EDUCATION CANNOT WAIT: EDUCATION FIRST PROJECT

AN ENDLINE REPORT FOR EARLY GRADE READING ASSESSMENT (EGRA) IN PALABEK REFUGEE SETTLEMENT, LAMWO DISTRICT

NOVEMBER, 2019

LUIGI GIUSSANI INSTITUTE OF HIGHER EDUCATION

EDUCATION CANNOT WAIT
a global fund for education in emergencies

AVSI
People for development

ACKNOWLEDGEMENTS

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

APPCO	African Partners for Child Poverty
CRS	Catholic Relief Services
DID	Difference in Differences
DIDE	Difference in Differences Estimate
ECW	Education Cannot Wait
EGM	Early Grade Mathematics
EGMA	Early Grade Mathematics Assessment
EGR	Early Grade Reading
EGRA	Early Grade Reading Assessment
FCA	Finn Church Aid
HI	Humanity and Inclusion
LGIHE	Luigi Giussani Institute of Higher Education
NRC	Norwegian Refugee Council
QED	Quasi Experimental Design
RTI	Research Triangle Institute
TICC	Teachers in Crisis Context
UNESCO	United Nations Educational, Scientific and Cultural Organization
WIU	Windle International Uganda

EXECUTIVE SUMMARY

Introduction

AVSI Foundation and Luigi Giussani Institute of Higher Education (LGIHE) have been piloting Early Grade Reading (EGR) in World View Nursery and Primary School in Palabek refugee settlement, Lamwo district. The EGR intervention was based on Jolly Phonics integrated in Weekly Foundation Story. The other school – Awich Nursery and Primary School was used as a control to facilitate the evaluation design. In order to assess the effectiveness of the EGR intervention, an endline was conducted in the stated schools.

Objectives

The EGRA was conducted to specifically (i) assess the learners' early grade reading skills in the classes of P1, P2 and P3; and (ii) ascertain the level of improvement in learners' reading abilities as a result of the EGR intervention.

Methods used

A quasi experimental design, where measurements was done in the 2 schools both before and after exposure to the intervention, was adopted to assess the effectiveness of the EGR intervention. During the baseline, a total of 339 learners from P.1 to P.3 of the 2 schools were assessed (treatment=170 learners; control=169 learners). Thereafter, an intervention was implemented with 12 teachers of the treatment school. After that, an endline was conducted with a total of 297 learners whereby in the treatment school a total of 157 learners (P1=47, P2=56, P3=54) were assessed while the rest 140 (P1=45, P2=41, P3=54) were assessed in the control school.

The EGRA tool used was adapted from the Research Triangle Institute (RTI) for purposes of validity, reliability and comparability. This tool included the following subtask areas: Letter Sound Knowledge, Segmenting, Nonword Decoding, Oral Passage Reading and Reading Comprehension, and English Vocabulary (body parts and words from the environment). The assessments were conducted in pairs by trained Assessors - a teacher of a respective school and an AVSI/LGIHE staff. The endline was conducted from 28th October to 1st November, 2019.

This being a paper-and-pencil assessment, learners' scores were captured into an Epidata (Version 3.1) system and exported to STATA (Version 13.0) for cleaning and analysis. The analysis was done in terms of proportions, percentages and means. Furthermore, in order to investigate the casual effects of the EGR intervention on the treatment school over the control school, difference in differences (DID) analysis was conducted.

Summary of Findings

i. Subtask 1: Letter Sound Knowledge
   - The project improved the letter sound knowledge in the treatment than control school, in terms of percentage of learners who sounded correctly at least 1 letter per minute (DIDE=67%).
   - The project improved the letter sound knowledge in the treatment than control school, in terms of the mean number of letters sounded correctly per minute by the learners (DIDE=10.9 letter sounds).

ii. Subtask 2: Segmenting
   ▪ The project improved the segmenting knowledge in the treatment than control school, in terms of percentage of learners who could correctly tell all the sounds of one or more words (DIDE=64%).
   ▪ The project improved the segmenting knowledge in the treatment than control school, in terms of the mean number of words correctly segmented by the learners (DIDE=1.9 words).

iii. Subtask 3: Nonword Decoding
   ▪ The project improved the nonword decoding knowledge in the treatment than control school, in terms of the percentage of learners who read correctly at least 1 word per minute (DIDE=13%).
   ▪ The project improved the nonword decoding knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the learners (DIDE=0.7 words).

iv. Subtask 4: Oral Passage Reading and Reading Comprehension
   ▪ The project improved the oral passage reading knowledge in the treatment than control school, in terms of the percentage of learners who read correctly at least 1 word per minute (DIDE=3%).
   ▪ The project improved the oral passage reading knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the learners (DIDE=1.4 words).
   ▪ Only 0.3% of the learners assessed managed to answer at least one out of five questions in the oral passage reading and comprehension subtask.

v.   Subtask 5: English Vocabulary
   ▪ The project improved the English vocabulary knowledge in the treatment than control school, in terms of the percentage of learners who could identify at least 50% of the English vocabulary words (DIDE=15%).
   ▪ The project improved the English vocabulary knowledge in the treatment than control school, in terms of the mean number of English vocabulary words that were correctly identified by the learners (DIDE=1.2 body parts and objects).

Conclusion

The results reveal improvements in the treatment school compared to the control over the period of 4 months of EGR intervention. Specifically, significant improvements were observed in the subtask areas of: letter sound knowledge, segmenting knowledge, and nonword decoding. A debrief with teachers reveals these improvements to be attributed to the EGR intervention that was implemented in the treatment school, immediately after the baseline assessment.

Based on the findings of this pilot, it is important to increase on the length of the intervention in order to allow teachers to fully comprehend and effectively implement the knowledge and skills acquired. Moreover, there is also need to equip teachers with knowledge and skills of creating authentic and appropriate teaching materials for enhancing early grade literacy skills in refugee contexts in Uganda.

CHAPTER ONE: INTRODUCTION

1.1. Background to the Assessment

Education Cannot Wait (ECW) is a one-year project being implemented by AVSI Foundation as a lead implementer and Luigi Giussani Institute of Higher Education (LGIHE), as a co-implementer. The ECW project is being implemented in a consortium of 13 implementing partners, including: APPCO, AVSI, CRS, FCA, HI, Plan International, NRC, Save the Children, Street Child, UNESCO, WIU and ZOA. However, the activities of this project which LGIHE is directly involved in include; Adaptation and development of training materials, and training of teachers using the adapted Teachers in Crisis Context (TiCC) materials.

AVSI and LGIHE have been piloting Early Grade Reading (EGR) in one of the schools (World View Nursery and Primary School) in Palabek refugee settlement, Lamwo district. The second school (Awich Nursery and Primary School) was used as a control to facilitate the evaluation design. LGIHE's interventions are expected to contribute to the achievement of the following outcomes:

- **Outcome 1**: Improved Equitable Access to Inclusive Relevant Learning Opportunities
- **Outcome 2**: Improved Delivery of Quality Education and Training.

1.2. Description of the EGR Intervention

The EGR intervention was based on Jolly Phonics integrated in Weekly Foundation Story. Jolly Phonics is a practical and fun method of teaching letter sounds to children using interesting games and activities. The children first learn the letter sounds and are then taught to read words by 'synthesising, or blending, the sounds together. This method allows children to work out unknown words for themselves, rather than being asked to memorise them. This is an important step towards independent reading. Jolly phonics comprises 42 sounds of the English language divided into seven sets that are taught to children in the order of simplicity to complexity. Every sound is accompanied by fun activities to enable the children sound it properly and remember it. This method also helps children to learn the digraphs and be able to read words fluently. Jolly phonics promotes fluency in speaking and reading as children are exposed to the different ways the letter sounds are produced. With the ability to produce the sounds, the children are also able to recognise the sounds and words produced. This builds their ability to write words with those sounds correctly; hence developing their spelling skills.

The WFS approach is an approach where teachers use stories from textbooks or those that they have formulated based on the themes they teach in the respective classes. This approach focuses on reading, writing and comprehension in all classes and subjects. The learners read a story based on a given theme, answer questions about it and either write sequels to the story or draw pictures of what they like, dislike, or fear about the story. This is to foster learners' comprehension. The WFS integrates the teaching of sounds and the Whole Word. While teaching a given sound or set of sounds, the teacher can create a story with a number of words that have the sound that s/he is targeting to teach in that lesson. The teacher first reads the story for the pupils, then lets them read it, and later draws their attention to the words with the sound that s/he wants to teach in that lesson. The teacher must bear in mind that the story remains within the topic of the day or week.

These approaches can be used in any context including resource constrained, crisis and emergency. It supports the application of multicultural and multilingual pedagogies. Hence, in the refugee context, such as in Uganda, this approach blends well into any curriculum and fosters creativity and innovation among teachers through their involvement in developing of learning materials from within their environments.

This intervention was implemented in a series of trainings and follow-up sessions as follows:

✶ 1st Training

The first training was conducted in a period of 2 days. This training was intended to respond to the observations made during the pre-intervention assessment and daily debriefs with teachers. It covered the following:

- Training on:
  - Child development milestones,
  - The literacy processes,
  - Introduction to strategies for teaching,
  - Introduction to jolly phonics,

✶ 1st Follow-up Training

The training was conducted in a period of 5 days: It covered the following:

- Training on:
  - The meaning of a phoneme, grapheme, phonetics, phonics, phonological and phonemic awareness,
  - The levels of phonemic awareness complexity
  - Using Jolly Phonics
  - Exploring strategies for teaching phonemic awareness
  - How blending and segmentation greatly aid reading and spelling
  - Learning the importance of connecting phonemic awareness to phonics and systematic ways to strengthening sound/symbol relationships
- Creation of Community of practice (CoP): This is a platform where teachers share ideas on how to improve classroom practices based on the knowledge and skills acquired from the EGR intervention. COP meetings were held weekly.

✶ 2nd Follow-up Training

The follow-up was conducted in a period of 5 days: It covered the following:

- Training on:
  - Scheming
  - Lesson planning
  - Material creation and use
  - Mode of teaching
  - Record of work
  - Assessment for learning

★ 3rd Follow-up Training

The follow-up was conducted in a period of 4 days: It covered the following:

- Classroom observations
- Community of practice.

Based on this, AVSI and LGIHE organized to conduct an Endline for EGRA in the 2 schools (World View Nursery and Primary School and Awich Nursery and Primary School). The endline is the post-intervention assessment whose results are compared with the baseline (as of June/July, 2019) in order to assess the effectiveness of the EGR interventions in Palabek refugee settlement.

1.3. Purpose of the Assessment

EGRA is aimed at assessing the effectiveness of the EGR interventions in Palabek refugee settlement.

1.4. Objectives of the Assessment

The EGRA was conducted to address the following objectives:

i. To assess the learners' early grade reading skills in the classes of P1, P2 and P3.
ii. To ascertain the level of improvement in learners' reading abilities as a result of the EGR intervention.

