USAID
FROM THE AMERICAN PEOPLE

PERFORMANCE & IMPACT EVALUATION (P&IE) OF THE USAID/UGANDA SCHOOL HEALTH AND READING PROGRAM: RESULT 1 INTERVENTIONS

Impact Evaluation Report, Cluster 2, Year 1

30 APRIL 2015

This publication was produced at the request of the United States Agency for International Development. It was prepared independently by Alicia Menendez and Yvonne Cao at NORC at the University of Chicago.

PERFORMANCE & IMPACT EVALUATION (P&IE) OF THE USAID/UGANDA SCHOOL HEALTH AND READING PROGRAM: RESULT 1 INTERVENTIONS

IMPACT EVALUATION REPORT, CLUSTER 2, YEAR 1

30 April 2015

PN 7384; USAID Contract N0: AID-617-C-12-00006

PRESENTED TO:
USAID/Uganda
Joseph Mwangi

PRESENTED BY:
NORC at the University of Chicago
4350 East-West Highway, 8th Floor
Bethesda, MD 20814
Telephone: (301) 634-9413
Fax: (301) 634-9301

CONTENTS

INTRODUCTION

NORC at the University of Chicago, in collaboration with Panagora, is conducting the Impact and Performance Evaluation of USAID/Uganda's School Health and Reading Program (SHRP), implemented by RTI. RTI (referred to as the implementing Partner or IP in the rest of the document) is implementing SHRP as two separate activities: (1) Activities related to Result 1, Improved Early Grade Reading and Transition to English, and (2) activities related to Result 2, Improved HIV/AIDS Knowledge, Attitudes and Practice. This report focuses on activities related to Result 1 (early grade reading) and presents the findings of the impact of the intervention on Cluster 2 students after one year of intervention.[1]

The report begins with a description of SHRP's Improved Early Grade Reading and Transition to English (Result 1) activities, followed by a description of the evaluation design, an analysis of baseline balance, and findings from the impact analysis.

A. PROJECT DESCRIPTION

For its Improved Early Grade Reading and Transition to English (Result 1) intervention, SHRP focuses on the nexus of language, pedagogy, and instructional materials to significantly improve students' early grade reading and literacy scores, as well as bring to scale a "Ugandan led 'reading policy'" (RTI International, 2012, p. 1). To this end, the program will build institutional capacity, support policy development and help institutionalize the training, support structures and policies necessary for sustainability. The intervention works at multiple levels. In early grade reading (Result 1), they consist of the following support to the MoES:

- At the school level, the intervention is comprised of training teachers in early grade literacy instruction using students' mother-tongue in P1-P3 and with a transition to English in P4; distribution of textbooks and instructional materials and teacher guides in local languages and English; support supervision provided to teachers; trainings in leadership for head teachers; and reading competitions and community awareness activities (reading awareness days, literacy week).
- At the district level, training for district education officials is raising awareness and building district-level support for the new curriculum.
- At the national level, MoES systems and pedagogical and language frameworks aim to strengthen and support mother-tongue based EGR and transition to English. The intent is to support the strengthening of policies related to reading; increase advocacy for reading at multiple levels (e.g. student, teacher, school, district, and national); and generate and use data for programmatic decision making.

Together, these interventions are expected to improve the instruction and learning environment of students and eventually lead to improved literacy skills.

Teacher trainings under Result 1 are implemented through Coordinating Center Tutors (CCTs), who are school support workers in charge of monitoring education quality within their Coordinating Centers (CC). Under the district education structure, each CCT is responsible for a certain number of schools

[1] At the time of submission of this report, the Result 1 team had collected data for: Round 1 for Cluster 1 (Feb-Mar 2013), Round 2 for Cluster 1 (Oct 2013), Round 3 for Cluster 1 (Oct 2014), Round 1 for Cluster 2 (Feb 2014) and Round 2 for Cluster 2 (Oct 2014), and the Result 2 team had collected baseline KAP data, but no follow-up data. Therefore, this report focuses on the impact of Result 1 on Cluster 2 students, using Cluster 2 Round 1 and Round 2 data. The results of the impact evaluation for Cluster 1 students is presented in a separate report. The first impact evaluation for Result 2 will only be carried out after the first follow-up data collection scheduled for 2015.

within a district (one district typically has multiple CCTs). The CCTs selected for the intervention receive training directly from SHRP and, in turn, deliver teacher training and program support in their schools, thus following a Training of Trainers (TOT) model.

