet in what we have identified as a "choral" response. It is critical to teach discernment of letters if students are to move forward in reading development.

Figure 11: Mastery of subtask Letter/Sound recognition



Subtask Familiar Word Reading

This EGRA subtest reflects the ability of the students to decode the printed word and sound it out correctly. Consisting of 30 separate common words, the score is the number of words read correctly in one minute, and thus reflects both the accuracy and fluency of reading. Both factors are fundamental to developing a real ability to read and comprehend that which is read.

Given the dismal results of the letter recognition subtest, it is worth highlighting the similar results in this subtest: the percentages with total non-performance are very similar, in both provinces and both grades. Again, third grade scores are significantly better, but hardly adequate at this level. Only a very small percentage of students have attained any fluency. The mean number of correct words per minute simply reflects what we know: most students have not learned to decode words at all, those who have are in fact developing reading skills, some better than others. Table 13 displays the results for this subtask.

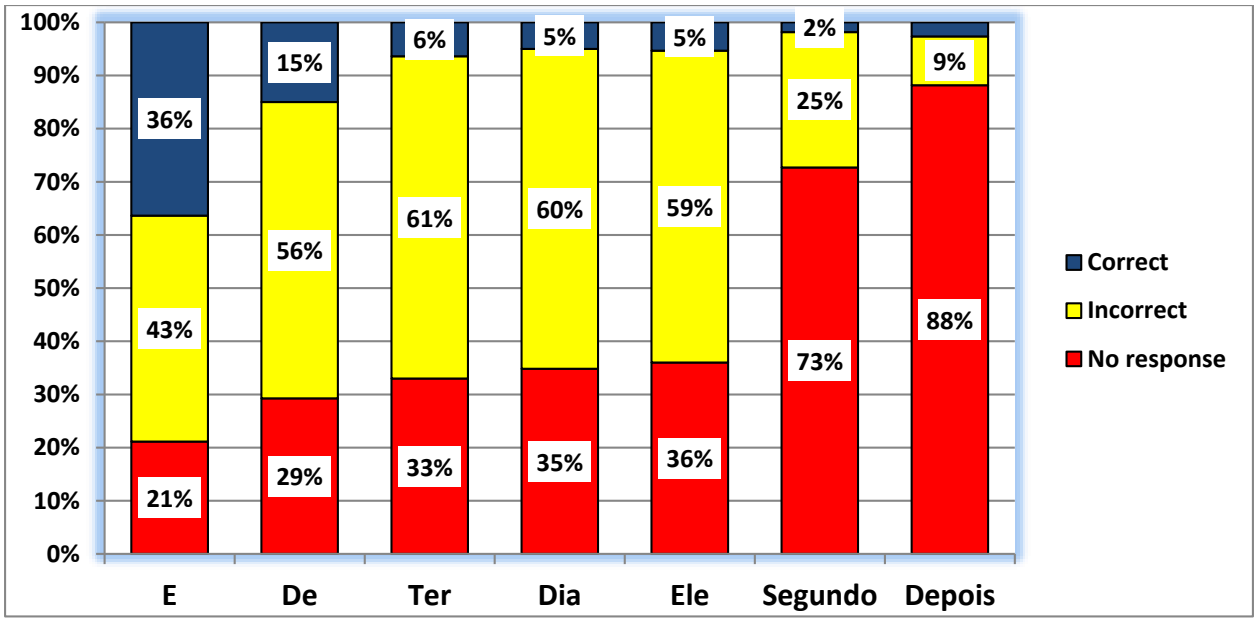
Table 13: Mean scores and distribution for subtask Word Recognition

Words Read Correctly (of 30) in One Minute	Second Grade		Third Grade	
	Nampula	Zambézia	Nampula	Zambézia
0	85.9%	80.4%	67.2%	62.7%
1-2	11.7%	18.7%	21.9%	29.8%
3-10	1.5%	0.3%	5.3%	4.1%
11-20	0.4%	0.5%	1.4%	1.5%
21 or more	0.5%	0.0%	4.3%	1.9%
		Sig.		Sig.
	Comparison between Provinces, Same Grade			
N	940	859	938	861
Mean	0.41	0.31	1.98	1.28
Std. Dev.	2.28	1.14	5.97	4.17
t	1.598		2.8593	
df	1797		1797	
p	0.2463 NS		0.0043	
			Change between Grade 2 and 3	
t			7.5312	6.5768
df			1876	1718
p			< 0.0001	< 0.0001