1.5. Assessment Hypotheses

The EGRA was guided by the following hypotheses:

- *Ha_1*: We anticipated that more learners in the treatment group will sound correctly one or more letters in 60 seconds compared to those in the control group, after the EGR intervention.
- *Ha_2*: We anticipated that more learners in the treatment group will tell correctly the sounds of one or more words in the list compared to those in the control group, after the EGR intervention.
- *Ha_3*: We anticipated that more learners in the treatment group will read correctly one or more words per minute compared to those in the control group, after the EGR intervention.
- *Ha_4*: We anticipated that more learners in the treatment group will read correctly one or more words per minute in the passage compared to those in the control group, after the EGR intervention.
- *Ha_5*: We anticipated that more learners in the treatment group will identify at least 50% of the English vocabulary words compared to those in the control group, after the EGR intervention.

CHAPTER TWO: ASSESSMENT METHODOLOGY

2.1.   Assessment Tool

The EGRA tool developed by RTI was adapted to ensure issues of validity, reliability and comparability with existing assessments. The subtask areas included in this assessment are: Letter Sound Knowledge, Segmenting, Nonword Decoding, Oral Passage Reading, Reading Comprehension and English Vocabulary (body parts and words from the Environment).

The tool was later translated into Acholi language to suit the demands of the thematic curriculum but also to cater for the diversities in language for the learners in the settlement. It should be noted that only the instructions were translated, and at some points some translations had to be made in Arabic to cater for the few learners who could not understand either English or Acholi.

2.2.   Target Population

The target population consisted of learners in P1, P2 and P3 from 2 schools in Palabek refugee settlement, Lamwo district, as shown in Table 1.

2.3.   Assessment Design

To establish a causal relationship between the intervention and changes in outcomes, the assessment adopted a Quasi Experimental Design (QED) in which measurements were done in the 2 groups (treatment and control schools) both before and after exposure to the programme. The schools were grouped as follows:

- **G1**: World View Nursery and Primary School, which was the treatment school.
- **G2:** Awich Nursery and Primary School, which was the control school.

2.4.   Externalities and Selection Criteria for the Schools

Experimental interventions may generate spill-over effects whenever untreated learners are affected by the treatment programme. The study is cognizant of spill-over effects and recognizes the difficulty of its quantification. Attempts were made to ensure that selected learners for the group were not from the same school; and that the selected school for the control was not from the same zone. This was to minimize possibilities of spill-over effects. Therefore, the schools were selected based on the following conditions:

- Being on the double shift system
- Using temporary structures for classrooms
- Enrolled learners from various tribes hence multi-lingual teaching demands
- Being located in different zones to minimize issues of spill-over effects
- Recruited both national and refugee teachers
- Enrolled both host and refugee children with the majority being refugees

2.5.   Sample Size and Number of Learners Assessed

The following records as realized from the head teachers of the respective schools in the month of May/June, 2019, were used in determining the desired sample size at baseline level.

Table 1: Population per school and class

School	Type of Assessment	P1	P2	P3	Total
World View Nursery and Primary School	EGRA	105	120	180	405
Awich Nursery and Primary School	EGRA+ EGMA	245	192	307	744
Total					1,149

In reference to the sample size for the Early Grade Mathematics Assessment that was conducted in the same settlement, the minimum desired number for EGRA was also set to 50 learners per class per target school. Note that EGRA and Early Grade Mathematics Assessment were conducted at Awich Nursery and Primary School as a control school, but with a separate group of learners. This therefore necessitated dividing the learners into 2 mutually exclusive groups.

At baseline, 339 learners were assessed in all the 3 classes. At endline, all these learners were targeted, however due to some unavoidable circumstances like transfer cases and absenteeism, some learners were not found during the data collection dates. At the end of the endline data collection exercise, 297 learners were assessed. The characteristics of these learners and the extent of attrition are explored in the **methodology sections 3.1** and **3.2**. The number of learners assessed at baseline and endline per group (school, sex) and class were as below.

Table 2: Number of Learners Assessed Per Group and Class

School	Class	Baseline						Endline		
		Sex		Nationality				Sex		
		Male	Female	UG	SS	Total		Male	Female	Total
World View Nursery and Primary School	P1	23	28	4	47	51		22	25	47
	P2	31	30	0	61	61		28	28	56
	P3	35	23	0	58	58		32	22	54
	Total	89	81	4	166	170		82	75	157
Awich Nursery and Primary School	P1	34	22	0	56	56		27	18	45
	P2	22	32	0	54	54		12	29	41
	P3	33	26	2	57	59		28	26	54
	Total	89	80	2	167	169		67	73	140

2.6.   Selection of Learners

The sampling of the learners was done by LGIHE officials at baseline level. This was done through simple random sampling having obtained a list of learners for each class, from the Head teachers. Note that sampling of learners was done per target class. At endline, only learners assessed at baseline were assessed.

2.7.   Training of Assessors

The same teachers who participated in the baseline assessment were invited and retrained to carry on the endline assessment. The refresher training of assessors was conducted on the 26th/10/2019 at World View Nursery and Primary School. This training comprised of 21 participants from World View Nursery and Primary School (9 teachers) and Awich Nursery and

Primary School for EGRA (12 teachers). Three teachers from the treatment schools did not manage to attend the training as well as the assessment due to some unavoidable circumstances. The training was facilitated by 3 experienced facilitators from LGIHE who included: The Senior Monitoring and Evaluation Officer - for the aspects of assessment methodology; the Deputy Principal, who was assisted by the Education Officer - for the aspects of English language and tasks included in the tool.

The training was intended to make assessors become familiar with the administration of the tool and with the specific implementation and coding practices associated with it. The training covered the following aspects:

- Self-introductions, expectations and setting ground rules.
- Assessment protocol – about ECW project, the assessment purpose, assessment methodology; and data quality control issues.
- Key points to consider when conducting the assessment – seeking consent; ensuring confidentiality, safety and security, and psychosocial wellbeing of others.
- EGRA tool – question by question explanations.
- Mock assessment for the EGRA tool – practice among assessors.

An interactive strategy was utilized during the trainings whereby participants were taken through the assessment process and tools. The facilitators emphasised how questions in each subtask were to be asked and answered. In this process, assessors practiced reading the instructions and answering the corresponding questions. Any incorrect response was addressed by the facilitators.

Having gone through all the instructions and questions in the respective tools, participants were divided into 3 groups comprising of 7 teachers and 1 LGIHE staff to practice administering the assessment. The participants in a rotational way chose 2 teachers to act as Assessors while 1 teacher acted as a learner. While others were observing and taking note of the proceedings, the 3 participants simulated the assessment process. At the end of each subtask, participants had an opportunity to discuss and correct areas where there were errors.

2.8. The Fieldwork Process

The endline assessment was conducted from 28th October to 1st November, 2019, in the 2 target schools. The assessment was based on a pen/pencil and paper approach. As of the baseline which was conducted from the 24th June, 2019 to 1st July, 2019, teachers were tasked to assess their learners in collaboration with an external member. This meant that each learner was assessed by a pair of Assessors, comprising of a teacher from the respective school and an AVSI/LGIHE staff. This was considered in order to ensure that:

- They become aware of how individualized assessments are conducted
- They become aware of the individual learner challenges hence able to help the child after the assessment and at an appropriate time/level.
- The child feels comfortable being assessed in the presence of their teacher.
- Good quality data is collected as one helps to assist or check on the other.

After the data collection, the test papers were taken to LGIHE offices in Kampala, for data entry. The entry was done by LGIHE experienced data entrants, using a pre-designed system – Epidata

(Version 3.1) that controls for data input errors and safeguards data integrity, and also eases the process of export to analytical packages.

2.9. Data Analysis

Before any data analysis was performed, rigorous cleaning of the data was performed in order to identify erroneous records. Any missing bio-data e.g. sex, nursery attendance, repetition, class, etc. was corrected as per the baseline characteristics.

The cleaning and analysis were done using STATA (Version 13.0) statistical package with the best practice of using 'do-files' which ease the process of code correction. The results were analysed in terms of proportions, percentages, means, and difference in differences estimates (DIDE) for each class and all the classes combined.

In order to investigate the casual effects of the EGR intervention on the treatment school over the control school, difference in differences (DID) analysis was conducted. The DID method removes the difference in the outcome between treatment and control groups at the baseline. It was implemented as an interaction term between the time (where 1 is assigned to the endline period and 0 is assigned to baseline period) and treatment (where 1 is assigned to the treatment group and 0 is assigned to control group) variables in a regression model as below:

$$Y = \beta_0 + \beta_1 * [Time] + \beta_2 * [Treatment] + \beta_3 * [Time * Treatment] + \varepsilon$$

Where; β_0 is the baseline average; β_1 is the time trend in control group; β_2 is the difference between two groups (treatment vs control) at baseline and β_3 is the difference in changes over time.

And with covariates as in the regression model below:

$$Y = \beta_0 + \beta_1 * [Time] + \beta_2 * [Treatment] + \beta_3 * [Time * Treatment] + \beta_4 * [Covariates] + \varepsilon$$

Furthermore, differential impact of the EGR intervention by sex of the learners is examined. This is done by estimating the impact of treatment on the outcome variables at endline by estimating the equation below:

$$Y = \beta_0 + \beta_1 * [female] + \beta_2 * [Treatment] + \beta_3 * [female * Treatment] + \varepsilon$$

2.10. Limitations to the Assessment

Application of sampling weights to reflect the probability of learners sampled when using the difference in differences estimation commands in STATA is not supported. However, the standard errors were clustered at school-class level in order to reflect the notion that, learners were chosen from their respective classes.

Secondly, much as quasi-experimental design was appropriate for this kind of pilot (due to the project duration), it has limitations when the parallel trend assumption is not fulfilled. This assumption requires that in absence of the EGR intervention, then the difference in observed outcomes between the treatment and control groups is constant over time. In other words, we need to ask the question: would the treated learners have experienced the same outcomes

as the control group learners if their teachers did not receive the EGR intervention? This is definitely the hardest assumption to fulfil when there are only 2 data points. The best practice is always to acquire more data points before and after so as to have a visual inspection of this assumption. That said, care must be taken when making conclusions about the program effects as it was not possible to investigate this assumption. Nonetheless, the strongest points for the analyses performed in this assessment are:

- Robust standard errors are used to account for autocorrelation between pre-intervention and post-intervention in the same individual;
- The characteristics of respondents in treatment and control groups were examined, before and after intervention in order to investigate whether the retained or dropped respondents from the 2 groups (treatment vs control) had similar characteristics;
- Furthermore, key variables are included as covariates in the estimation of the project effects.

CHAPTER THREE: PRESENTATION AND DISCUSSION OF THE FINDINGS

3.1.   Implication of Attrition: Baseline-Endline missingness

Considering 339 learners who were assessed at baseline, and merging it with the endline data using "learner's id" as a unique identifier, only 297 learners were matched. This therefore resulted into 42 baseline learners (29 control; 13 treatment) that were not found at endline because they were either transfer cases or absent during the data collection dates.

