



# Impact Evaluation of the USAID/ Aprender A Ler Project in Mozambique

## Year 3 IE/RCT

## Final Report – Executive Summary

May 23, 2016

This report was prepared for USAID/Mozambique by Magda Raupp, Bruce Newman, Luis Revés, Carlos Lauchande and Edward Jay Allan under Contract AID- 656-C-15-00002 awarded to International Business & Technical Consultants, Inc. (IBTCI). The authors' views expressed in this report do not necessarily reflect the views of the United States Agency for International Development or the United States Government.



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

## **IE/RCT Report Final Report – Executive Summary**

### **International Business & Technical Consultants Inc.**

Magda Raupp, Team Leader  
Bruce Newman, Statistician/Data Analyst  
Luis Revés, GSC Research, Deputy Team Leader  
Carlos Lauchande, Statistician  
Edward Jay Allan, Project Director  
Merrill Anne Jordan, Program Associate

May 23, 2016

#### **DISCLAIMER**

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



# TABLE OF CONTENTS

<b>LIST OF FIGURES</b> .....	iiii
<b>LIST OF TABLES</b> .....	iiii
<b>MAP OF APAL</b> .....	v
<b>ACRONYMS</b> .....	vi
<b>EXECUTIVE SUMMARY</b> .....	1
Did the ApaL reading intervention improve reading outcomes? .....	2
What contributed to this achievement?.....	4
Are the effects sustainable after the cessation of the intervention? .....	6
Summary Recommendations.....	9
<b>1. CONTEXT</b> .....	11
<b>2. THE USAID/APRENDER A LER INTERVENTION</b> .....	133
<b>3. METHODOLOGY OF THE IMPACT EVALUATION/ RCT</b> .....	14
3.1 Impact Evaluation Question for 2015.....	166
3.2 Data Collection Instruments.....	177
3.2.1 Early Grade Reading Assessment (EGRA) Instrument.....	17
3.2.2 School Management Assessment (SMA) Instrument.....	19
3.2.3 Interview Protocol for interviews with school directors.....	19
3.3 Instrument Administration .....	19
3.3.1 EGRA and Student Interview Administration .....	19
3.3.2 Teacher Interview.....	20
3.3.3 School Director Interview .....	20
3.3.4 Portuguese Class Observation.....	200
3.3.5 Classroom Inventory.....	200
3.3.6 Interview with School/Pedagogical Directors.....	200
3.4 Field Data Collection Procedures.....	200
3.5 Data Analysis.....	211
3.5.1 Analysis of the EGRA reading scores.....	211
3.5.2 Analysis of the Class Observation Data .....	211
3.5.3 Analysis of the School/Pedagogical Director Interview Data .....	211
3.6 Limitations of the Study .....	222
<b>4. CHARACTERISTICS OF THE POPULATION</b> .....	233
4.1 Description of Students Assessed by the EGRA and interviewed.....	233
4.2 Grade 2 and Grade 3 Teachers Interviewed .....	27
<b>5. FINDINGS</b> .....	28
5.1 2015 EGRA results by treatment group and grade.....	29
5.1.1 Comparing 2014 with 2015 EGRA scores, by treatment group and grade.....	34
5.2 Differences in performance of boys and girls .....	38
5.3 Difference in Performance for Students in Urban and Rural Schools.....	40
5.4 Unplanned additional training and training activities provided to some school directors, pedagogical directors and teachers during 2015.....	43
5.5 USAID Indicators.....	44
5.6 Teachers.....	45
5.6.1 Classroom Observation (SMA/ApaL).....	47
5.7 The classroom environment.....	49

5.8 School Directors/Pedagogical Directors and Their Perspectives on Sustaining ApaL Interventions.....	54
5.8.1 ApaL strategies sustained one year after the end of the intervention .....	55
5.8.2 Respondents' view of the training they require to sustain the intervention.....	56
5.8.3 Trained teachers who remain in school teaching the same grades.....	56
5.8.4 Training sustained in 2015.....	56
5.8.5 Support received to continue activities promoted by ApaL.....	57
5.8.6 The Use and Replacement of Teaching and Learning Aids (TLAs) .....	57
5.9 Evaluation Question # 4: Management Sustainability.....	58
5.10 Costs and Cost-Effectiveness over the 2014-2015 Period.....	60
<b>6. CONCLUSIONS.....</b>	<b>62</b>
<b>7. RECOMMENDATIONS.....</b>	<b>65</b>
<b>ANNEXES.....</b>	<b>67</b>
Annex A. Scope of Work.....	67
Annex A.1. Scope of Work for the Evaluation Services Task Order.....	67
Annex A.2. Scope of Work for the Follow-On Contract.....	86
Annex B. Reports Submitted for Work Done Under the Task Order.....	99
Annex C. EGRA Instrument (Student Interview).....	100
Annex D. SMA Package (Teacher Interview).....	101
Annex E. Field-level Sampling Strategy for Classrooms & Students for EGRA Administration .....	103
Annex F. Semi-Structured Interview Protocol.....	105
Annex G. Statistical Annexes.....	107
Annex G.1 Deciles Histogram Grade 2 (Text Reading Comprehension, Reading Fluency, Familiar Word Reading Fluency, Letter Recognition, Concepts about Print) .....	107
Annex G.2 Deciles Histogram Grade 3 (Text Reading Comprehension, Reading Fluency, Familiar Word Reading Fluency, Letter Recognition, Concepts about Print) .....	116
Annex G.3 Scatter Plots – Grades 2 & 3.....	125
Annex H. Urban-Rural Differentiations.....	140
Annex H.1 Comparison of 2015 EGRA Mean Scores by Urban/Rural Locality.....	140
Annex H.2 Intermediate variables that may affect EGRA scores by Urban/Rural Locality .....	140
Annex H.3 Use of Portuguese with Family and Friends.....	140
Annex H.4 Comparison of Treatment Group Effects within Urban/Rural Areas.....	141
Annex H.5 Effects of Treatment Group on Grade 2 Urban Students - Detail .....	142
Annex H.6 Effects of Treatment Group on Grade 2 Rural Students - Detail .....	145
Annex H.7 Effects of Treatment Group on Grade 3 Urban Students - Detail .....	148
Annex H.8 Effects of Treatment Group on Grade 3 Rural Students - Detail .....	151
Annex H.9 Student Days Attended in July 2015 by Grade, Area and Treatment Group.....	154
Annex H.10. Total Enrollment, Grades 2 and 3, by Urban-Rural Treatment.....	155
Annex H.11 Enrollment, Grades 2 and 3, by Urban-Rural Treatment - Detail .....	156
Annex H.12 Retention of Teacher Practices.....	158
Annex H.13 Retention of Portuguese Reading Booklets.....	158
Annex H.14 Retention of Classroom Inventory Materials (TLA) .....	159
Annex I. References.....	160
Annex J. Conflict of Interest Statements .....	162
Annex K. USAID Perspectives on the Impact Evaluation.....	166