Why such low scores? At an item level, we looked at what was causing the problem. The following graph looks at the response pattern for just the first seven words in the list from the third grade students from Zambézia. As can be seen, the words, listed at the bottom of the chart, are not particularly challenging. The red portion of each bar shows the percentage of students who did not respond, the yellow those who got the word incorrect, and the blue at the top those who got it right. Non-response rises rapidly with the sixth word at 73% and is 88% by the seventh, suggesting extremely limited fluency. Examination of the remaining 23 items showed a great similarity with item seven: a few students had the fluency accuracy improved somewhat, but these students are but a handful among many.

Also notable with this finding is that, among the first seven words, among those who do respond, wrong answers outstrip correct answers with every word, achieving a remarkable 12:1 ratio with most. It is clear that, for almost all students, they have neither the fluency nor the accuracy of word recognition to proceed to the next step of reading development, that of reading a text and understanding what it means.

Figure 12: Word Response Pattern, First Seven Words, Zambézia, Third Grade



Subtask Reading Fluency and Comprehension

This subtest consists of two short passages that contain stories that are of appropriate level and interest to students, and of roughly equal length (120 and135 words, respectively). For each story, students were given one minute to read as much as they could, then, based on how much they had read, up to four comprehension questions were asked and the results recorded. The first comprehension question required that at least 13 words of the story had been read.

> *"To understand a simple passage given the capacity of short-term memory, average students should read a minimum of 45-60 words per minute by the end of third grade. Learning research and existing data suggest that this standard is possibly usable worldwide."* Helen Abadzi, Education for all, 2011

The results from the prior section on Familiar Word Reading clearly predict the results obtained: most students could not read with fluency sufficient to pass to the comprehension stage, less than 5% in grade 3, and less than 1% in grade 2. The majority of the students who were able to read enough words of the stories (14 students in grade 2, 85 students in grade 3, out of roughly 1,800 students in each grade), answered the comprehension questions correctly. Simply for documentation purposes, the following tables record the number of words read for each story.

Table 14: Number of words read for story # 1

First Story: Words Correctly Read in One Minute	Second Grade		Third Grade	
	Nampula	Zambézia	Nampula	Zambézia
0	98.8%	72.9%	92.5%	56.0%
1-10	0.4%	26.7%	1.6%	40.3%
11-20	0.5%	0.2%	2.8%	1.9%
21-30	0.1%	0.1%	0.4%	0.7%
31 or more	0.1%	0.1%	2.7%	1.2%
		Sig.		Sig.
	Comparison between Provinces, Same Grade			
N	939	859	938	861
Mean	0.17	0.43	2.73	1.67
Std. Dev.	1.92	1.77	14.51	7.17
t	2.9769		1.9374	
df	1796		1797	
p	0.003		0.0529 NS	
			Change between Grade 2 and 3	
t			5.3596	4.9213
df			1876	1718
p			< 0.0001	< 0.0001