In order to assess the overall level of attrition among the learners sampled at baseline, balance testing was done and the results are as shown below.

Table 3: Comparison of attrition across the study groups

	Control=29; Treatment=13; Total=42				
	(1)	(2)	(3)	(4)	(5)
Variable(s)	Mean Control	Mean Treatment	Difference	t-value	Pr(T>t)
Age	10.828	10.154	-0.674	0.79	0.4330
Nursery attendance	0.517	0.462	-0.056	0.33	0.7460
Repeated class	0.414	0.231	-0.183	1.13	0.2633
Female	0.759	0.538	-0.220	1.43	0.1612
*** p<0.01, ** p<0.05, * p<0.1					

These results suggest that learners who are missing are not so much different from those that are present, indicating that missingness or attrition wouldn't significantly affect the estimation results. A little detailed look at the results of the balance between the study arms, shows no differences at 5% level. This means that, similar individuals are missing from both treatment and control groups.

3.2.   Baseline Characteristics of the Learners Considered for Further Analyses

In order to match the endline with the baseline findings, only learners assessed at endline were retained for the subsequent analyses. The rest of the learners were dropped. This therefore meant the **297** learners assessed at endline. Their baseline characteristics (used as covariates) are shown below.

Table 4: Baseline Characteristics of the Retained Learners

	(1)	(2)	(3)	(4)	(5)
Variable(s)	Mean Control	Mean Treatment	Difference	t-value	P-value
Primary One (Control=45; Treatment=47; Total=92)					
Age	9.24	8.94	-0.308	0.94	0.3482
Nursery attendance	0.62	0.79	0.165	1.75	0.0840*
Repeated class	0.44	0.23	-0.210	2.17	0.0330**
Male	0.60	0.47	-0.132	1.26	0.2092
Primary Two (Control=41; Treatment=56; Total=97)					
Age	10.59	9.89	-0.693	2.27	0.0255**
Nursery attendance	0.54	0.54	-0.001	0.01	0.9933
Repeated class	0.49	0.38	-0.113	1.11	0.2713
Male	0.29	0.50	0.207	2.07	0.0409**

	(1)	(2)	(3)	(4)	(5)
Variable(s)	Mean Control	Mean Treatment	Difference	t-value	P-value
Primary Three (Control=54; Treatment=54; Total=108)					
Age	11.93	11.39	-0.537	1.53	0.1290
Nursery attendance	0.46	0.48	0.019	0.19	0.8489
Repeated class	0.44	0.44	0.000	0.00	1.0000
Male	0.52	0.59	0.074	0.77	0.4433
All Classes (Control=140; Treatment=157; Total=297)					
Age	10.67	10.12	-0.550	2.43	0.0155**
Nursery attendance	0.54	0.59	0.057	0.98	0.3272
Repeated class	0.46	0.36	-0.100	1.76	0.0787*
Male	0.48	0.52	0.044	0.75	0.4536
*** p<0.01, ** p<0.05, * p<0.1					

- **Sex**: A half (50%) of the 297 learners assessed were male (48% in control; 52% in treatment). The proportion of male learners assessed in P1, P2 and P3 are: 53%, 41% and 56%, respectively.

- **Age**: The mean age of the learners assessed at was 10.4 years (10.7 for control; 10.1 for treatment), ranging from 6 to 19 years. The mean ages of the learners assessed in P1, P2 and P3 are: 9.1, 10.2, and 11.7 years, respectively.

- **Nursery Attendance**: Most (57%) of the 297 learners assessed attended nursery (54% in control; 59% in treatment). The proportion of learners who attended nursery, disaggregated by class – P1, P2 and P3 are: 71%, 54% and 47%, respectively.

- **Class Repetition**: More than a third (40%) of the 297 learners assessed have ever repeated a class (46% in control; 36% in treatment). The proportion of learners who have ever repeated a class, disaggregated by class – P1, P2 and P3 are: 34%, 42% and 44%, respectively.

In order to improve on effect size estimations, baseline values of these variables are included in the final analyses as covariates. Also, the results without the covariates are shown.

3.3. Subtask 1. Letter Sound Knowledge

3.3.1. Commentary on Letter Sound Knowledge

This subtask assesses the ability of a child to fluently identify the sounds each letter makes. A learner was presented with a sheet of paper with capital and lower-case letters of the English alphabet whereby he/she was asked to tell the sound associated with as many of the letters as they could identify within 60 seconds. The letters that were presented to the learners are as in the extract below.

I	K	N	A	G	S	E	I	E	F
O	T	E	R	o	E	C	A	M	N
T	D	S	H	A	D	W	L	B	I

D	h	N	I	E	T	F	A	N	E
S	D	P	O	U	N	A	E	T	a
T	E	V	T	E	R	H	U	X	L
A	T	C	L	S	Y	G	J	E	T
H	L	O	M	I	R	E	I	S	A
i	Y	H	C	Z	N	P	U	R	O
H	T	O	N	W	R	S	O	E	Q

Extract 1: Letters for the letter sound knowledge subtask

3.3.2. Measure of the Skill of Telling Letter Sounds

The achievement of learners in telling letter sounds is presented as (i) the percentage of learners who sounded correctly one or more letters in 60 seconds and (ii) the mean number of letters sounded correctly per minute by the learners.

Note that a learner was discontinued if he/she gave no single correct answer on the first ten letters.

3.3.3. Effectiveness of the EGR Intervention on Letter Sound Knowledge

Study Hypothesis: We anticipated that more learners in the treatment group will sound correctly one or more letters in 60 seconds compared to those in the control group, after the EGR intervention.

(i) The percentage of learners who sounded correctly one or more letters in 60 seconds

At endline, the percentage of learners who sounded correctly at least 1 letter per minute in the treatment school (87%) was higher than that of the control school (26%).

	Control (Baseline)	Treatment (Baseline)	Control (Endline)	Treatment (Endline)
Primary one	16%	0%	16%	83%
Primary two	15%	14%	20%	86%
Primary three	20%	17%	41%	93%
All classes	17%	11%	26%	87%
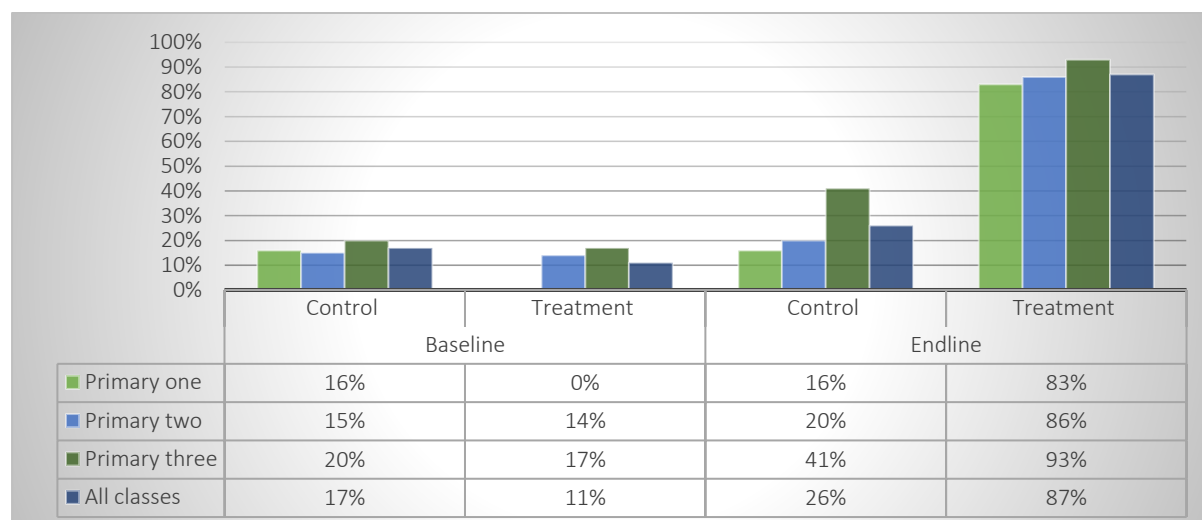
Figure 1: Letter sound knowledge – percentage of learners who sounded correctly at least 1 letter per minute

The impact of the project on the percentage of learners who sounded correctly one or more letters in 60 seconds on the letter sound knowledge subtask is investigated using the intent-to-treat (ITT) estimate. The ITT estimates are calculated through a difference in differences

approach of the dependent variable (letters sounded correctly) on: period (baseline vs endline) and a vector of control variables (age, ever repeated a class, nursery attendance and sex).

Table 5: Letter sound knowledge - proportion of learners who sounded correctly at least 1 letter per minute

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	0.83	0.67	0.56	0.67
Std. Errors	0.095	0.107	0.107	0.066
P-value	0.000***	0.000***	0.000***	0.000***
Panel 2: ITT Estimates with covariates				
DID Estimates	0.84	0.66	0.56	0.67
Std. Errors	0.096	0.106	0.106	0.068
P-value	0.000***	0.000***	0.000***	0.000***
Observations	184	194	216	594
R-square	0.51	0.44	0.38	0.42

Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates".
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Primary One: The results in column (1) show that, the project improved the letter sound knowledge in the treatment than control school, in terms of percentage of P1 learners who sounded correctly at least 1 letter per minute (DIDE=83%). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the letter sound knowledge in the treatment than control school, in terms of percentage of P2 learners who sounded correctly at least 1 letter per minute (DIDE=67%). This effect is statistically significant at 5% level.

Primary Three: The results in column (3) show that, the project improved the letter sound knowledge in the treatment than control school, in terms of percentage of P3 learners who sounded correctly at least 1 letter per minute (DIDE=56%). This effect is statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the letter sound knowledge in the treatment than control school, in terms of percentage of learners who sounded correctly at least 1 letter per minute (DIDE=67%). This effect is statistically significant at 5% level.

During the debrief with the teachers of the treatment school, it was noted that most learners have improved in the letter sound knowledge especially the Primary 1 learners. For the Primary

2 and Primary 3, the teachers felt that they gave little time to this aspect as they had a lot to cover with the learners. However, they also acknowledged the positive trend in the performance of the learners in this subtask. These views indeed explain why there is a larger project effect in Primary 1 as compared to other classes.

Furthermore, teachers of the treatment school have noted a positive change in the way they prepare and deliver their lessons. Before the EGR intervention, most of them acknowledged to enter class without teaching and learning materials but after going through the EGR trainings, they now deliver interesting lessons and the learners appreciate. They use Weekly Foundation Stories to help their learners learn the various letter sounds, pronunciations and understand new words, among others.