## LIST OF FIGURES

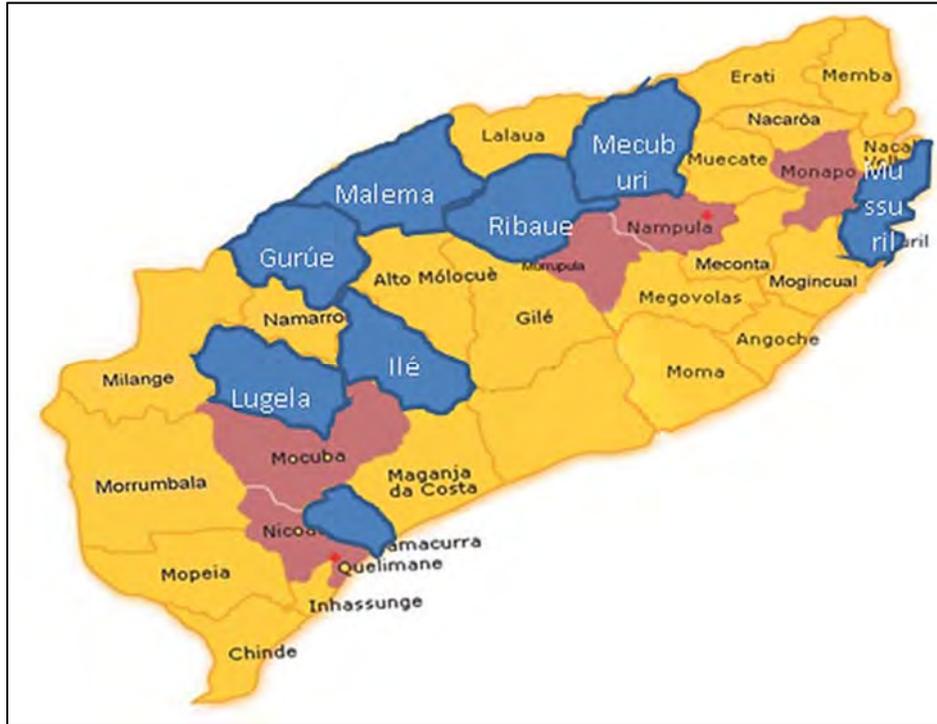
Figure 0-1 Letter Recognition deciles as a predictor of Reading Fluency performance - Grade 2.....	5
Figure 0-2 Letter Recognition deciles as a predictor of Reading Fluency performance - Grade 3 .....	6
Figure 0-3 Familiar Words correctly read by treatment group and year – Grade 3 .....	7
Figure 0-4 Distribution of Familiar Words Correctly read by minute per treatment group.....	8
Figure 0-5 Teacher observed practice by year and treatment (Full).....	9
Figure 0-6 Teacher observed practice by year and treatment (Medium).....	9
Figure 3-1 Design & Timeline of the Impact Evaluation .....	16
Figure 4-1 Self-reported repetition by grade repeated, of 2015 second and third grade students.....	26
Figure 5-1 Distribution of Oral Comprehension scores for Grade 2 students by treatment group.....	31
Figure 5-2 Distribution of Letter Recognition scores for Grade 2 students by treatment group.....	32
Figure 5-3 Distribution of Concept about Print scores for Grade 3 students by treatment group .....	33
Figure 5-4 Distribution of Words Correctly Read per minute for Grade 3 students by treatment group.....	33
Figure 5-5 Relationship between Letter Recognition and Familiar Words Reading – Grade 3.....	35
Figure 5-6 Oral Comprehension scores for Grades 2 and 3 by treatment group and year .....	36
Figure 5-7 Concepts about Print scores for Grades 2 and 3 by treatment group and year .....	37
Figure 5-8 Letter Recognition scores for Grades 2 and 3 by treatment group and year .....	38
Figure 5-9 Familiar Words correctly read scores for Grades 2 and 3 by treatment group and year .....	38
Figure 5-10 Fluency scores for Grades 2 and 3 by Treatment Group and Year .....	39
Figure 5-11. Reading Comprehension scores for Grades 2 and 3 by Treatment Group and Year .....	40
Figure 5-12 Retention of Teachers in Grade 2 or Grade 3 from 2014.....	48

## LIST OF TABLES

Table 0-1. Scores at Midline 1 (Sep 2013), Midline 2 (Sep 2014) and Endline (Sep 2015).....	2
Table 0-2. Differences in performance of girls and boys by treatment group and grade (2015).....	3
Table 0-3. Comparison of 2015 Grade 3 EGRA Mean Scores by Urban/Rural Locality .....	4
Table 0-4. Variables that affect performance of students in rural settings .....	4
Table 0-5. Correlations between Subtests and Percentage of Variance Explained .....	6
Table 0-6. Results obtained in 2014 vs. 2015 .....	7
Table 4-1. Population and sample count of schools where data were collected in 2015.....	24
Table 4-2. School enrollment by grade and by sex .....	24
Table 4-3. EGRA participants by grade and by sex .....	25
Table 4-4. Self-reported Age of students taking the EGRA .....	25
Table 4-5. Age of Students by Treatment Group .....	26
Table 4-6. Family situation as reported by students who took the EGRA.....	27
Table 4-7. Overall absentee rates for boys and girls in both provinces.....	27
Table 4-8. Student reported Portuguese use with parent.....	27
Table 4-9 Student reported Portuguese use with siblings and friends .....	28
Table 4-10. Age of teachers interviewed .....	28
Table 4-11. Teacher Years of Experience.....	28
Table 4-12. Teacher native language.....	29
Table 4-13 Teacher use of local language to facilitate teaching of Portuguese.....	29
Table 5-1. Reading Scores at Midline 1, 2, and Final.....	30
Table 5-2. EGRA Scores obtained by treatment and Control groups in 2015.....	32
Table 5-3. Correlations between Subtests and Percentage of Variance Explained .....	34
Table 5-4. Unit costs per student enrolled for 2014 and combined 2014-2015.....	34
Table 5-5. Oral Comprehension Scores for Grades 2 and 3 by Year and by Treatment Group .....	36