Teacher trainings at the school level focus on pedagogy with an emphasis on using structured lesson plans and learner books. These lesson plans provide teachers with a practical step-by-step process for implementing the transitional bilingual approach mandated by the Ugandan EGR policy. In addition, SHRP is developing materials to support early grade reading. These materials are being adapted in order to take into account the different needs of learners at different stages of cognitive and academic development, and the linguistic characteristics of the different local languages, rather than be translated directly from one language to another. Furthermore, in order to develop these materials, SHRP works with MoES and Local Language Boards (LLBs) to standardize orthographies of the target languages. All materials follow the same general pedagogical framework to facilitate guidelines for textbook development and teacher training. Teachers receive these instructional materials in their local language of instruction and English. Teachers also receive ongoing monitoring support supervision from SHRP staff who observe classes and provide constructive feedback.

Additionally, at the district level, workshops and training for district education officials are serving as a forum for raising awareness and building district-level support for the new curriculum. Communication campaigns directed at Members of Parliament and district leaders are also organized.

Finally, at the national level, SHRP is working with MoES and the Sector Policy Management Working Group to develop a Uganda-specific reading strategy, which will include policies in the areas of Local Language Board development, textbook development, printing as well as Special Needs. In particular, SHRP is working to develop a national literacy strategy, national reading standards and benchmarks, as well as to harmonize reading assessment efforts with the Uganda National Examination Board (UNEB). SHRP is also working to strengthen Local Language Boards (LLB) and to raise awareness about special needs education. SHRP is also assisting the MoES in advocating for reading outside the classroom. Together with MoES, SHRP aims to raise awareness of local language development, reading instructions and special needs learners by using national communication campaigns through mass media and mobilizing local communities.

In total, SHRP will work in 12 local languages, developing instructional materials for each language and providing training and in-service support to teachers in areas where these 12 languages are spoken and used for mother-tongue based instruction in the early grades. The intervention is implemented following a pipeline roll-out, with the intervention targeting the districts associated with the first four local languages starting in Year 1 (Cluster 1), then targeting the districts associated with an additional four local languages starting Year 2 (Cluster 2), and finally targeting the districts associated with the final four local languages starting in Year 3 (Cluster 3). As such, Cluster 1 students will receive the intervention for a total of 4 years, Cluster 2 students for 3 years and Cluster 3 students for 2 years.

Over the 5 years of the project, SHRP will target the following local languages and associated districts:

Cluster	Local Language	Region	Districts
1	Luganda	Central	Wakiso, Gomba
	Runyankore/Rukiga	South West	Kiruhura, Bushenyi, Kabale
	Ateso	Eastern	Kumi, Katakwi, Serere
	Leblango	Northern	Apac, Lira, Kole
2	Runyoro/Rutoro	Mid-Western	Masindi, Kyenjojo, Kbarole
	Acholi	Mid-Northern	Gulu, Pader, Kitgum
	Lugbarati	West Nile	Arua
	Lumasaaba	Mid-Eastern	Mbale, Sironko, Manafwa
3	Lugwere	Mid-Eastern	Budaka, Pallisa, Kibuku
	Nkarimojong	North East	Nakapiripirit, Napak, Abim
	Lukhonzo	Mid-Eastern	Kasese
	Lusoga	East Central	Iganga and Kamuli

A.1 SHRP IN 2014 FOR CLUSTER 2

In 2014, SHRP started rolling out the intervention for Cluster 2 students working in the following local languages: Runyoro-Rutooro, Lumasaaba, Lugbarati, and Acoli. SHRP developed pupil primers and teacher guides for P1 and P2 teachers in all the Cluster 2 languages, as well as in English. All materials were distributed on time before the start of the school year in February. A total of 67,338 pupil primers in the Cluster 2 languages were distributed. In addition, SHRP delivered trainings through a three-tiered cascade model, to master trainers, teacher trainers and teachers and conducted two trainings in 2014, a 5-day training in January before the start of the school year, and a 3-day refresher training in May between the second and third school terms. In preparation for the first training in January, SHRP trained more than 100 Master Trainers and almost 300 Teacher Trainers. Overall, SHRP trained 3,012 teachers in January (including both Cluster 1 and Cluster 2 teachers). SHRP also conducted a leadership and management training in August-September 2014 attended by 3,997 teachers. SHRP continued support supervision monitoring to provide in-service support to teachers on Early Grade Reading through lesson preparation and classroom observation. SHRP also worked with International Book Bank and Books for Africa to select supplementary readers from various publishers and distributed them to schools in Kabarole and Kyenjojo districts. Teachers from those schools received training in library management in June-July 2014.