⇨ Impact of EGR Intervention by sex

Overall, the results show no differential impact of the EGR intervention on the sex of the learners, in terms of percentage of learners who sounded correctly at least 1 letter per minute ($female * treatment\ coef = -0.03$; $P - value = 0.640$), at 5% level of significance

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	0.72***	0.74***	0.48***	0.62***
Female	-0.09	0.06	-0.04	-0.04
Treatment X Female	-0.06	-0.13	0.09	-0.03
Constant	0.19	0.15	0.43	0.28
Observations	92	97	108	297
R-squared	0.47	0.44	0.30	0.38
Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** p<0.01, ** p<0.05, * p<0.1				

(ii) The mean number of letters sounded correctly per minute by the learners

At endline, the mean number of correct letter sounds per minute by the treatment school learners (mean=10.9 letters) was higher than that of the control school (mean=1.5 letters).
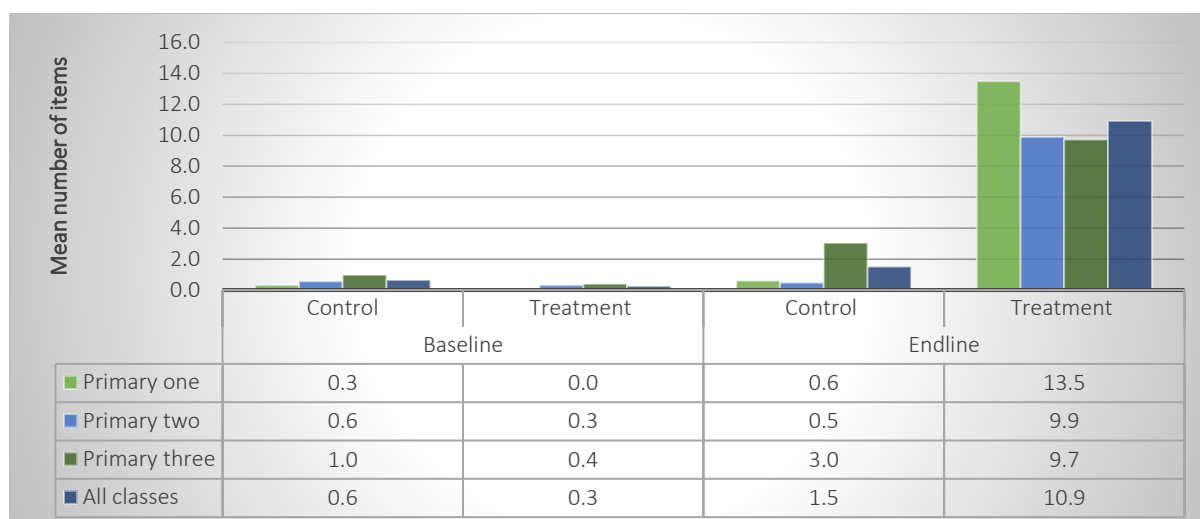
Figure 2: Letter sound knowledge - mean number of sounds stated per minute

The impact of the project on the mean number of letters sounded correctly per minute by the learners on the letter sound knowledge subtask is as shown below:

Table 6: Letter sound knowledge - mean number of sounds stated per minute

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	13.2	9.6	7.3	9.8
Std. Errors	1.863	1.223	1.459	1.309
P-value	0.000***	0.000***	0.000***	0.001***
Panel 2: ITT Estimates with covariates				
DID Estimates	13.2	9.6	7.3	9.8
Std. Errors	1.824	1.226	1.470	1.317
P-value	0.000***	0.000***	0.000***	0.001***
Observations	184	194	216	594
R-square	0.45	0.44	0.33	0.39
Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates". *** p<0.01, ** p<0.05, * p<0.1				

Primary One: The results in column (1) show that, the project improved the letter sound knowledge in the treatment than control school, in terms of the mean number of letters sounded correctly per minute by the P1 learners (DIDE=13.2 letter sounds). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the letter sound knowledge in the treatment than control school, in terms of the mean number of letters

sounded correctly per minute by the P2 learners (DIDE=9.9 letter sounds). This effect is statistically significant at 5% level.

Primary Three: The results in column (3) show that, the project improved the letter sound knowledge in the treatment than control school, in terms of the mean number of letters sounded correctly per minute by the P3 learners (DIDE=9.7 letter sounds). This effect is statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the letter sound knowledge in the treatment than control school, in terms of the mean number of letters sounded correctly per minute by the learners (DIDE=10.9 letter sounds). This effect is statistically significant at 5% level.

⇨ **Impact of EGR Intervention by sex**

Overall, the results show differential impact of the EGR intervention on the sex of the learners (in favour of males), in terms of the mean number of letters sounded correctly per minute by the learners ($female * treatment\ coef = -3.87$; $P - value = 0.019$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	15.95***	12.18***	7.42***	11.26***
Female	0.33	0.04	-0.29	-0.17
Treatment X Female	-5.74	-5.54**	-1.85	-3.87**
Constant	0.46	0.46	3.18	1.60
Observations	92	97	108	297
R-squared	0.37	0.40	0.18	0.29
Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** p<0.01, ** p<0.05, * p<0.1				

3.4.   Subtask 2. Segmenting

3.4.1.   Commentary on Segmenting

The Assessors read out aloud the *words* listed and the learners were asked to tell all the sounds in the word read. This was not a timed task. Each sound in a particular word was marked. The words that were presented to the learners are as in the extract below.

If		/i/	/f/
Too		/t/	/oo/
Up		/u/	/p/
Me		/m/	/ee/

Say	/s/	/ae/	
Dog	/d/	/o/	/g/
Map	/m/	/a/	/p/
Bet	/b/	/e/	/t/
Fish	/f/	/i/	/sh/
Lick	/l/	/i/	/ck/

Extract 2: Words for the segmenting subtask

3.4.2. Measure of the Skill of Segmenting

The achievement of learners in segmenting is presented as (i) the percentage of learners who could correctly tell all the sounds of one or more words in the list and (ii) the mean number of words whose sounds were correctly stated by the learners.

Note that a learner was discontinued if he/she gave no single correct answer, either partly or fully, on the first five words.

3.4.3. Effectiveness of the EGR Intervention on Segmenting Knowledge

Study Hypothesis: We anticipated that more learners in the treatment group will tell correctly the sounds of one or more words in the list compared to those in the control group, after the EGR intervention.

(i) The percentage of learners who could correctly tell all the sounds of one or more words

At endline, the percentage of learners who could correctly tell all the sounds of one or more words in the treatment school (74%) was higher than that of the control school (13%).

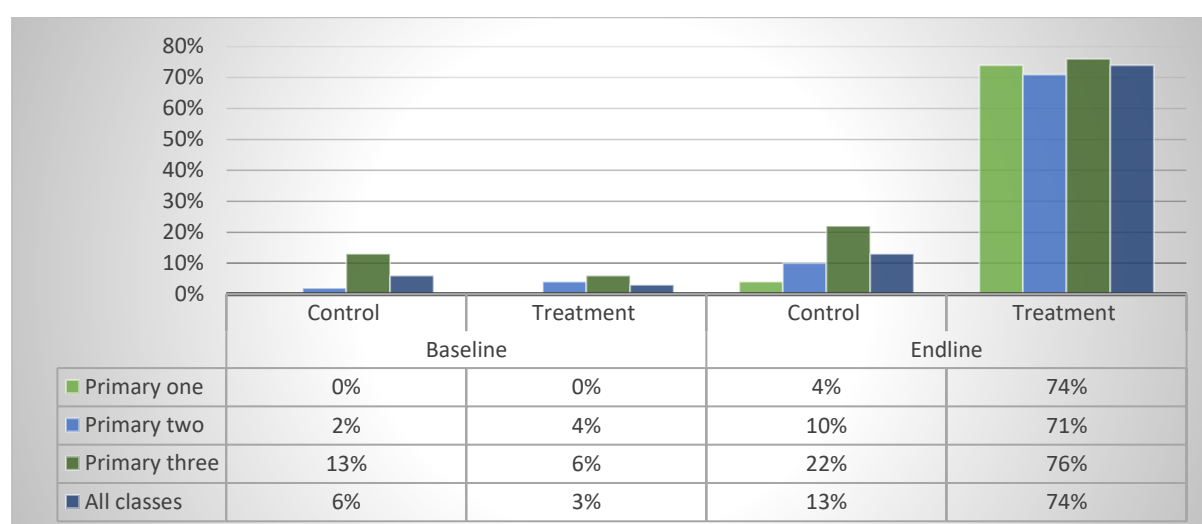|	Baseline		Endline	
	Control	Treatment	Control	Treatment
■ Primary one	0%	0%	4%	74%
■ Primary two	2%	4%	10%	71%
■ Primary three	13%	6%	22%	76%
■ All classes	6%	3%	13%	74%

Figure 3: Segmenting – percentage of learners who segmented correctly at least 1 word

The impact of the project on the percentage of learners who could correctly tell all the sounds of one or more words in the segmenting subtask is investigated using the intent-to-treat estimate. The ITT estimates are calculated through a difference in differences approach of the

dependent variable (words correctly segmented) on: period (baseline vs endline) and a vector of control variables (age, ever repeated a class, nursery attendance and sex).

Table 7: Segmenting - proportion of learners who segmented correctly at least 1 word

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	0.70	0.61	0.61	0.64
Std. Errors	0.071	0.084	0.099	0.021
P-value	0.000***	0.000***	0.000***	0.000***
Panel 2: ITT Estimates with covariates				
DID Estimates	0.72	0.60	0.61	0.64
Std. Errors	0.068	0.085	0.098	0.026
P-value	0.000***	0.000***	0.000***	0.000***
Observations	184	194	216	594
R-square	0.63	0.50	0.37	0.47
Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates". *** p<0.01, ** p<0.05, * p<0.1				

Primary One: The results in column (1) show that, the project improved the segmenting knowledge in the treatment than control school, in terms of percentage of P1 learners who could correctly tell all the sounds of one or more words (DIDE=70%). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the segmenting knowledge in the treatment than control school, in terms of percentage of P2 learners who could correctly tell all the sounds of one or more words (DIDE=61%). This effect is statistically significant at 5% level.

Primary Three: The results in column (3) show that, the project improved the segmenting knowledge in the treatment than control school, in terms of percentage of P3 learners who could correctly tell all the sounds of one or more words (DIDE=61%). This effect is statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the segmenting knowledge in the treatment than control school, in terms of percentage of learners who could correctly tell all the sounds of one or more words (DIDE=64%).

This improvement was also noted by the teachers of the treatment schools during a debrief with them. They however added that, they still needed more time to explain the concepts of diagraphs and trigraphs to the learners to help them effectiveness improve on their reading

skills. This opinion as well explains why most learners in the treatment school as well as control could still not manage to read the 'oral passage reading' that was presented to them.

⇨ Impact of EGR Intervention by sex

Overall, the results show no differential impact of the EGR intervention by sex of the learners, in terms of percentage of learners who could correctly tell all the sounds of one or more words ($female * treatment\ coef = 0.05$; $P - value = 0.248$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	0.70***	0.64***	0.47***	0.58***
Female	-0.08	0.03	-0.06	-0.04
Treatment X Female	0.02	-0.03	0.16	0.05
Constant	0.08	0.08	0.25	0.15
Observations	92	97	108	297
R-squared	0.51	0.37	0.29	0.38
Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** p<0.01, ** p<0.05, * p<0.1				

(ii) The mean number of words whose sounds were correctly stated by the learners

At endline, the mean number of words correctly segmented by the treatment school learners (mean=2.3 words) was higher than that of the control school (mean=0.4 words).