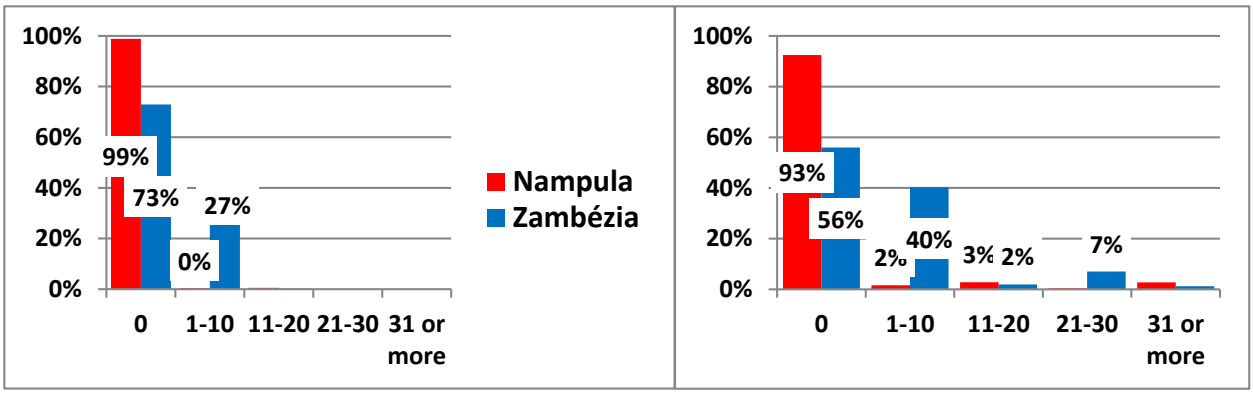
Table 15: Number of words read for story # 2

Second Story: Words Read Correctly in One Minute	Second Grade		Third Grade	
	Nampula	Zambézia	Nampula	Zambézia
0	99.0%	71.0%	92.3%	54.8%
1-10	0.2%	28.1%	1.3%	39.4%
11-20	0.3%	0.8%	2.8%	3.6%
21-30	0.3%	0.1%	1.1%	0.9%
31 or more	0.1%	0.0%	2.6%	1.3%
		Sig.		Sig.
	Comparison between Provinces, Same Grade			
N	940	859	936	861
Mean	0.21	0.62	3.01	2.29
Std. Dev.	2.60	1.83	16.07	8.79
t	3.8345		1.1642	
df	1979		1795	
p	0.0001		0.2445 NS	
			Change between Grade 2 and 3	
t			5.2732	5.4517
df			1874	1718
p			< 0.0001	< 0.0001

The following charts show that, in this subtest, Nampula is again well behind Zambézia in terms of the number of students who could read any words of the story. While Zambézia shows more students with the ability to read at least a few words of this page in a minute, the lack of fluency is such that few are able to focus more on comprehension and less on the decoding of the words. Reading fluency refers to the ability to read text accurately, quickly, and with good expression so that time can be allocated to understanding what is read (Meyer & Felton, 1999) and worldwide studies have shown that fluency increases comprehension. [31] In both provinces and both grades, fluency is a serious problem for the development of reading ability. The origins of the serious deficiencies observed can easily be traced backwards in the EGRA subtests in the prior sections, which clearly demonstrate that the necessary skills to learn how to read have not been developed in the necessary sequence, and at the appropriate developmental stage of the students.

The IE baseline has both established what the challenges are in Mozambique and shown where USAID/APAL interventions are necessary. Lastly, and key to the purpose of such a baseline, we have shown that the effects of such inputs can be measured at the different stages of individual student reading skills development. Figure 13 shows the results for this subtest and the magnitude of the challenges that MINED, and USAID/APAL, face.

Figure 13: The challenge: reading fluently with comprehension



EGRA results by sex of the student

EGRA results were also examined by sex of the students assessed (Annex I). On the first oral comprehension sub-test of the EGRA, no significant differences were observed between boys and girls in either grade in Nampula. Third grade girls in Zambézia performed slightly better (mean score of 3.9) on this sub-test than did boys (3.7), although the reverse was true for the second oral comprehension sub-test (girls, 4.6; boys 4.8). Boys outperformed girls on the concepts about print sub-test only in second grade in Nampula (girls, 3.0; boys 3.4). Boys also scored higher on the letter recognition sub-test in second grade in Zambézia (1.0 girls; 2.2 boys) and in third grade in Nampula (5.0 girls; 8.1 boys). Familiar word reading scores of girls and boys were not different in either province in second grade, but in third grade more girls than boys had scores of zero in both Nampula (girls, 71%; boys, 64%) and Zambézia (girls, 68%; boys, 58%). Scores of zero on the first story reading sub-test were also more