In Year 2, SHRP organized one communication campaign targeted at Members of Parliament and district leaders. Field Assistants also organized advocacy meetings in 60 communities. Program staff, education officials and field assistants initiated greater community mobilization and advocated for literacy instruction at Parent Teacher Association and School Management Committee meetings.

At the national level, SHRP continued to work on Cluster 2 language orthographies with the local Language Boards. The program also supported the Special Needs Education unit of the MoES to develop programs and activities, such as development of instructional materials that raise teachers' awareness of special needs students.

B. EVALUATION DESIGN

An impact evaluation (IE) is conducted to assess the causal effect of a specific intervention on a set of outcomes. It allows us to attribute changes in an outcome to a specific intervention or set of interventions by answering the counterfactual question "What would have happened to program

participants in the absence of the intervention?" Ideally, this is done by observing the same program participants both with and without the intervention at the same point in time. Of course, this is not possible; at any given time, a participant either receives the intervention or not. Therefore, we can never directly observe the counterfactual and, instead, need to create a comparison group to serve as the counterfactual. Identifying a credible comparison group is a critical aspect of an impact evaluation.

The ideal comparison group stems from the use of experimental methods in which eligible participants are randomly assigned to receive the intervention or not. Randomization ensures that, on average, characteristics of the treatment and control groups are statistically identical, with the only difference being their participation in the intervention. In this case, any measured difference in outcomes between the groups over time can be attributed to the program.

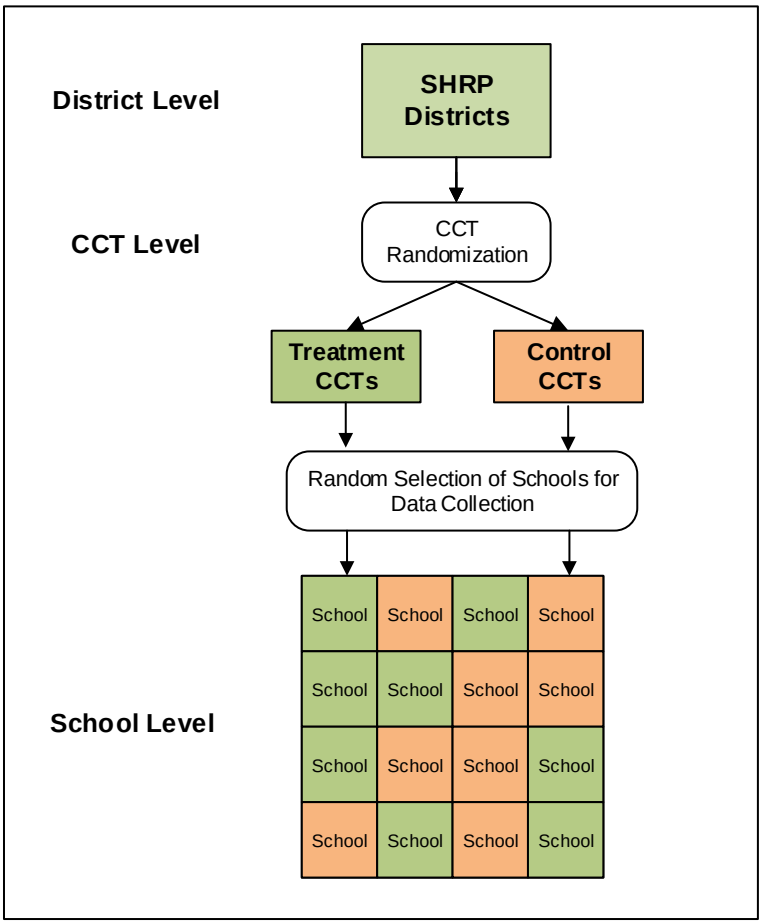
To assess the impact of the SHRP school-level intervention on Cluster 2 students, NORC is implementing a randomized controlled trial (RCT).

B.1 ASSIGNMENT OF SCHOOLS TO TREATMENT AND CONTROL GROUPS AND ESTIMATION STRATEGY

The CCTs are responsible for training teachers and providing follow-up support and assistance in implementing the Result 1 interventions. Towards this end, SHRP conducted training workshops (training of trainer workshops) for CCTs in different regions. Since each CCT is responsible for several schools, randomizing at the school level would imply that a CCT would have to treat schools under his or her jurisdiction differently if some were designated as treatment schools and others as controls. Therefore, randomization at the school level had a high risk of 'contamination' between treatment and control groups. To avoid this potential problem, we opted to randomize at the CCT level, assigning the entire cluster of schools under a CCT to either the treatment or the control group.