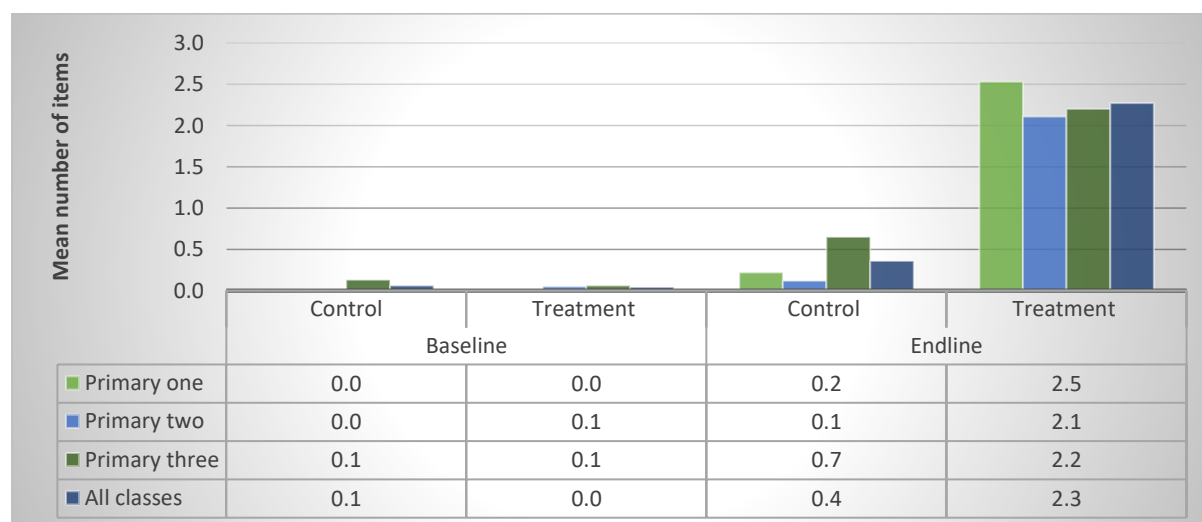|	Baseline Control	Baseline Treatment	Endline Control	Endline Treatment
Primary one	0.0	0.0	0.2	2.5
Primary two	0.0	0.1	0.1	2.1
Primary three	0.1	0.1	0.7	2.2
All classes	0.1	0.0	0.4	2.3

Figure 4: Segmenting- mean number of words correctly segmented

The impact of the project on the mean number of words whose sounds were correctly stated by the learners on the segmenting subtask is as shown below:

Table 8: Segmenting- mean number of words correctly segmented

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	2.31	1.96	1.63	1.93
Std. Errors	0.380	0.260	0.341	0.171
P-value	0.000***	0.000***	0.000***	0.000***
Panel 2: ITT Estimates with covariates				
DID Estimates	2.37	1.96	1.62	1.95
Std. Errors	0.376	0.265	0.337	0.187
P-value	0.000***	0.000***	0.000***	0.000***
Observations	184	194	216	594
R-square	0.41	0.45	0.33	0.38
Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates". *** p<0.01, ** p<0.05, * p<0.1				

Primary One: The results in column (1) show that, the project improved the segmenting knowledge in the treatment than control school, in terms of the mean number of words correctly segmented by the P1 learners (DIDE=2.3 words). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the segmenting knowledge in the treatment than control school, in terms of the mean number of words correctly segmented by the P2 learners (DIDE=2.0 words). This effect is statistically significant at 5% level.

Primary Three: The results in column (3) show that, the project improved the segmenting knowledge in the treatment than control school, in terms of the mean number of words correctly segmented by the P3 learners (DIDE=1.6 words). This effect is statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the segmenting knowledge in the treatment than control school, in terms of the mean number of words correctly segmented by the learners (DIDE=1.9 words). This effect is statistically significant at 5% level.

⇨ Impact of EGR Intervention by sex

Overall, the results show no differential impact of the EGR intervention by sex of the learners, in terms of the mean number of words correctly segmented by the learners ($female * treatment\ coef = 0.34\ ; P - value = 0.291$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	1.98***	2.24***	1.24**	1.73***
Female	-0.38	-0.05	-0.43	-0.35***
Treatment X Female	0.70	-0.52	0.70	0.34
Constant	0.38	0.15	0.86	0.54
Observations	92	97	108	297
R-squared	0.29	0.34	0.18	0.25

Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** $p<0.01$, ** $p<0.05$, * $p<0.1$

3.5.   Subtask 3. Nonword Decoding

3.5.1.   Commentary on Nonword Decoding

A learner was presented with a sheet of made-up English words where he/she was asked to read the words aloud, quickly and carefully in 60 seconds. The words that were presented to the learners are as in the extract below.

Maz	Ver	Lut	Paf	Nom
Yod	Fut	Et	Zib	Mib
Dag	Fol	San	Leb	Huz
Teb	Ved	Bif	Lef	Vom
Ret	Nep	Riz	Lus	Rop
Dit	Nup	Kad	Hig	Yag
Wix	Tob	Tib	Gax	Jod
Nad	Gof	Sig	Ral	Reg
Tup	Fim	Peb	Sen	Kib
Sim	Fid	Sal	Zon	Tat

Extract 3: Words for the nonword decoding subtask

3.5.2.   Measure of the Skill of Nonword Decoding

The achievement of learners in nonword decoding is presented as (i) the percentage of learners who could read correctly one or more words per minute and (ii) the mean number of words read correctly per minute by the learners.

Note that a learner was discontinued if he/she gave no single correct answer on the first five words.

3.5.3.   Effectiveness of the EGR Intervention on Nonword Decoding Knowledge

Study Hypothesis: We anticipated that more learners in the treatment group will read correctly one or more words per minute compared to those in the control group, after the EGR intervention.

(i) The percentage of learners who could read correctly one or more words per minute

At endline, the percentage of learners who read correctly at least 1 word in 60 seconds in the treatment school (17%) was higher than that of the control school (6%).

	Control	Treatment	Control	Treatment
	Baseline		Endline	
■ Primary one	0%	0%	0%	21%
■ Primary two	0%	2%	2%	14%
■ Primary three	9%	4%	13%	15%
■ All classes	4%	2%	6%	17%
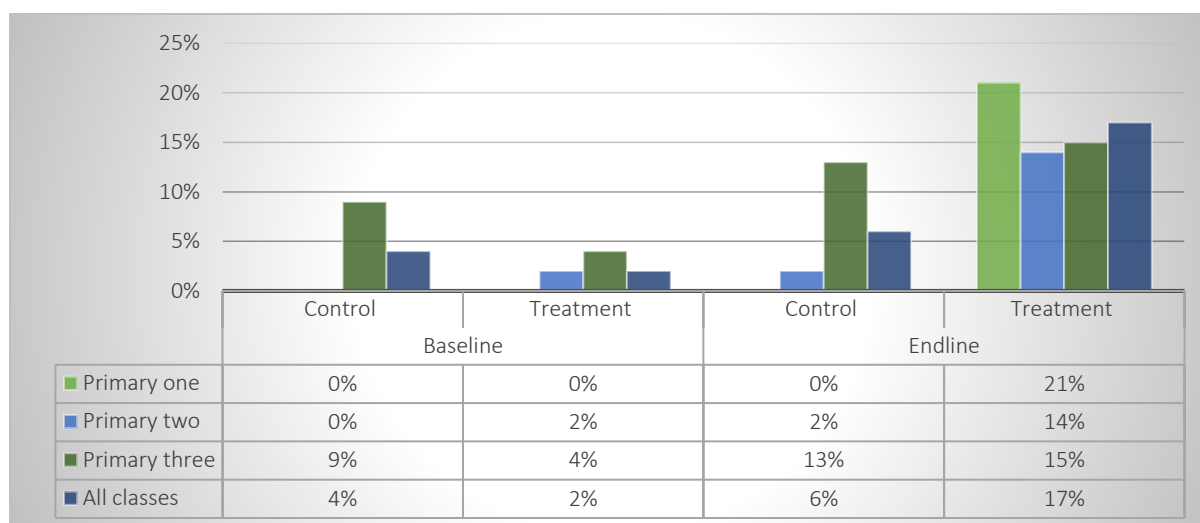
Figure 5: Nonword decoding – percentage of learners who read correctly at least 1 word per minute

The impact of the project on the percentage of learners who could read correctly one or more words per minute in the nonword decoding subtask is investigated using the intent-to-treat estimate. The ITT estimates are calculated through a difference in differences approach of the dependent variable (words correctly read) on: period (baseline vs endline) and a vector of control variables (age, ever repeated a class, nursery attendance and sex).

Table 9: Nonword decoding - proportion of learners who read correctly at least 1 word per minute

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	0.21	0.10	0.07	0.13
Std. Errors	0.060	0.056	0.082	0.029
P-value	0.001***	0.074*	0.369	0.007***
Panel 2: ITT Estimates with covariates				
DID Estimates	0.22	0.10	0.07	0.13
Std. Errors	0.061	0.056	0.082	0.030
P-value	0.000***	0.077*	0.373	0.008***
Observations	184	194	216	594
R-square	0.17	0.07	0.02	0.05
Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates". *** p<0.01, ** p<0.05, * p<0.1				

Primary One: The results in column (1) show that, the project improved the nonword decoding knowledge in the treatment than control school, in terms of the percentage of P1 learners who read correctly at least 1 word per minute (DIDE=21%). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the nonword decoding knowledge in the treatment than control school, in terms of the percentage of P2 learners who read correctly at least 1 word per minute (DIDE=10%). This effect is statistically significant at 10% level.

Primary Three: The results in column (3) show that, the project improved the nonword decoding knowledge in the treatment than control school, in terms of the percentage of P3 learners who read correctly at least 1 word per minute (DIDE=7%). This effect is however not statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the nonword decoding knowledge in the treatment than control school, in terms of the percentage of learners who read correctly at least 1 word per minute (DIDE=13%). This effect is statistically significant at 5% level.

As of the debrief with the teachers in the treatment school, most of their learners need support in diagraphs and trigraphs. This would facilitate them [learners] put together the various sounds to be able to read.