Table 5-6. Concepts about Print Scores for Grades 2 and 3 by Year and by Treatment Group.....	36
Table 5-7. Letter Recognition Scores for Grades 2 and 3 by Year and by Treatment Group.....	37
Table 5-8. Scores on familiar words read correctly.....	38
Table 5-9. Scores on Text Reading Fluency .....	39
Table 5-10. Scores on Reading Comprehension.....	39
Table 5-11. Differences in Performance of Girls and Boys by Treatment Group and Grade .....	41
Table 5-12. Comparison of 2015 Grade 2 EGRA Mean Scores by Urban/Rural Locality.....	42
Table 5-13. Comparison of 2015 Grade 3 EGRA Mean Scores by Urban/Rural Locality.....	42
Table 5-14. Intermediate variables that may affect EGRA scores by Urban/Rural Locality.....	42
Table 5-15. Use of Portuguese with Family and Friends.....	43
Table 5-16. Comparison of Treatment Group Effects within Urban/Rural Areas – Grade 2.....	43
Table 5-17. Comparison of Treatment Group Effects within Urban/Rural Areas – Grade 3.....	44
Table 5-18. Student Days Attended in July 2015 by Grade, Area and Treatment Group.....	44
Table 5-19. EGRA Scores by whether the SD received additional training in 2015 .....	46
Table 5-20. Average Score of Grade 3 Students on Familiar Words Correctly Read per Minute.....	47
Table 5-21. Grade 3 Students Correctly Reading 20 or more Words per Minute by Sex.....	47
Table 5-22. Grade 3 Students Correctly Reading 20 or more Words per Minute by Treatment Group .....	47
Table 5-23. Retention at the school, teaching the same grade by grade and by treatment group .....	48
Table 5-24. Classroom Practices and Behaviors Observed in 2015 by Treatment Group .....	50
Table 5-25. Classroom Practices and Behaviors Observed in 2014 by Treatment Group .....	50
Table 5-26. 2015-2014 Comparisons of Classroom Practices and behaviors by Treatment Group.....	51
Table 5-27. July 2015 Class Days Offered and Attended in July 2015 by Grade .....	52
Table 5-28. Attendance rate of students assessed .....	52
Table 5-29. Teacher-reported days missed in July 2015.....	53
Table 5-30. Portuguese Class Duration in 2015 .....	53
Table 5-31. Observed Portuguese Class Duration in 2014.....	53
Table 5-32. Attendance, reading materials by sex and treatment group .....	54
Table 5-33. Attendance, reading materials and supplies by sex and treatment group (2014).....	54
Table 5-34. Possession of supplementary materials between 2015 and 2014.....	55
Table 5-35. Materials observed in the classroom .....	55
Table 5-36. Materials sustained in 2015 .....	56
Table 5-37. Summary of interviews conducted.....	57
Table 5-38. Training received in 2015 .....	60
Table 5-39. Unit costs per student enrolled for 2014 and combined 2014-2015 .....	64
Table 5-40. Cost effectiveness Grade 2 “one year with intervention, one year without” model .....	64
Table 5-41. Cost effectiveness Grade 3 “one year with intervention, one year without” model .....	65

# MAP OF APAL



**2012-2014 ApaL Program Districts**  
Shown in red

<b>Nampula</b>	<b>Zambézia</b>
Murrupula	Mocuba
Nampula City	Quelimane
Monapo	Nicoadala

Source: ApaL

## ACRONYMS

ApaL	<i>Aprender a Ler</i> (Learn to Read)
CAP	Concepts about Print
CEA	Cost Effectiveness Analysis
clpm	Correct Letters Per Minute
cwpm	Correct Words Per Minute
df	Degrees of Freedom
DPEC	<i>Direcção Provincial de Educação e Cultura</i> (Provincial Directorate of Education and Culture)
EGRA	Early Grade Reading Assessment
IBTCI	International Business & Technical Consultants, Inc.
IE	Impact Evaluation
IFP	<i>Instituto de Formação de Professores</i> (Teacher Training College)
INSET	In-Service Teaching
LEI	Local Education Institution
M&E	Monitoring and Evaluation
MINEDH	<i>Ministério da Educação e Desenvolvimento Humano</i> (Ministry of Education and Human Development)
NS	Not Significant
PD	Pedagogical Director
RCT	Randomized Controlled Trial
RSA	Rapid School Assessment
SD	School Director
SDEJT	<i>Serviços Distritais de Educação, Juventude e Tecnologias</i> (District Services for Education, Youth, and Technology)
SMA	School Management Assessment
SMT	School Management Tool
TIMSS	Trends in International Mathematics and Science Study
TLA	Teaching-Learning Aid
TOT	Training of Trainer
UIS	UNESCO Institute for Statistics
UNESCO	United Nations Educational, Scientific and Cultural Organization
USAID	United States Agency for International Development
WEI	World Education, Inc.
ZIP	<i>Zona de Influência Pedagógica</i>

## EXECUTIVE SUMMARY

Sustainability is frequently mentioned as a development project goal but efforts to sustain interventions beyond the conclusion of funding have only recently emerged as a growing area of emphasis for evaluation. And yet, the sustainability of interventions following the conclusion of formal support should be viewed as an essential element of programs that attempt to bring about meaningful behavior change — especially in the variety of settings with severe resource limitations, as is the case with development projects. Despite widespread acknowledgment that sustainability should be central to intervention efforts<sup>1</sup> (e.g., August et al. 2006; Johnson et al. 2004), specific attention to the process of assessing the sustainability of development interventions remains relatively sparse.

In 2012, the United States Agency for International Development (USAID), in collaboration with the Mozambique Ministry of Education and Human Development (MINEDH), designed the USAID/Aprender a Ler (ApaL) project to improve early grade reading skills and selected World Education, Inc. (WEI) as the implementer and International Business & Technical Consultants Inc. (IBTCI) to conduct an Impact Evaluation (IE) of the project. Both contracts, the one to implement ApaL and the one to conduct the IE, were awarded at essentially the same time. USAID believes that this Impact Evaluation of ApaL is the first external and independent IE of its kind for USAID in Africa. It is also one of the relatively few IEs designed from the beginning to measure the relative impacts of the interventions not only at the conclusion of the interventions but also to measure the lasting impacts of the interventions. Previous reports submitted to USAID/ Mozambique in 2013 and 2014 by IBTCI, under the Evaluation Services IQC task order, have focused on the effects of the ApaL program during implementation. This report specifically addresses the extent to which the effects observed at the conclusion of ApaL interventions in 2014 were sustained once direct support from the implementer ceased.

ApaL began with delivery of the program in 35 *Zonas de Influência Pedagógica* (ZIPs)<sup>2</sup> in six districts along the economic corridors of the provinces of Nampula and Zambézia. Based on experience during initial implementation, expansion to close to five hundred schools in the area was anticipated.<sup>3</sup> USAID selected these provinces because at the time of project design, they contained 42 percent of the school-age population of Mozambique. Furthermore, these heavily populated and rural provinces had posted weak education performance results compared to national averages. At the beginning of the 2013 school year, 180 schools clustered around the 35 ZIPs in the two provinces were randomly assigned by the IE to three groups—60 schools to Full treatment, 60 to Medium treatment, and 60 to a no-treatment or Control group. The Randomized Control Trial (RCT) methodology utilized by the IE ensured that the three groups were equivalent at Baseline. Thus, any differences in reading outcomes could be ascribed to the intervention, and the findings for Control schools would reflect what would have happened in the absence of any intervention.

This report presents the findings, conclusions, and recommendations of an IE of the USAID/ApaL project and constitutes the final phase of this three-year effort to determine the magnitude of the effects of the project on its beneficiaries. Specifically, the analyses (i) examined the degree to which gains obtained by students on the Early Grade Reading Assessment (EGRA) in 2014 have been sustained; (ii) documented which activities, processes and procedures implemented by ApaL in 2014 have remained after the cessation of direct support to the schools; and, (iii) explored whether school directors have the technical skills and the schools have the resources needed to continue activities implemented by ApaL.