[31] Berninger, V.W., Abbott, R.D., Billingsley, F., & Nagy, W. (2001). Processes underlying timing and fluency of reading: Efficiency, automaticity, coordination, and morphological awareness. In M.Wolf (Ed.), *Time, Fluency, and Dyslexia*. Timonium, MD: York Press.

common among girls in second grade only in Nampula (girls, 77%; boys, 68%); in third grade more girls scored zero in Nampula (girls, 96%; boys, 90%) and also lagged, though not significantly, in Zambézia (girls, 59%; boys, 53%). Similar results were obtained in the second story reading sub-test. While performance on prereading skills is not consistently different between boys and girls, by third grade boys outperform girls in familiar word reading and in passage reading in both provinces. However these performance advantages are small but significant as shown is Annex I.

Comparability of the EGRA results of the RCT groups

We also examined the EGRA sub-test results to assess whether *a priori* differences existed between the RCT treatment and control groups (Annex J). Our flagship sub-tests for this purpose, based on the results presented earlier, were letter recognition and word recognition. These sub-test scores were not statistically different between the groups. The two oral comprehension sub-tests exhibited small but significant differences between groups in both provinces and in both grades, as did results for concepts about print in all but Nampula second grade students. Seven of the eight story-reading comprehension questions were not significantly different, and, given the extremely low scores overall, are of no practical importance—a few higher scoring outliers can easily make the mean scores appear different when almost all of the students scored zero on the sub- task.

As indicated above, these variations, along with other end-of-year independent variables, will be statistically taken into account when performing comparisons betweem the baseline and end-of-year results of EGRA in 2013 and 2014. We conclude that the randomization process was successful in minimizing between-group differences in student characteristics and reading performance.

5.5 About the few students able to read

So few students demonstrated letter and word recognition abilities that further analyses were conducted to examine the socio-demographic characteristics of those who did show reasonable performance in comparison to those who did not. Letter recognition scores of 26 or more correct answers were used to differentiate second grade students; correctly recognized word counts of 21 or more were used for third grade students. The number of students performing adequately was deemed too small to conduct analyses in either second (11 students) or third (16 students) grade in Zambézia and in second grade (15 students) in Nampula. Third grade in Nampula included a total of 40 students scoring 21 or more words read correctly (out of a total of 938 students, or 4.3%), and a selection of variables from the EGRA student cover sheet were examined in relation to performance. We also looked at whether the better performing students were "clustered" in specific schools, or were concentrated in schools that were the head schools of ZIPs. Contained in Annex K, these analyses help to understand the student and home characteristics associated with better performance in Nampula.

Males made up a significantly larger (85%) proportion of better performing students in Nampula third grade than did females (34 of the 40 students were boys). Better performing students were significantly older (11.6 years versus 10.3 years), although no association with self-reported repetition was found. No differences were found for the variables Live with Mother, Live with Father, Attended Preschool, Work at Home, Work Outside of Home, Speak Portuguese with Mother or Speak Portuguese with Father. These results pertain only to comparisons among Nampula third grade students: the number of better-performing students is too small to comment on these relationships in either grade in Zambézia or in second grade in Nampula. The variables Books Read to Student at Home and Student Reads at Home were extremely associated with word reading performance: 70% of better performing students reported being read to at home (versus 32% of their less-well performing peers), while reading at home

by themselves was reported by 73% of successful students on the word reading sub-test versus 29% of those who lagged behind. The 40 successful students were not concentrated in specific schools but distributed among 28 schools in Nampula, half from schools in which only one student of the ten assessed achieved 21 or more correct words. The intra-class correlation of 0.04 indicates that virtually all of the variation in achievement is due to differences between individual students rather than differences between schools. Similarly, studying at the head school of a ZIP was associated with only six of the 40 students. It is our intention to continue the process of identifying student characteristics and other factors related to better performance among students at trhe end of the school year in both grades.

5.6 Interview Findings: Potential for sustainability

In the context of the USAID/APAL activity, we defined sustainability as an improvement in the ability or disposition of the local change agents—district officials, pedagogical directors, school directors, and especially directors of "head" schools of the ZIPs—to take actions that will increase the likelihood that activities implemented by USAID/APAL become institutionalized. Of course, sustainability of an initiative regardless of how beneficial it is proven to be is highly dependent on the level of resources that the change agents have at their disposal. The IE's challenge is how to "tell the story" of the existing capacity in a manner that would provide insights on the likelihood that processes and activities implemented become sustainable.