⇨ Impact of EGR Intervention by sex

Overall, the results show no differential impact of the EGR intervention by sex of the learners, in terms of the percentage of learners who read correctly at least 1 word per minute ($female * treatment\ coef = -0.03 \ ; P-value = 0.598$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	0.27**	0.10	0.02	0.12*
Female	0.00**	-0.08	0.05	0.00
Treatment X Female	-0.11	0.01	0.01	-0.03
Constant	0.00	0.08	0.11	0.06
Observations	92	97	108	297
R-squared	0.13	0.06	0.01	0.03
Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** p<0.01, ** p<0.05, * p<0.1			

(ii) The mean number of words read correctly per minute by the learners

At endline, the mean number of words correctly read per minute by the treatment school learners (mean=1.2 words) was higher than that of the control school (mean=0.6 words).
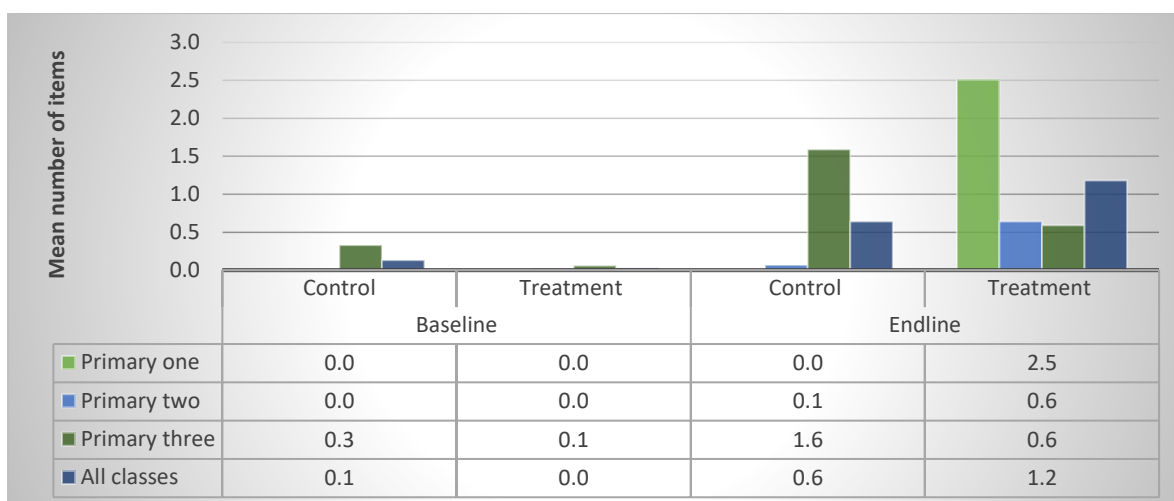
Figure 6: Nonword decoding – mean number of words correctly read per minute

The impact of the project on the mean number of words correctly read per minute by the learners on the nonword decoding subtask is as shown below:

Table 10: Nonword decoding- mean number of words correctly read per minute

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	2.51	0.55	-0.72	0.65
Std. Errors	0.813	0.288	0.758	0.670
P-value	0.002***	0.057*	0.342	0.375
Panel 2: ITT Estimates with covariates				
DID Estimates	2.58	0.54	-0.78	0.68
Std. Errors	0.833	0.275	0.772	0.688
P-value	0.002***	0.049**	0.316	0.369
Observations	184	194	216	594
R-square	0.13	0.06	0.04	0.04
Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates". *** p<0.01, ** p<0.05, * p<0.1				

Primary One: The results in column (1) show that, the project improved the nonword decoding knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the P1 learners (DIDE=2.5 words). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the nonword decoding knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the P2 learners (DIDE=0.6 words). This effect is statistically significant at 10% level.

Primary Three: The results in column (3) show that, the project did not improve the nonword decoding knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the P3 learners (DIDE=-0.7 words). This effect is however not statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the nonword decoding knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the learners (DIDE=0.7 words). This effect is however not statistically significant at 5% level.

⇨ Impact of EGR Intervention by sex

Overall, the results show no differential impact of the EGR intervention by sex of the learners, in terms of the mean number of words correctly read per minute by the learners ($female * treatment\ coef = 0.08 ; P - value = 0.865$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	2.32**	0.80	-0.98	0.51
Female	0.00	-0.23	0.04	-0.13
Treatment X Female	0.36	-0.55	-0.05	0.08
Constant	0.00	0.23	1.57	0.70
Observations	92	97	108	297
R-squared	0.09	0.07	0.02	0.01

Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** p<0.01, ** p<0.05, * p<0.1

3.6. Subtask 4. Oral Passage Reading and Comprehension

3.6.1. Commentary on Oral Passage Reading and Comprehension

A learner was presented with a sheet of paper with a passage in English language whereby he/she was expected to read the passage aloud, quickly and carefully in 60 seconds. Those learners who managed to read any 44 words (75%) correctly were asked some oral questions about the passage. Below is the passage that the learners were asked to read.

> Susan lost her coat.
> It was a cold day. She was worried. She looked in her desk. She looked in her seat. The coat was not there.
> She ran to the flag. She looked under the tree. It was not there.
> She told her teacher she had lost her coat. The teacher pointed to Susan's neck.
> Susan laughed.

Extract 4: Passage for the oral passage reading and comprehension subtask

3.6.2. Measure of the Skill of Oral Passage Reading and Comprehension

The achievement of learners in reading a passage is presented as (i) the percentage of learners who could read correctly one or more words per minute in the passage and (ii) mean number of words correctly read in the passage per minute.

For the comprehension task, the percentage of learners who answered correctly one out of five questions is presented. Note that a learner was discontinued if he/she gave no single correct answer on the first twelve words.

3.6.3. Effectiveness of the EGR Intervention on Oral Passage Reading Knowledge

Study Hypothesis: We anticipated that more learners in the treatment group will read correctly one or more words per minute in the passage compared to those in the control group, after the EGR intervention.

(i) The percentage of learners who read correctly one or more words per minute in the passage

At endline, the percentage of learners who read correctly one or more words in the passage per minute in the treatment school (15%) was higher than that of the control school (14%).

	Baseline Control	Baseline Treatment	Endline Control	Endline Treatment
Primary one	0%	0%	2%	21%
Primary two	0%	0%	15%	16%
Primary three	6%	0%	22%	7%
All classes	2%	0%	14%	15%
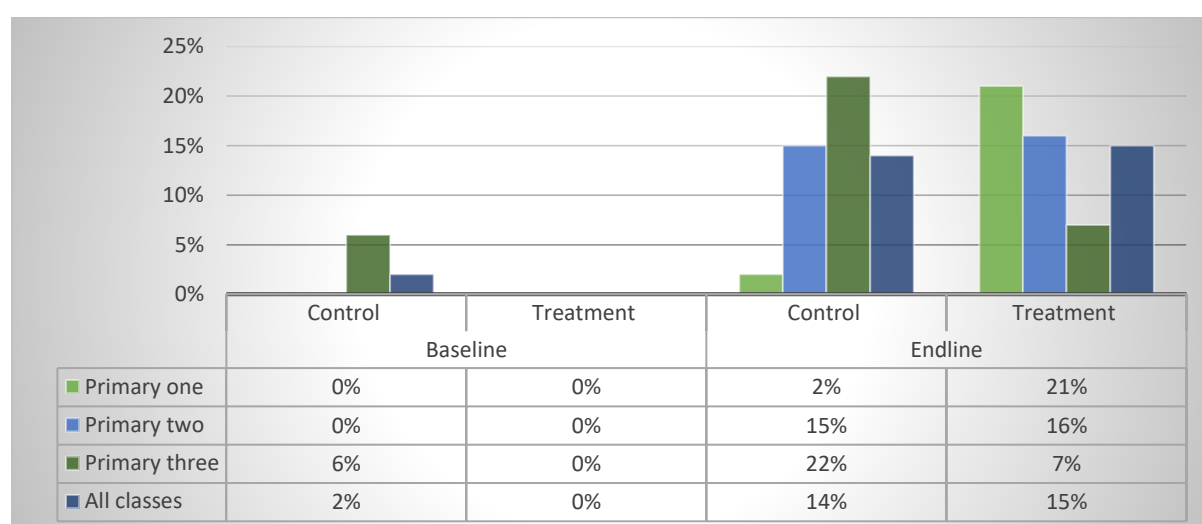
Figure 7: Oral passage reading – percentage of learners who read correctly at least 1 word in the passage per minute

The impact of the project on the percentage of learners who could read correctly one or more words per minute in the oral passage and reading comprehension subtask is investigated using the intent-to-treat estimate. The ITT estimates are calculated through a difference in differences approach of the dependent variable (words correctly read from the passage) on: period (baseline vs endline) and a vector of control variables (age, ever repeated a class, nursery attendance and sex).

Table 11: Oral passage reading - proportion of learners who read correctly at least 1 word in the passage per minute

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	0.19	0.01	-0.09	0.03
Std. Errors	0.064	0.075	0.074	0.054
P-value	0.003***	0.847	0.215	0.576
Panel 2: ITT Estimates with covariates				
DID Estimates	0.19	0.01	-0.10	0.03
Std. Errors	0.065	0.075	0.075	0.053
P-value	0.004***	0.903	0.207	0.582
Observations	184	194	216	594
R-square	0.14	0.08	0.08	0.06
Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates". *** p<0.01, ** p<0.05, * p<0.1				

Primary One: The results in column (1) show that, the project improved oral passage reading knowledge in the treatment than control school, in terms of the percentage of P1 learners who read correctly at least 1 word per minute (DIDE=19%). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the oral passage reading knowledge in the treatment than control school, in terms of the percentage of P2 learners who read correctly at least 1 word per minute (DIDE=1%). This effect is however not statistically significant at 5% level.

Primary Three: The results in column (3) show that, the project did not improve the oral passage reading knowledge in the treatment than control school, in terms of the percentage of P3 learners who read correctly at least 1 word per minute (DIDE=-9%). This effect is however not statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the oral passage reading knowledge in the treatment than control school, in terms of the percentage of learners who read correctly at least 1 word per minute (DIDE=3%). This effect is however not statistically significant at 5% level.

In addition to the limited knowledge in diagraphs and trigraphs that teachers in the treatment school felt has not yet been developed in the learners, cases of absenteeism of learners was also cited as a major factor affecting teaching and learning in the schools. This is evident during the data collection days where some learners did not appear in school over a period of 1 week.

⇨ **Impact of EGR Intervention by sex**

Overall, the results show no differential impact of the EGR intervention by sex of the learners, in terms of the percentage of learners who read correctly at least 1 word per minute ($female * treatment\ coef = -0.11 ; P-value = 0.086$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	0.27**	0.06	-0.12	0.07
Female	0.05	-0.01	0.09	0.06*
Treatment X Female	*-0.17*	*-0.10*	*-0.06*	*-0.11**
Constant	0.00	0.15	0.18	0.10
Observations	92	97	108	297
R-squared	0.10	0.01	0.05	0.01
Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** p<0.01, ** p<0.05, * p<0.1				

(ii) The mean number of words correctly read in the passage per minute

At endline, the mean number of words correctly read in the passage per minute by the treatment school learners (mean=1.9 words) was higher than that of the control school (mean=1.3 words).