---

<sup>1</sup> Johnson K, Hayes C, Center H, Daley C. *Building capacity and sustainable prevention innovations: A sustainability planning model. Evaluation and Program Planning. 2004; 27:135–149.*

<sup>2</sup> *In Mozambique, a head school and surrounding 5-10 schools form a Zona de Influência Pedagógica (ZIP).*

<sup>3</sup> *ApaL expanded to 538 schools in 2015.*

Detailed reports of the work done in 2013 and 2014 under the Evaluation Services IQC task order are on file on the Development Experience Clearinghouse (DEC); Annex B presents a list with links. Here we refer to the main findings of the evaluation, summarize findings observed in 2015, and discuss the conclusions reached, and the implications and recommendations for future direction. The information provided could assist MINEDH, USAID, and other donors in the allocation of additional resources both towards improving early grade reading outcomes per se and the broader learning environment. The findings included in this report can be grouped under three main areas.

## Did the ApaL reading intervention improve reading outcomes?

Measurements using the Early Grade Reading Assessment (EGRA) methodology had compared Midline 1 (2013), Midline 2 (2014) reading scores obtained by Grade 2 and Grade 3 students at the treatment schools to scores of a Control group that received no intervention. The results showed that while overall performance remains low, large absolute and relative gains in reading performance were achieved in the intervention schools that were not achieved in Control schools. It must be noted that the EGRA scores at Baseline had shown no significant statistical differences between the three groups, thus reaffirming the success of the randomized sampling strategy.<sup>4</sup> In all comparisons made between 2013 and 2014, children in treatment schools, especially those in the Full treatment schools, outperformed their counterparts in the Control schools. Table 0-I shows scores obtained by students by year and by treatment group for four EGRA subtests. All differences observed between Full and Medium treatment and Control groups are statistically significant.

Table 0-I. Scores at Midline 1 (Sep 2013), Midline 2 (Sep 2014) and Endline (Sep 2015)

EGRA Subtests and Maximums	Midline 1 - 2013			Midline 2 - 2014			Endline – 2015		
	Full	Med.	Cont.	Full	Med.	Cont.	Full	Med.	Cont.
Grade 2									
Letter Recognition (100 clpm)	9.5	7.4	4.7	19.9	17.2	5.4	10.7	7.2	5.9
Familiar Words (30 cwpm)	1.9	1.0	0.8	3.3	2.6	1.1	2.4	1.5	1.3
Reading Fluency (120 cwpm)	2.2	1.0	1.0	5.2	4.2	1.7	3.4	2.4	2.0
Reading Comprehension (4 items)	.07	.02	.02	.21	.12	.03	.25	.22	.14
Grade 3									
Letter Recognition (100 clpm)	16.6	15.3	12.0	29.6	27.8	12.3	20.1	19.2	12.6
Familiar Words (30 cwpm)	3.9	3.3	2.8	8.0	6.0	3.2	5.7	4.8	3.3
Reading Fluency (120 cwpm)	5.3	4.4	4.3	14.6	12.0	5.2	8.8	7.4	5.0
Reading Comprehension (4 items)	.16	.12	.12	.53	.43	.15	.50	.36	.30

clpm = Correct Letters Per Minute; cwpm=Correct Words Per Minute

‘Better’ than Control or statistically significant, however, still does not mean “Good.” While these results were encouraging, despite the training provided to teachers and to school and pedagogical directors and the influx of materials to improve reading instruction and learning, the great majority of children did not demonstrate the skills on oral reading fluency needed to read with comprehension and were far from reading the number of words per minute that have been established as benchmarks for developing countries such as Mozambique. Although the effects of the intervention were clear, it is not realistic to think that a six-month intervention could bring children close to benchmark levels. ApaL alone is not enough to overcome many of the underlying recognized constraints to quality education that exist in Mozambican schools—high absenteeism rates among students, teachers, and school directors, teachers with limited training in teaching reading, a very short school day, and persistent

<sup>4</sup> Please note that comparisons between Baseline and Endline results are not meaningful. Due to delays in the implementer’s work plan, the Baseline study was not conducted at the end of 2012 as planned but in the beginning of the 2013 school year. That means that 3<sup>rd</sup> graders assessed at Baseline were 2<sup>nd</sup> graders in 2012 who after 2/3 months of vacation were arriving in February 2013 to start Grade 3. Same with 2<sup>nd</sup> graders, who were in fact 1<sup>st</sup> graders returning after the summer vacation to start Grade 2.

shortages in teaching and learning materials.

Among Grade 2 students there was no statistically significant performance difference between boys and girls in Full treatment schools. In the Medium and Control groups significant differences were found on the Letter Recognition and Familiar Word Reading subtests. Although individual subtest scores were not always significant, overall there is a clear tendency for favoring boys. In none of the 6 subtests across the three groups do girls outperform boys and this gap increases with greater subtest complexity. In Grade 3 these patterns crystalize. In all six subtests in Control schools, girls lag behind reaching about 60% of boys' performance on reading skills. In Full and Medium schools although boys outperform girls on four out of the six subtests, the differences favoring boys are much smaller than in Control schools. Table 0-2 compares scores obtained by boys and girls in three EGRA subtests: Letter Recognition, Familiar Word Reading and Fluency.

Table 0-2. Differences in performance of girls and boys by treatment group and grade (2015)

EGRA Subtest	Student Sex	SECOND GRADE			THIRD GRADE		
		Means by Treatment Group			Means by Treatment Group		
		Full	Medium	Control	Full	Medium	Control
Letter Recognition (clpm)	Girls	10.0	6.1	4.5	17.6	16.7	9.9
	Boys	11.4	8.4	7.4	22.7	21.7	15.7
	Sig. t-test	0.305 NS	<b>0.030</b>	<b>0.005</b>	<b>0.007</b>	<b>0.004</b>	<b>0.000</b>
	Girls/Boys	87.7%	73.3%	60.8%	77.7%	76.8%	62.8%
Familiar Words Read Correctly (cwpm)	Girls	2.2	1.1	1.0	4.8	4.4	2.5
	Boys	2.7	1.8	1.6	6.6	5.3	4.2
	Sig. t-test	0.241 NS	<b>0.015</b>	0.066 NS	<b>0.011</b>	0.128 NS	<b>0.003</b>
	Girls/Boys	81.9%	61.7%	63.8%	71.9%	81.8%	59.7%
Text Reading Fluency (wpm)	Girls	3.0	2.0	1.8	7.4	6.3	3.8
	Boys	3.8	2.8	2.3	10.3	8.5	6.3
	Sig. t-test	0.214 NS	0.137 NS	0.340 NS	<b>0.028</b>	<b>0.043</b>	<b>0.010</b>
	Girls/Boys	78.6%	73.6%	80.0%	71.7%	74.0%	60.7%

Thus IBTCI concludes that the ApaL program contributed significantly to reducing the increasing performance gap between boys and girls in Grade 3 and that there may be other factors that must be addressed in order to eliminate the continuing disadvantage of girls in school performance. We should also note that most of the Grade 3 students received support from ApaL in 2014, which explains the higher performance of boys and girls in the treatment groups when compared with that of boys and girls in Control schools.

On every single comparison of results from the EGRA, students in urban schools strongly outperformed their rural counterparts as shown in Table 0-2.