A total of 109 interviews were conducted with MINED provincial and district officers, school directors, ZIP coordinators, pedagogical directors and PTA presidents in the participating schools of the two provinces by IE staff and GSC supervisors: 57 in Nampula and 47 in Zambézia. In addition, five interviews were conducted with MINED central officers by the IE staff COP. Table 16 presents the breakdown of the interviews and Annex H lists all persons interviewed by province, district, school and function.

Table 16: Interviews conducted

Province	Nampula	Zambezia
Districts	2	2
EP1	10	9
EPC	40	47
	Interviews conducted	
Mined District Officers	2	3
School Directors	14	14
School Directors/ZIP Coordinators	18	14
Pedagogical Directors	4	3
PTA presidents	19	13
TOTAL	57	47
MINED Central officers	5	

N=109

The focus of the interviews was the potential for sustainability of USAID/APAL activities. One must bear in mind that this was a convenience sample: Central MINED interviewees were selected by USAID/Mozambique and, in the field, a sample was drawn among those present, available and willing to participate. Therefore, these are not appropriate data to draw inferences regarding gender equity among MINED administrative staff. MINED possesses gender disaggregated information for all staff at the level of school district and province.

The information gathered is grouped around the guiding questions included in the protocol interviews. An analysis of the information gathered by the interviews shows broad agreement about the role of the directors of the "head" schools of the ZIPs in sustaining USAID/APAL activities. Respondents consider that the "head" schools of the ZIPs have more resources, more educational materials, are better equipped to conduct training and have closer communication with MINED. At this time, only the potential for sustainability can be assessed. The residual effects of the intervention—Are the activities implemented by USAID/APAL still to be present at schools (Full and Medium treatment groups) even when USAID/APAL's resources are not?—can only be assessed as data are collected after one year of non-intervention. Six themes emerged from an analysis of the 109 interview narratives:

- **There is concern regarding the acquisition of reading skills in the early grades**. With few exceptions interviewees agree that students are not acquiring the reading skills that they need to succeed in later grades. Examples of typical statements include: "*…at this school students do not learn how to read…"* or *"…low levels of reading skills in grades 2 and 3…"* (PTA presidents); "*…we haven't been successful in teaching students to read because of the language they speak at home…" or "…about 50% of the students are able to read…"* (pedagogical directors); or "*…at the end of the second and third grades students often have not acquired basic reading skills…"*
- **The fund *Apoio Directo à Escola* (ADE) is essential.** Even though not sufficient to cover all school needs, the ADE is viewed by the great majority of interviewees as the only resource available for the acquisition of materials and other small items.
- **Teacher preparation and lack of materials are seen as great challenges.** Teachers need help on how to use materials, organize learning in unconventional settings (outside of the classroom) and teach students who arrive at school speaking a language other than Portuguese. It is considered that teachers need more in-service to deal with the initial grades. Typical statements include: *"…the teachers are not prepared to deal with the initial grades especially students who arrive at school speaking another language…" or "…acquiring materials for the classes is a challenge, books for third grade are never sufficient…"*
- **It is generally conceded that the Teacher Training Institutes** - Instituto de Formação dos Professores (IFP) need to do more to prepare teachers for the reality of the schools. IFPs need to address issues of large classes, unconventional settings, and bilingualism in their teaching preparation courses. At the same time, interviewees advocate a more active role of the IFPs in in-service training.
- **MINED materials are useful and welcome but often insufficient**. Apparently books for grade 1 and 2 students normally arrive on time and are sufficient for all students. Grade 1 and 2 students keep the books and take them home after the end of the year because they are consumable and include writing exercises. MINED then sends each year to the school as many books as students enrolled in those grades. For grades 3 and above the books belong to the school—a student uses a book during the year but must give it back at the end of each year. The problem of insufficient books arises from having students not returning books or books being lost or damaged. The Directors have the responsibility of reporting how many books are at school for grades 3 and above and what the enrollment is for those grades. Problems with the flow of information school-ZIP-MINED often result in insufficient number of books.
- **The ZIPs are seen as the correct strategy to implement MINED policies**. Interviewees point out that the head schools of the ZIPs already conduct much of the training and take initiatives to improve teaching and learning. Having schools grouped in a ZIP is seen as an correct strategy to facilitate teacher training and delivery of books and materials. The head school of the ZIP is seen as a guarantee that there will be open communication channels to the district and to the MINED. Typical statements include: "*…the ZIP helps us with in-service and exchange of information…"* or *"…the ZIP head school has competent people who they visit our school and help with the coordination between parents and teachers…"* However, the directors of head schools of the ZIPs point out that without