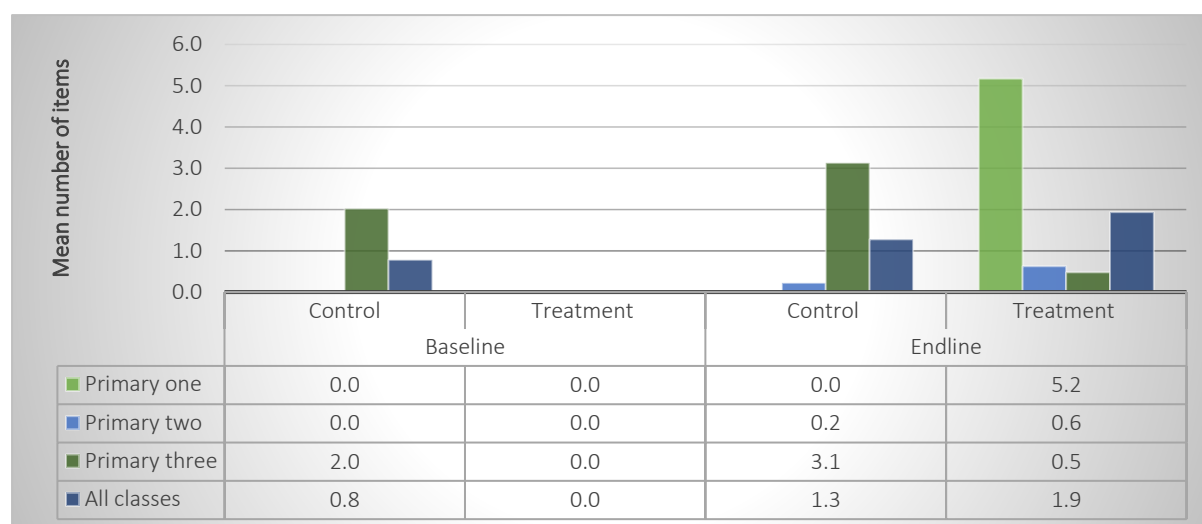|	Control	Treatment	Control	Treatment
	Baseline		Endline	
■ Primary one	0.0	0.0	0.0	5.2
■ Primary two	0.0	0.0	0.2	0.6
■ Primary three	2.0	0.0	3.1	0.5
■ All classes	0.8	0.0	1.3	1.9

Figure 8: Oral passage reading - mean number of words correctly read per minute

The impact of the project on the mean number of words correctly read in the passage per minute by the learners on the oral passage reading subtask is as shown below:

Table 12: Oral passage reading- mean number of words correctly read per minute

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	5.15	0.41	-0.63	1.44
Std. Errors	1.747	0.318	1.788	1.342
P-value	0.004***	0.204	0.725	0.333
Panel 2 ITT Estimates with covariates				
DID Estimates	5.25	0.35	-0.75	1.49
Std. Errors	1.790	0.296	1.777	1.370
P-value	0.004***	0.242	0.674	0.326
Observations	184	194	216	594
R-square	0.12	0.05	0.04	0.02

Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates".

*** p<0.01, ** p<0.05, * p<0.1

Primary One: The results in column (1) show that, the project improved the oral passage reading knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the P1 learners (DIDE=5.2 words). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the oral passage reading knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the P2 learners (DIDE=0.4 words). This effect is however not statistically significant at 5% level.

Primary Three: The results in column (3) show that, the project did not improve the oral passage reading knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the P3 learners (DIDE=-0.6 words). This effect is however not statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the oral passage reading knowledge in the treatment than control school, in terms of the mean number of words correctly read per minute by the learners (DIDE=1.4 words). This effect is statistically significant at 5% level.

⇨ **Impact of EGR Intervention by sex**

Overall, the results show no differential impact of the EGR intervention by sex of the learners, in terms of the mean number of words correctly read per minute by the learners ($female * treatment\ coef = -0.19\ ; P - value = 0.850$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	4.36**	0.99*	-1.79	0.76
Female	0.05	0.10	1.23	0.36
Treatment X Female	1.46	-1.13*	-1.89	-0.19
Constant	0.00	0.15	2.54	1.09
Observations	92	97	108	297
R-squared	0.09	0.06	0.05	0.00

Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** p<0.01, ** p<0.05, * p<0.1

3.6.4. Reading Comprehension

At endline, out of 42(14.1%) learners who managed to read one or more words in the passage in 60 seconds, only 1(0.3%) qualified to answer the comprehension questions. To be eligible to answer the comprehension questions, the learner must have read at least 44(75%) words in the passage in 60 seconds. This learner managed to answer correctly three out of the five questions.

Overall, only 0.3% of the learners assessed managed to answer at least one out of five questions in the oral passage reading and comprehension subtask.

3.7. Subtask 5: English Vocabulary

3.7.1. Commentary on English Vocabulary

Vocabulary was assessed using two tasks on knowledge of: (a) body parts and (b) words from the environment, in order to check learners' understanding of contextual terms in English. The Assessors read out the *words* listed and the learners were asked to show or touch parts of his/her body and objects in the environment, that matched the *word(s)* the Assessors had read out. This was not a timed task. Note that, the instructions were read in either the local language or English but the vocabulary words were stated in English. The words that were presented to the learners are as in the extract below.

Foot	Arm	Chin	Knee	Mouth	Back	Elbow	Shoulder
Pencil		Shoes	Book	Rubber	Paper	ground (floor)	

Extract 5: Words for the English vocabulary subtask

3.7.2. Measure of the Skill of English Vocabulary

The achievement of learners in English vocabulary is presented as (i) the percentage of learners who could identify at least 50% of the English vocabulary words and (ii) the mean number of body parts and objects that were correctly identified.

3.7.3. Effectiveness of the EGR Intervention on English Vocabulary Knowledge

Study Hypothesis: We anticipated that more learners in the treatment group will identify at least 50% of the English vocabulary words compared to those in the control group, after the EGR intervention.

(i) The percentage of learners who could identify at least 50% of the English vocabulary words

At endline, the percentage of learners who identified at least 50% of the English vocabulary words (body parts and objects from the environment) in the treatment school (87%) was higher than that of the control school (64%).

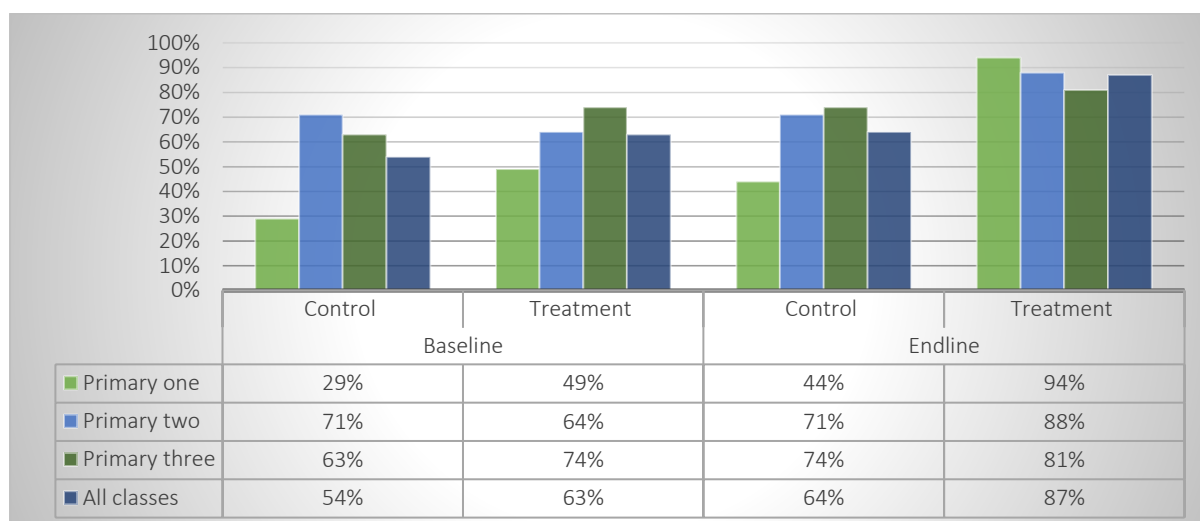|	Baseline Control	Baseline Treatment	Endline Control	Endline Treatment
Primary one	29%	49%	44%	94%
Primary two	71%	64%	71%	88%
Primary three	63%	74%	74%	81%
All classes	54%	63%	64%	87%

Figure 9: English vocabulary- percentage of learners who could identify at least 50% of the English vocabulary words

The impact of the project on the percentage of learners who could identify correctly at least 50% of the body parts and objects from the environment in the English Vocabulary subtask is investigated using the intent-to-treat estimate. The ITT estimates are calculated through a difference in differences approach of the dependent variable (English vocabulary words) on: period (baseline vs endline) and a vector of control variables (age, ever repeated a class, nursery attendance and sex).

Table 13: English vocabulary- proportion of learners who could identify at least 50% of the English vocabulary words (body parts and objects from the environment)

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	0.29	0.23	-0.04	0.15
Std. Errors	0.130	0.128	0.120	0.100
P-value	0.027**	0.072*	0.759	0.196
Panel 2: ITT Estimates with covariates				
DID Estimates	0.30	0.23	-0.04	0.15
Std. Errors	0.128	0.128	0.119	0.098
P-value	0.021**	0.073*	0.757	0.179
Observations	184	194	216	594
R-square	0.24	0.04	0.02	0.07

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes

Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates".
*** p<0.01, ** p<0.05, * p<0.1

Primary One: The results in column (1) show that, the project improved the English vocabulary knowledge in the treatment than control school, in terms of the percentage of P1 learners who could identify at least 50% of the English vocabulary words (DIDE=29%). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the English vocabulary knowledge in the treatment than control school, in terms of the percentage of P2 learners who could identify at least 50% of the English vocabulary words (DIDE=23%). This effect is statistically significant at 10% level.

Primary Three: The results in column (3) show that, the project did not improve the English vocabulary knowledge in the treatment than control school, in terms of the percentage of P3 learners who could identify at least 50% of the English vocabulary words (DIDE=-4%). This effect is however not statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the English vocabulary knowledge in the treatment than control school, in terms of the percentage of learners who could identify at least 50% of the English vocabulary words (DIDE=15%). This effect is statistically significant at 5% level.

⇨ Impact of EGR Intervention by sex

Overall, the results show no differential impact of the EGR intervention by sex of the learners, in terms of the percentage of learners who could identify at least 50% of the English vocabulary words ($female * treatment\ coeff = 0.05 ; P-value = 0.724$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	0.56***	-0.06	0.06	0.22
Female	0.23	-0.20	-0.02	0.05
Treatment X Female	-0.18	0.38**	0.02	0.05
Constant	0.35	0.85	0.75	0.61
Observations	92	97	108	297
R-squared	0.32	0.10	0.01	0.08

(ii) The mean number of body parts and objects from the environment that were correctly identified

At endline, the mean number of body parts and objects from the environment that were correctly identified by the treatment school learners (mean=9.1 items) was higher than that of the control school (mean=7.6 items).