Table 0-3. Comparison of 2015 Grade 3 EGRA Mean Scores by Urban/Rural Locality

EGRA Subtest - Grade 3	Urban	Rural
Oral Comprehension	9.7	7.6
Concepts about Print	7.4	5.8
Letter Recognition (clpm)	25.2	14.4
Familiar Words Read Correctly (wpm)	7.2	3.7
Text Reading Fluency (cwpm)	10.8	5.7
Text Reading Comprehension	0.58	0.32

The Student Assessment Coversheet identifies some of the factors that could be related to these performance differences.

Table 0-4. Variables that affect performance of students in rural settings

Other Variables	Urban	Rural
Class Days Offered in July	20.9	11.8
Student Days Attendance in July	11.9	8.3
Frequent Portuguese use with Parents (%)	73%	48%

Given these large urban-rural differences, IBTCI also examined the relative effects of each of the two interventions within Urban and within Rural schools in 2015. In Urban schools Grade 2, we detected no performance differences between intervention and control groups. In Grade 3 small (12% – 14%) but statistically significant differences were observed only on the Oral Comprehension and the Concepts about Print subtests.

In Rural schools, on the other hand, large differences, of 50% or more, were found between the two intervention groups and the Control schools. In Grade 2 Full schools outperformed Control on all six subtests often by 85% or more—Medium schools did not outperform Control on any of the subtests. In Grade 3 both Full and Medium groups outperformed the Control group on Concepts about Print, Letter Recognition, Familiar Word Reading, Reading Fluency and, for Full treatment, also on Reading Comprehension.

But even then, Full treatment schools in rural areas generally do not perform as well as Control schools in Urban areas. From the data available, we see that Full treatment Rural schools had a significantly higher (50% more) number of student days attended; in Urban schools no differences were found between any of the groups.

### What contributed to this achievement?

The RCT methodology utilized allows for direct attribution of the USAID/ApaL intervention to improved outcomes because the random sampling component eliminates the effects of potential unobservable differences between treatment and control groups on the outcomes. This allows one to conclude that being in a school receiving the ApaL reading program is the only significant factor that explains the differences between the performance of students in treatment and in Control schools. Thus, the significant statistical differences between scores observed between 2013 and 2014 are a result of the ApaL program.

Due to delays in start-up, the intervention was implemented for two months in the 2013 school year. Full implementation only occurred in 2014. ApaL was designed to integrate into Grades 2 and 3 classrooms of treatment schools major inputs that have been shown effective in improving reading scores: increased time to read, appropriate and sufficient teaching and learning materials, and improved methods to teach reading.<sup>5</sup> While both the Medium and the Full treatment schools received these inputs, a component specifically designed to improve school management processes and routines that supported these initiatives was added to the Full treatment schools. The reading program includes four major result areas related to early grade reading: strengthening the teaching of reading, improving teacher performance, increasing the availability of teaching and learning materials (TLAs), and enhancing school management processes to support reading instruction.

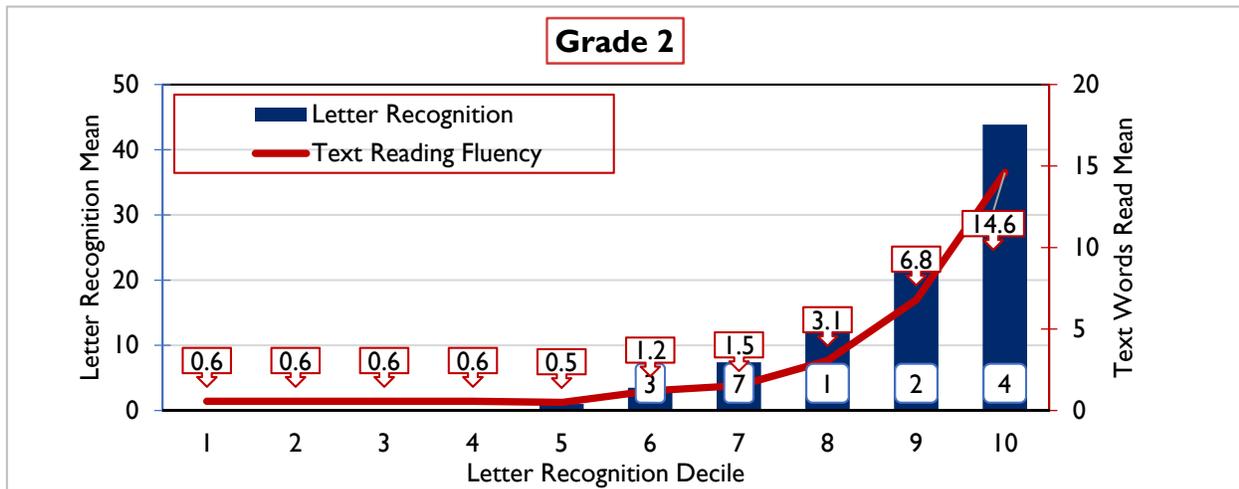
The results obtained allow us to conclude that by addressing fundamental reading skills such as phonological awareness and alphabetic principle in the early grades, in a systematic and sustained way, ApaL can and does improve student reading performance. Figure 0-1 shows the results of the quantitative analysis confirming that children who succeed in basic skills such as letter recognition and

<sup>5</sup> Gove, A. and P. Cvelich (2011). *Early Reading: Igniting Education for All. A report by the Early Grade Learning Community of Practice. Revised Edition. Research Triangle Park, NC: Research Triangle Institute [www.eddataglobal.org](http://www.eddataglobal.org)*

familiar word reading have higher oral reading fluency that allows them to read with comprehension—the ultimate objective of reading.

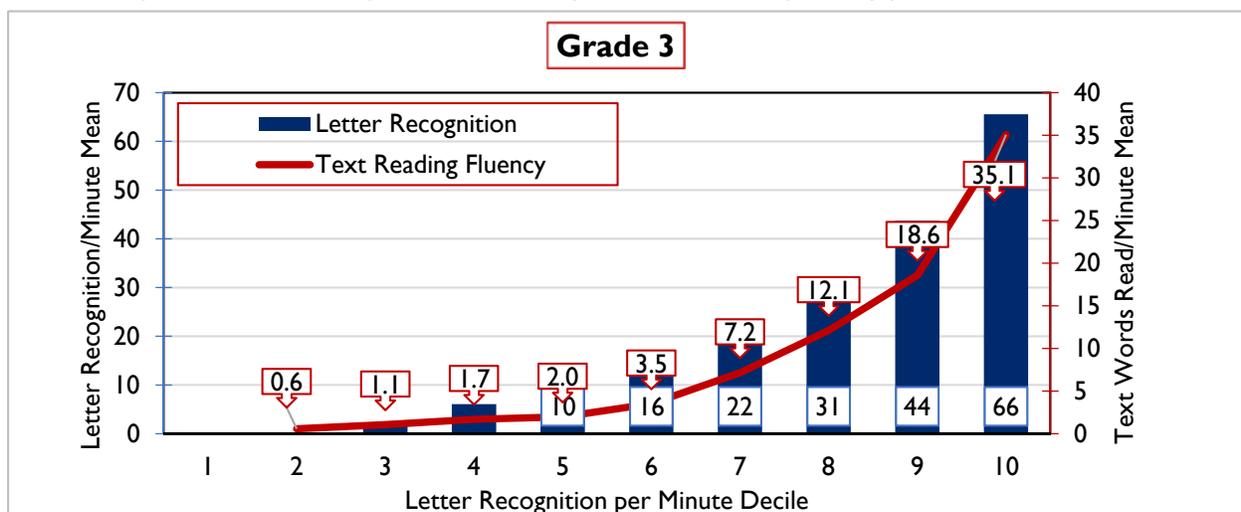
In Grade 2, for example, Letter Recognition is highly associated with Reading Fluency. The correlation between Letter Recognition and Text Reading Fluency is quite high ( $r = 0.65$ ), which means that 42% of all variance in reading fluency is explained by letter recognition scores alone. As shown in Figure 0-1, extremely low levels of performance on this task mean that few students can recognize enough letters to be able to read, even slowly, words in a text passage.