additional resources, the expectations that they influence schools by assisting with training and monitoring may be frustrated. Whereas the ZIPs seem to be the right point of intervention for USAID/APAL, the lack of resources mentioned by every ZIP director/coordinator and the lack of incentives for the additional work that needs to be conducted could be seen as threats to USAID/APAL implementation and to the sustainability of the activities implemented.

6. CONCLUSIONS

The aim of the baseline study was to assess the early reading skills among grade 2 and grade 3 students in a sample of Mozambique's public schools to allow comparisons between students in the two treatment groups and students in the control group. The study also sought to discuss school and student characteristics that could be related to student performance The conclusions that emerged from the analysis of the data collected among 3,598 second and third graders, 341 teachers, and 171 school directors are backed by data collected in the schools that formed the sample and should be applied only to that group of sampled schools in the USAID/APAL target areas in each province of Nampula and Zambézia.

- The EGRA is sensitive enough to detect differences between student performance in grades 2 and 3. This suggests that the instrument will also be sensitive to detect differences in the performance of students who benefitted from the intervention and those in the control group to which no treatment was applied.
- In spite of the gains observed between second and third grades, students—especially third graders—did not demonstrate the reading level that a student should achieve at the end of first and second grades. The great majority of the 3,598 second and third graders[32] assessed in the 180 randomly selected schools are not acquiring the foundational skills that will allow them to become fluent readers. Only a very small group of students (generally less than 2%) demonstrated a notably higher level of performance.
- The results of the EGRA reveal that by the end of grade 2 (third graders), the majority of students have not yet acquired sufficient foundational skills that will allow them to read fluently with comprehension. Students showed poor oral vocabulary, had trouble following instructions, demonstrated limited knowledge of the letters, were unable to correctly identify the names associated with the letters. Given students' difficulties in identifying letter names, it is not surprising that they could not decode words. Taken together, these findings suggest that reading instruction in the early grades is not focusing on helping students to learn the sounds associated with each letter and applying this knowledge to sound out familiar or unfamiliar words.
- Because students have not acquired the basic building blocks for reading, their oral reading fluency scores were low. As a consequence, reading comprehension was also low, with fewer than 5% of grade 3 students being able to correctly answer even the first of the eight reading comprehension questions.
- As repeatedly pointed out in the school effectiveness literature previously cited and reported by the SMA inadequate infrastructure, attendance rates of 70%, lack of reading material, teacher and school director tardiness, absenteeism and lack of preparation, low level of commitment to the task of improving reading on the part of school directors, limited number of days and hours spent at school, and little time-on-task dedicated to reading instruction, all may contribute to the unsatisfactory results obtained by the students on the EGRA.

[32] The baseline data collection was conducted in February/March 2013, at the beginning of the 2013 school year. Students assessed were beginning second and third grades.

- The high level of self-reported confidence demonstrated by teachers is surprising when one examines second and third grade results on the EGRA. The predominance of the 10+1/2 and the limited in-service opportunities focused on early grade literacy may be one of the elements that fuel this confidence—teachers may be unaware of their own limited skills—and suggests the need for a specific strategy to improve teaching to read in the early grades.
- In Nampula, where students' reported knowledge of Portuguese is less prevalent, students may take longer to acquire the skills necessary to read. Considering that Portuguese is the official language of instruction, the limited use of Portuguese outside of school seems to be one of the factors that influence the acquisition of early literacy skills. In Zambézia, where Portuguese is more prevalent, second grade students did significantly better on the EGRA oral subtasks than in Nampula, where the local language is, reportedly, more frequently spoken outside of school. In the beginning of third grade—after two years of instruction conducted in Portuguese—Nampula students start to catch up and the percentage of third graders who could not respond to any of the 14 oral language prompts decreases from 16% to 10% (subtask #1) and from 32% to 27% (subtask # 2) reducing the gap between the two provinces.