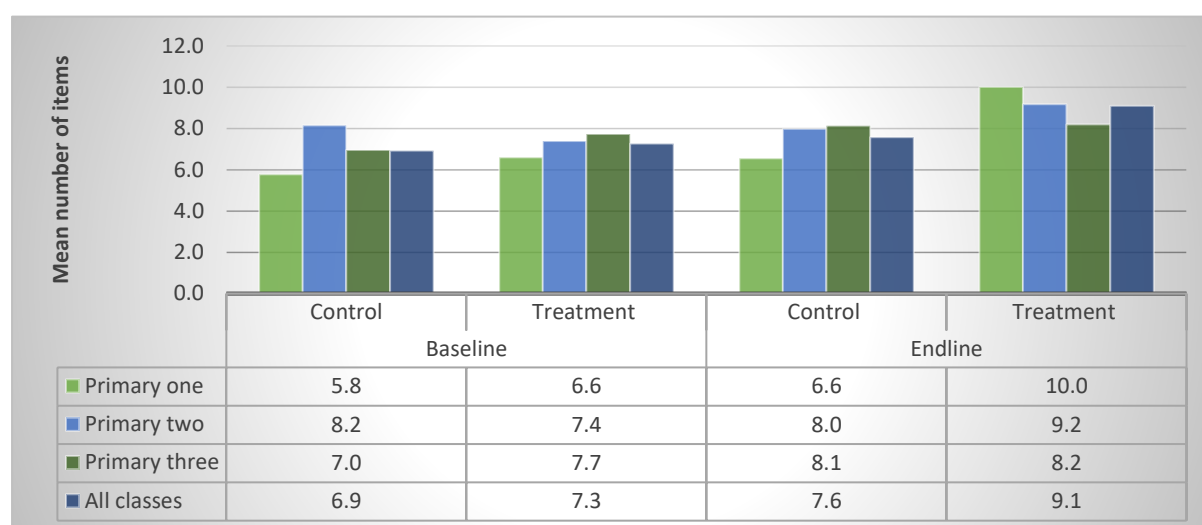|	Control (Baseline)	Treatment (Baseline)	Control (Endline)	Treatment (Endline)
Primary one	5.8	6.6	6.6	10.0
Primary two	8.2	7.4	8.0	9.2
Primary three	7.0	7.7	8.1	8.2
All classes	6.9	7.3	7.6	9.1

Figure 10: English vocabulary - mean number of body parts and objects from the environment that were correctly identified

The impact of the project on the mean number of body parts and objects from the environment that were correctly identified by the learners on English vocabulary subtask is as shown below:

Table 14: English vocabulary - mean number of body parts and objects from the environment that were correctly identified

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
Panel 1: ITT Estimates without covariates				
DID Estimates	2.65	1.96	-0.70	1.17
Std. Errors	0.684	0.806	0.545	0.813
P-value	0.000***	0.016**	0.198	0.209
Panel 2: ITT Estimates with covariates				
DID Estimates	2.75	1.96	-0.68	1.20
Std. Errors	0.669	0.789	0.512	0.821
P-value	0.000***	0.014**	0.184	0.203
Observations	184	194	216	594
R-square	0.34	0.07	0.06	0.11

Notes: Columns (1) to (4) report the intent-to-treat estimates of the project impact. Standard errors are robust, and clustered at school-class level for estimates in Column (4). The ITT is based on the

	(1)	(2)	(3)	(4)
	Primary one	Primary two	Primary three	All classes
baseline and endline samples for only learners who were assessed at these 2 points. The baseline covariates in Panel 2 include: age of the learner, sex of the learner, whether learner ever attended nursery and whether learner ever repeated any class. The R-squared estimates are for "ITT without covariates". *** p<0.01, ** p<0.05, * p<0.1				

Primary One: The results in column (1) show that, the project improved the English vocabulary knowledge in the treatment than control school, in terms of the mean number of English vocabulary words that were correctly identified by the P1 learners (DIDE=2.7 body parts and objects). This effect is statistically significant at 5% level.

Primary Two: The results in column (2) show that, the project improved the English vocabulary knowledge in the treatment than control school, in terms of the mean number of English vocabulary words that were correctly identified by the P2 learners (DIDE=2.0 body parts and objects). This effect is statistically significant at 5% level.

Primary Three: The results in column (3) show that, the project did not improve the English vocabulary knowledge in the treatment than control school, in terms of the mean number of English vocabulary words that were correctly identified by the P3 learners (DIDE=-0.7 body parts and objects). This effect is however not statistically significant at 5% level.

All classes: The results in column (4) show that, the project improved the English vocabulary knowledge in the treatment than control school, in terms of the mean number of English vocabulary words that were correctly identified by the learners (DIDE=1.2 body parts and objects). This effect is however not statistically significant at 5% level.

⇨ **Impact of EGR Intervention by sex**

Overall, the results show no differential impact of the EGR intervention by sex of the learners, in terms of the mean number of English vocabulary words that were correctly identified by the learners ($female * treatment\ coeff = -0.03;\ P-value = 0.966$), at 5% level of significance.

	(1)	(2)	(3)	(4)
	Primary 1	Primary 2	Primary 3	All classes
Treatment	3.76	0.05	0.47	1.54
Female	0.86	-1.16	0.27	0.25
Treatment X Female	-0.73	1.88	-0.92	-0.03
Constant	6.19	8.77	8.00	7.45
Observations	92	97	108	297
R-squared	0.40	0.09	0.02	0.10
Notes: The estimates are based on endline samples of learners who were only assessed at both baseline and endline phases, without covariates. The standard errors are robust, and clustered at school-class level for estimates in Column (4). *** p<0.01, ** p<0.05, * p<0.1				

3.8. Other findings from the debrief with the teachers

In order to understand in-depths the changes in the treatment school as a result of the EGR intervention, teachers of World View Nursery and Primary School were engaged into a debriefing session. Their views and perceptions in regards to EGR intervention are as below:

According to most of the teachers in the treatment school, what they have learnt has become part of their work. Teachers stated to have learnt how to teach letter sounds and learners are using them. According to one of the teachers: "I did not know letter sounds and how to teach it but after receiving the training, I now know about 80% of the letter sounds and I am able to transfer this knowledge to my learners" (Teacher, treatment school). During the course of the intervention, the LGIHE team observed that the trained teachers were applying the teaching of sounds in their lessons.

For most of the teachers in the treatment school, their skills of writing on manila papers and chalk boards have been improved. During the course of the intervention, the LGIHE team observed that all the teachers trained taught using displayed teaching and learning materials that were neat. Additionally, teachers were observed to exhibit legible and clear hand writings on the chalk board. This has influenced on how their learners write notes. Moreover, using charts in the teaching and learning process makes learning incidental.

Before the EGR intervention, teachers of the treatment schools acknowledged to give too much content to the leaners however, after receiving that intervention, they have learnt how to breakdown the content into consumable items. This has enabled their learners to effectively learn and has also reduced on stuffing.

Related to this, teachers in the treatment school, before the EGR intervention acknowledged to scheme following the text books however, after receiving the intervention, the scheme based on the curriculum. They also try to integrate the skills acquired and knowledge learnt in other subjects. This all happens during planning and delivery of the lessons. During the course of the intervention, the LGIHE team observed that teachers had up-to-date lesson plans that were matching with the respective schemes, and with relevant learning materials.

In regards to the Community of Practice that was formed as a result of the EGR intervention, teachers in the treatment school perceived it as being useful as it has helped them: to find out the appropriate solutions to their difficulties; to share knowledge; and to reflect on what they need to do in order to improve on the quality of teaching and learning of EGR in their school. For continuity of the COP, teachers proposed to have fortnight meetings however, they still felt need for additional support from LGIHE and AVSI that could be provided when need arises. One of the teachers had this to say: "I think the COP should continue but additional support from LGIHE and AVSI is needed" (Teacher, treatment school).

For most of the aspects in EGR intervention, teachers as well as learners need more time to fully comprehend and apply. They envisaged like 1 year to be able to fully transfer the acquired knowledge to the learners. During the course of the intervention, the LGIHE team observed

that teachers also need support in record keeping, storage of teaching and learning materials, assessment and evaluation, planning activities of the lesson in consideration of its duration.

In summary, teachers of the treatment school have observed the following most significant changes as a result of the EGR intervention:

- The EGR intervention has promoted teamwork among teachers. Some of the teachers cited examples of where they invited their colleagues to help with a particular topic – sound knowledge. This to them had never happened before the EGR intervention.
- Teachers look at themselves as resource persons on EGR and EGRA.
- Teachers now use child centred approaches coupled with interesting lessons that motivate the learners to learn.
- Teachers have acknowledged the fact that they now write well on the manila papers and chalk boards.
- Teachers' confidence on EGR has improved as they view themselves with a potential of helping their learners to improve on their reading and writing skills.

CHAPTER FOUR: CONCLUSIONS AND RECOMMENDATIONS

4.1. Conclusions

The results realized through difference in differences analyses reveal improvements in the treatment school compared to the control over the period of 4 months of EGR intervention. A debrief with teachers reveals these improvements to be attributed to the EGR intervention that was implemented in the treatment school, immediately after the baseline assessment.

The results therefore support the stated hypotheses with statistically significant improvements (at 5% level) being observed in the subtask areas of: letter sound knowledge, segmenting knowledge, and nonword decoding. Smaller project effects were realized in the subtask areas of: oral passage reading and comprehension, and English vocabulary.

Furthermore, variations in the project effects were realized across the 3 target classes whereby for all the subtask areas, larger effects were observed in Primary 1, as compared to other classes. This could be attributed to the positive attitude and abilities of the Primary 1 teacher for English and Literacy that was observed during the course of the intervention.

4.2. Recommendations

Based on the findings of this assessment, the following recommendations are made:

☑ **Increase on the length of the intervention**

For teachers to fully comprehend and effectively implement the knowledge and skills acquired from the EGR interventions, they need more time and practice. This would allow the project implementors to have series of follow-up sessions and it would also help teachers in fostering the learning culture initiated by the project through the creation of the community of practice cycles.

☑ **Training of teachers on materials development**

To maximize the impact of the EGR intervention, there is need for an extensive training of the teachers on materials development to enable them create authentic and appropriate teaching materials to reinforce the training on the use of Jolly Phonics and the Weekly Foundation Story. In addition to this, the schools also need support in terms of text books for reference.

☑ **Timing of the training**

The timing for the training needs to be agreed upon by the headteacher and the teachers. This should lead to suitable days when the teachers find it easy to attend the trainings. This would also reduce on cases where teachers find it difficult to attend the trainings as evidenced during one of the follow-up sessions when the turn-up of teachers was very low.

☑ **Extend the intervention to other classes**

During the EGR intervention, teachers recommended a similar training for teachers of transition and upper classes. This is mainly because they commended the content and approaches that were implemented in the lower classes. To the teachers, a training of this kind in other classes would facilitate a holistic approach to improving teaching and learning of English in their schools.

4.3. Areas for future research

Further research could be done to investigate how teachers' attitudes and pedagogical skills influence acquisition, retention and application of knowledge and skills acquired from the intervention.

Further research could also be done to investigate the extent and scope of teamwork among teachers instilled by the EGR intervention and how it influences performance in literacy in refugee contexts.