Figure 0-1 Letter Recognition deciles as a predictor of Reading Fluency performance - Grade 2



This pattern is even stronger in Grade 3, where more students perform higher on both the Letter Recognition and Text Reading Fluency subtests as shown in Figure 0-2. Students who performed in the seventh decile (average 22 cpm) of the Letter Recognition subtest read 7 words per minute. Those students in the top decile (average 66 cpm) read five times as many words—35 cwpm. The correlation between Letter Recognition and Text Reading Fluency is stronger than in Grade 2 ( $r = 0.77$ ), accounting for 59% of variance in reading fluency scores.

Figure 0-2 Letter Recognition deciles as a predictor of Reading Fluency performance - Grade 3



The relationships observed between all the pre-reading and reading skills measured by the EGRA for Grade 3 are displayed in Table 0-5. The information presented leads to the conclusion that (1) pre-

reading skills (Oral Comprehension and Concepts about Print are generally not strong predictors of reading skills (the variance explained ranges from 14 to 20%; (2) Letter Recognition is a strong predictor of Familiar Word Reading (69%), of Text Reading Fluency (59%), and contributes somewhat (30%) to Reading Comprehension; (3) Familiar Word reading is an excellent predictor of Text Reading Fluency (70%) and strongly associated with Reading Comprehension (43%); (4) Text Reading Fluency predicts fully 49% of Reading Comprehension scores.

**Table 0-5. Correlations between Subtests and Percentage of Variance Explained**

Grade 3	Concepts about Print	Letter Recognition	Familiar Word Reading	Text Reading Fluency	Reading Comprehension
Oral Comprehension	<b>r = 0.58</b> (R <sup>2</sup> = 34%)	<b>r = 0.38</b> (R <sup>2</sup> = 44%)	<b>r = 0.38</b> (R <sup>2</sup> = 14%)	<b>r = 0.34</b> (R <sup>2</sup> = 11%)	<b>r = 0.45</b> (R <sup>2</sup> = 20%)
Concepts about Print		<b>r = 0.54</b> (R <sup>2</sup> = 29%)	<b>r = 0.49</b> (R <sup>2</sup> = 24%)	<b>r = 0.43</b> (R <sup>2</sup> = 18%)	<b>r = 0.45</b> (R <sup>2</sup> = 20%)
Letter Recognition			<b>r = 0.83</b> (R <sup>2</sup> = 69%)	<b>r = 0.77</b> (R <sup>2</sup> = 59%)	<b>r = 0.55</b> (R <sup>2</sup> = 30%)
Familiar Word Reading				<b>r = 0.89</b> (R <sup>2</sup> = 70%)	<b>r = 0.65</b> (R <sup>2</sup> = 43%)
Text Reading Fluency					<b>r = 0.70</b> (R <sup>2</sup> = 49%)

The importance of learning letters and their sounds, which facilitate the reading of words in order to reach the fluency necessary to read with comprehension—the ultimate goal of reading—cannot be overstated and must be at the core of any reading program. The inputs provided by ApaL in 2014 had a clear impact on student reading skills as shown by the consistent higher performance observed in the treatment groups, especially in the Full schools. However, scripted lessons, teacher and school director training, coaching and providing TLAs and reading materials are only one aspect of the solution to a persistent reading problem as observed in schools in Mozambique. Our findings suggest that additional efforts need to be made to reduce student, teacher and school director tardiness and absenteeism and to expand the time students spend learning to read in order to profit from interventions such as the ApaL reading program. This is not an easy job given the variety of factors and the many challenges surrounding school absenteeism.

**Are the effects sustainable after the cessation of the intervention?**

The USAID/ApaL reading program was able to significantly improve foundational reading skills in 2014, and some results persist in 2015 after the intervention has ceased for one year. Table 0-5 reflects the percent of retention of learning (significant differences are **bolded**) obtained at the end of 2014 after one school year of implementation and retained at the end of 2015 one year after ApaL had ceased its involvement at the schools selected for the IE sample.

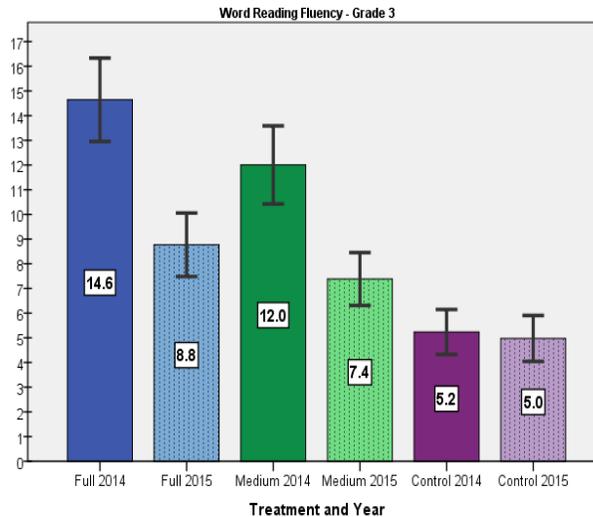
**Table 0-6. Results obtained in 2014 vs. 2015**

EGRA Subtests 2015 as % of 2014	Grade 2			Grade 3		
	Full	Medium	Control	Full	Medium	Control
Oral Comprehension	<b>92%</b>	<b>90%</b>	95%	<b>92%</b>	<b>64%</b>	98%
Concepts about Print	<b>88%</b>	<b>80%</b>	94%	<b>89%</b>	<b>92%</b>	97%
Letter Recognition	<b>54%</b>	<b>42%</b>	109%	<b>68%</b>	<b>69%</b>	102%
Familiar Word Reading	<b>71%</b>	<b>55%</b>	109%	<b>70%</b>	<b>74%</b>	102%
Text Word Reading	<b>65%</b>	<b>57%</b>	118%	<b>60%</b>	<b>61%</b>	95%

Figure 0-3 illustrates the sustainability of 2014 scores for text reading fluency in Grade 3, by treatment group. In the treatment groups, scores significantly declined in the no-intervention year: 60% of 2014

scores remained in 2015 for Full school students; 61% for Medium. Control groups did not change. While Full schools maintain a lead over Control schools, Medium schools have become almost indistinguishable from Control schools.