REFERENCES

Abadzi, H. (2011). Reading Fluency Measurements: in EFA FTI Partner Countries: Outcomes and Improvement Prospects, EFA FTI Secretariat Update September 2011.

Abadzi, H. (2007). "Absenteeism and Beyond: Instructional Time Loss and Consequences", World Bank Policy Research Working Paper No. 4376, p. v.

Abadzi, H. (2009). "Instructional Time Loss in Developing Countries: Concepts, Measurement, and Implications." World Bank Research Observer. 24 (2): 267-290.

Abadzi, H. (2011). Reading Fluency Measurements in EFA Partner Countries: Otcomes and Improvement Prospects, EFA, FTI Secretariat Update September 2011.

Adelman, E., A.M. Moore, S, Manji. (2011) USAID EQUIP2. 2011. "Using Opportunity to Learn and Early Grade Fluency to Measure School Effectiveness in Mozambique."

Bandura, A. (1997). Self-efficacy: The exercise of self-control. New York: W.H. Freeman and Company.

Berninger, V.W., Abbott, R.D., Billingsley, F., & Nagy, W. (2001). Processes underlying timing and fluency of reading: Efficiency, automaticity, coordination, and morphological awareness. In M.Wolf (Ed.), Time, Fluency, and Dyslexia. Timonium, MD: York Press. EPT: "Rapport Mondiale de Suivi sur L'Éducation Pour Tous."

Clark, R.E. & Sugrue, B.M. (1991). "Research on instructional media, 1978-1988" in G.J.Anglin (ed.) *Instructional technology: past, present, and future* ch.30 pp.327-343 (Libraries unlimited: Englewood, Colorado).

EQUIP2. (Working Paper) "Using Opportunity to Learn and Early Grade Fluency to Measure School Effectiveness in Ethiopia, Guatemala, Honduras, and Nepal."

Fehrler, S., K. Michealova and A. Wechtler (2009). The cost-effectiveness of inputs in primary education: Insights from the literature and recent surveys for sub-Saharan Africa. The Journal of Development Studies, Taylor & Francis Journals, vol 45 (9), pages 1545-1578.

Glewwe, P., E. Hanushek, S. Humpage, and R. Ravina (2011). School Resources and Educational Outcomes in Developing countries: A Review of Literature from 1990 to 2010. http://www.nber.org/papers/w17554.

UNESCO. L'éducation pour Tous. Aperçu regional: Afrique Subsaharienne. www.efareport.unesco.org.

Passos, A., Nahara, T., Magaia, F., & Lauchande, C. (2005). The SACMEQ II Project in Mozambique: A Study of the Conditions of Schooling and the Quality of Education. Harare: SACMEQ.

Scriven, M. (2008). A Summative Evaluation of RCT Methodology and an Alternative Approach to Causal Research. Journal of Multidisciplinary Evaluation, Vol 5, No 9, March 8. http//www.jmde.com.

Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) Report III 2010-2011.

Spaull, N. (2011). Primary School Performance in Botswana, Mozambique, Namibia, and South Africa. SACMEQ Working Paper Nr. 8.

The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ). Report on Education 2009-2013. http://www.sacmeq.org/.

UNESCO. Education for All Global Monitoring Report. http://www.unesco.org/.

UNESCO Institute of Statistics. Montreal, Canada. Global Education Digest 2012. "Opportunities Lost: The impact of grade repetition and early school leaving." www.uis.unesco.org.

UNESCO Institute of Statistics. Data to Make a Difference. www.uis.unesco.org.

World Education. Projects by Region. www.worlded.org.