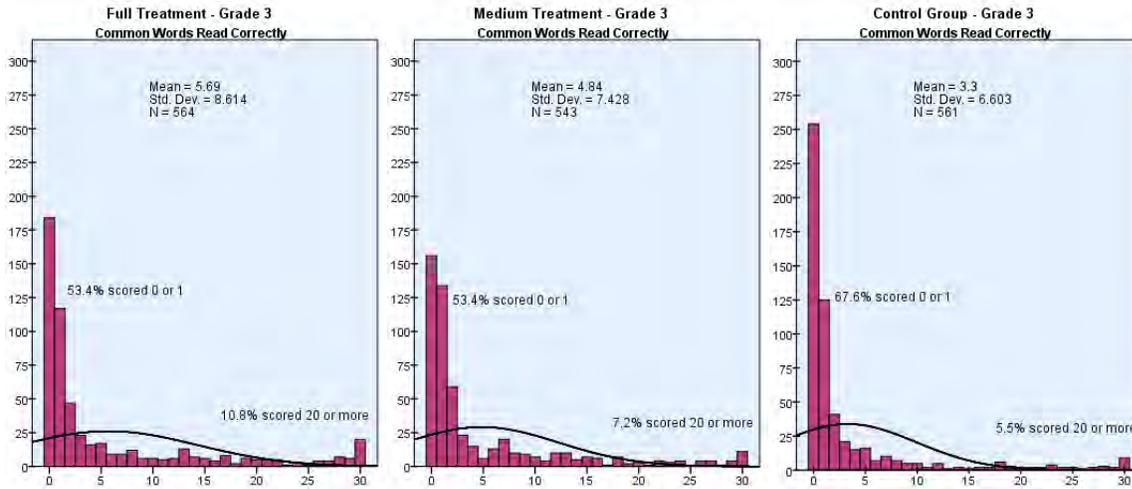
Figure 0-3 Text Reading Fluency by treatment group and year – Grade 3



Although as a result of the cessation of the program scores went down—as expected—on all EGRA subtests in both grades, the Full treatment schools continued to significantly out-perform Control schools in both Grade 2 and Grade 3. Medium schools generally did not perform better than Control schools in Grade 2, but did show better results relative to the Control schools in Grade 3 on four of the six EGRA subtasks. Grade 3 Medium schools were much more similar to the Full schools. The reason Grade 3 Medium treatment students performed better than Grade 2 students vis-à-vis Control students is probably due to the fact that most of Grade 3 students and most of their teachers would already have experienced a full year of ApaL intervention the year before. In 2015, Grade 2 students (except those who repeated) were not directly exposed to the ApaL program.

In short, from a technical standpoint, the intervention benefits students both at the conclusion of the intervention and also, to a lesser extent, one year after the intervention has ceased and schools have been functioning without further support. However, note that scores on the sub-test Text Reading Comprehension, the ultimate goal of learning to read, continue to be extremely low, as reading fluency remains a serious limitation for almost all students. Research shows that there is a strong correlation between fluency and reading comprehension and for this reason Oral Reading Fluency is often used as the best “composite” indicator of the ability to read. This is behind the establishment of the benchmarks of 20 words correctly read per minute at the end of Grade 2 and 40 at the end of Grade 3 for developing countries such as Mozambique. Students who do not reach these benchmarks will not be able to read with comprehension. Figure 0-4 shows the distribution of scores per treatment group and the percentage of students who scores 0-1 or 20 or more words correctly read per minute.

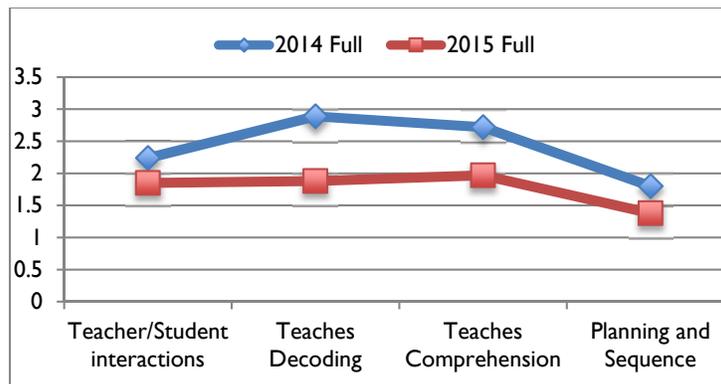
Figure 0-4. Distribution of Familiar Words Correctly read by minute per treatment group



Correct words per minute (cwpm)—both isolated words and words in a connected text—are the key skills to acquire in order to read with comprehension. This is where the efforts should be put if the goal is getting children to learn to read (which implies comprehending what you read). Note that in 2015 only 10.8% of third graders assessed in Full treatment schools, 7.2% of those in Medium and 5.5% of those in Control schools were able to read 20 or more correct words per minute, and even in treatment groups more than half of the children scored zero or read only one word per minute. Even though the differences between the groups are statistically significant, the real educational significance is doubtful when so many students score zero, even third graders that were part of the program in 2014.

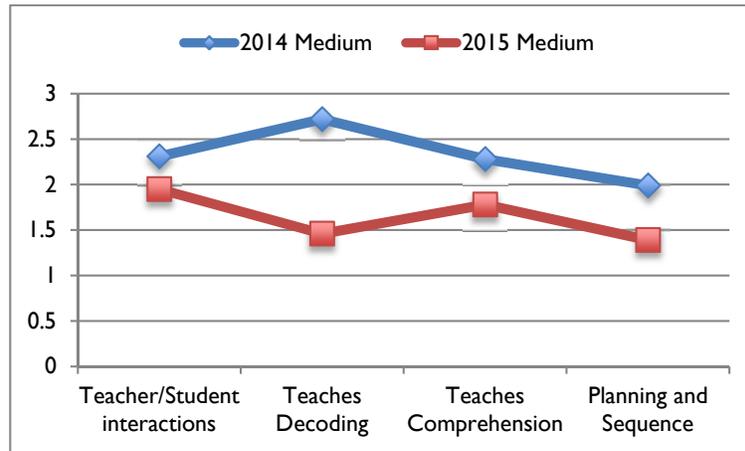
As noted, the treatment group advantages versus the Control group were not as great in 2015. About 70% of the teachers who had been trained in ApaL practices were retained at the same school teaching either Grade 2 or 3. Of these, they did not, for the most part, carry these practices on in 2015. Figures 0-5 and 0-6 contrast teacher observed instructional performance in 2014 and 2015 along four instructional behaviors measured by the SMA.<sup>6</sup>

Figure 0-5. Teacher observed practice by year and treatment (Full)



<sup>6</sup> Data provided to the IE by ApaL M&E with descriptive analysis. The IE conducted further analyses to reach a composite score that allow for multiple comparisons.

Figure 0-6 Teacher observed practice by year and treatment (Medium)



Training provided to Full school directors combined with the greater stability of school directors in their position contributed to the larger impacts observed in Full schools. Student reading material possession and classroom teaching and learning aid (TLA) materials dropped from 50% to 60% of their 2014 levels. Nonetheless, enough early grade reading gains remained throughout the two-year period so that, combined with the reduction of unit costs, the cost-effectiveness of the ApaL interventions improved relative to that reported for the 2014-only school year.

Over the two-year period, Full schools show stronger cost-effectiveness results relative to the Medium treatment, reconfirming that the best model of intervention is the Full treatment model. In 2014, the two interventions were examined from a cost-effectiveness perspective: dividing the unit cost per student by the gains observed on each EGRA subtest. To extend the cost-effectiveness analysis to cover the 2015 year required spreading the 2014 costs over two school years, recalculating the beneficiaries (enrollment totals) and developing new unit costs. These unit costs were significantly reduced from \$11.54 for Full treatment in 2014 to \$6.04 in the combined 2014-2015 period. Medium treatment per unit costs declined from \$9.13 to \$4.75. This combined with the gains observed reaffirms the cost-effectiveness of the ApaL program.

## Summary Recommendations

Many of the basic recommendations made in the Midline 1 and 2 reports are still valid. Specifically, (a) increasing time spent learning basic literacy skills, (b) teaching sounds of all letters of the alphabet, (c) providing students with books that focus on decoding and word-building skills, training teachers to use these strategies in the classroom, (e) putting into place strategies to encourage students and their families to increase the number of words read, and (f) especially providing school directors with the training needed to improve their schools. There are challenges to be faced when attempting to modify behaviors and procedures that have been in place for a long time but ApaL is making an effort to implement the recommendations.

Recommendations 1, 2, 3, 4, and 10 are primarily for USAID to consider in conjunction with its general project design and management activities. The other recommendations are primarily for USAID and MINEDH dialogue and consideration.

**I. Make sustainability part of the implementation plan of an intervention.** Sustainability should not be an afterthought to be addressed at the end of an intervention. A specific description of the measures that will be taken to make the intervention sustainable should be required as an important element of the design of a project. ApaL has worked with district officers and school directors to make

sure that trained teachers remained in their schools and classes, to ensure that skilled trainers are located in each ZIP/district, and that teaching and learning materials continue to be used.

**2. Strengthen project Monitoring and Evaluation (M&E).** A well-developed Monitoring and Evaluation (M&E) component with specific and measurable indicators should be required from implementers as part of the project design, independent of the external evaluation of project results and impact. In addition to the monitoring of project activities and other inputs, greater focus on the achievement of outcomes represented by well-defined and agreed upon indicators will enhance the implementation of a project and provide information to correct its course when necessary.<sup>7</sup>

**3. Adopt a non-linear implementation model to provide support more or less intensely as required in different situations.** Rather than follow an implementation-and-immediate-scale-up model, interventions such as ApaL may require a more extended period of support to allow for internalizing and routinization of activities. Withdrawing the support according to a fixed schedule (e.g., the end of one school year) without considering the level of routinization achieved may significantly decrease both the effects and the sustainability of an intervention.

**4. Assess the level of readiness for the innovation prior to the implementation of the intervention.** A range of instruments has been developed and used to identify specific concerns of potential adopters and stakeholders. These instruments provide insights into issues that must be addressed to ensure widespread acceptance, adoption and sustainability of an intervention.

**5. Implement the program where the effects are greater.** Findings suggest that the program functions better in some settings than in others. For example, on an absolute basis, children in rural schools consistently show lower scores than do children in urban schools—in fact, on average, students in Full treatment rural schools performed worse than students in urban Control schools. This notwithstanding, our findings indicate that, in 2015, ApaL’s effects in rural schools are larger 2015 the effects observed in urban settings. That is, although ApaL does not eliminate the relative deficit in learning for rural children, it significantly reduces it. The findings suggest that the program, as designed and implemented, may be more appropriate for rural schools, and that in the future, those schools should become ApaL’s main target.<sup>8</sup>

**6. Identify the reasons for girls’ consistent under-performance and include in all projects strategies to close the gender gap.** Both Full and Medium ApaL interventions reduced the difference in performance between boys and girls, when compared to Control schools in the same grade. The program benefited both boys and girls equally but has not eliminated the gender gap observed in 2013 and 2014. Findings show that while ApaL narrows the relative gap between boys and girls, the absolute gap widens as students advance to Grade 3 and when more complex skills are assessed. The increased magnitude of sex differences in Grade 3 suggests that greater efforts will be required to address the root causes of under-performance of girls.

**7. Work with MINEDH to improve the ways school directors are selected and prepared.** School directors are key to educational improvement. This is shown first by the higher scores of students in Full treatment schools, which is probably closely associated with greater attendance by teachers and students both, and confirmed by the finding that scores were higher at schools where school directors received additional training in 2015. We recommend that ApaL work with MINEDH and district leadership to identify staff with leadership potential to become a school director and

---

<sup>7</sup> Indicators for project outcomes need to be specific, measurable, attainable, relevant and time-bound. For example, *Improving reading skills* or *Expanding time available to teach reading* do not qualify as SMART indicators. Indicators developed during the design phase can be adjusted, if necessary, based on the findings obtained at Baseline.

<sup>8</sup> ApaL has indicated that, with the exception of treatment schools in Nampula City and Quelimane, all schools in the project 2015 and 2016 scale-up are rural schools.

develop a clear set of criteria for applicants. Second, because many principals learn the skills they need on-the-job, they need continuous in-service opportunities learn how to improve school management.

**8. Provide incentives to keep trained teachers at the school teaching the early grades.** The management component received by Full treatment schools contributed to a higher rate of retention of trained teachers in those schools. Nonetheless, approximately 30% of the teachers trained in 2014 either left the school or were assigned to other grades. This highlights the need to develop and strengthen incentives to keep trained teachers in the same grades in subsequent years.

**9. Use more effectively the data produced by the School Management Assessment (SMA) instruments, checklists, and assessments.** This would require the reduction of available data to rate schools in simple categories of school management aspects. The “grades” assigned would show clearly where a school stands in terms of factors that relate to student performance such as student, teacher and staff attendance, days of class offered, or hours of instruction per day. This will reinforce the use of data to identify and monitor school management factors causal to student performance. This type of exercise should also form part of the MINEDH school supervision process.

**10. Consider the sustainability of the various inputs provided as part of the program when designing future projects.** Some inputs provided by ApaL, such as TLAs, seem to have only a moderate level of sustainability (50% - 60%) while others were even less sustainable. For example, given the amount of teacher training conducted, only decoding activities showed any difference in the classroom a year after implementation. All inputs require on-going assistance in order to continue. We must identify ways to improve the sustainability of the inputs themselves and ensure that necessary funding will be available to maintain them. This needs to be a part of the intervention design.

The body of the report presents more details on the context, the intervention, the RCT methodology. Findings on the results of the intervention are followed by conclusions and recommendations.

U.S. Agency for International Development  
1300 Pennsylvania Avenue, NW  
Washington, DC 20523