Figure 1 illustrates the evaluation design for Cluster 2, i.e. random assignment into the treatment or control group within treatment districts.

Figure 1. Assignment to Treatment and Control Groups



The difference in outcome indicators between treatment schools (green) and control schools (orange) will show the effect of the school-level intervention. We estimate the impact of the intervention for each language within a given time period on a number of literacy outcomes using a difference in differences model. To evaluate the impact of the intervention package on literacy score, *Y*, for students in grade P1 within a specific language group, we use the scores collected at the beginning of the school year 2014 (baseline) and at follow-up (in this case, at the end of one year of intervention, October 2014) and regress the individual test score *Y* of student *i* on the treatment status *T* of the school *s* in district *d,* a dummy variable *P* indicating period (i.e. 0 or 1 respectively for baseline and endline) and their interaction *T*P*:

$$Y_{isdp} = \beta_0 + \beta_1 T_{sd} + \beta_2 P + \beta_3 T_{sd} * P + \beta_4 D_d + \beta_5 X_i + \beta_6 Z_s + \varepsilon_{isd}$$

where X_i are individual characteristics of the student *i*, such as age and sex, Z_s are school characteristics such as language of instruction, D_d is a vector of dummy variables indicating the districts. As we mentioned T_{sd} is a dummy equal to 1 if school *s* in district *d* received the intervention and 0 otherwise, *P* indicates the period –baseline or endline, therefore, our coefficient of interest is $\beta 3$, which shows the differential effect of the treatment at the endline.

B.2 IMPACT INDICATORS AND DATA COLLECTION INSTRUMENTS

Literacy is comprised of multiple skills, both receptive and productive. Successful readers must be able to identify letters and their corresponding sounds, segment and blend those sounds to form words and

sentences, master appropriate vocabulary, and make meaning from text, among other skills. They must also be able to demonstrate their understanding and engagement with text through writing. To assess the effectiveness of the SHRP in reaching its goal to improve early grade reading and transition to English, specific key literacy skills were assessed.

The consensus among the reading research community in the United States is that effective reading instruction attends to *at least* five main reading skill areas including alphabetics (letter knowledge and phonemic awareness), fluency, vocabulary, and comprehension and phonics (National Reading Panel, 2000; Snow, Burns & Griffin, 1998).[2] Based on this research, the Early Grade Reading Assessment, (EGRA) a brief oral reading assessment that tests these skills, is used to measure program impacts on literacy (RTI International, 2007).

EGRA is comprised of multiple sub-tests that focus on the five main reading skill areas outlined above. Within each of these five areas, there are multiple sub-tests that can be selected for inclusion, based on local needs and the goals of the assessment system.

No clear benchmarks for the EGRA tool have been established. That is, the EGRA tool provides a snapshot of early literacy skills but does not provide guidelines for interpreting which children can be considered "readers" or what level of performance should be expected on each sub-test. At the same time, EGRA has been used to assess early literacy skills in more than 50 countries around the world; thus, performance of students participating in SHRP can be compared with the range of performance of other children on EGRA in other low-income countries.

A notable component of SHRP is its transitional bilingual design. That is, literacy instruction begins in one of four mother tongue languages, with English introduced as a subject area nearly simultaneously (within 4-8 weeks after mother tongue instruction has begun). The language of instruction will then increasingly transition from mother tongue to English over the course of four years. Because of this transitional bilingual design, the impact evaluation necessarily requires a heteroglossic[3] approach to assessment. Therefore, early literacy skills are assessed in mother tongue and English.

This transitional bilingual design affected the selection of sub-tests included in the EGRA tool in each language. For example, because most grade 1 students cannot be expected to have prior knowledge in English language or literacy, the sub-tests that have been selected to assess literacy in English are aimed at capturing lower skill levels; in contrast, students are expected to already possess basic linguistic knowledge in their mother tongues and the EGRA sub-tests that have been selected aim to capture a distribution of literacy skills that include higher level abilities.

[2] These five skills are not meant to be all inclusive; however, considerable empirical research has been conducted in these skill areas that has indicated they are important predictors of reading.

[3] A "heteroglossic" approach conceptualizes literacy learning in both mother tongue and English as interconnected, co-existing, and mutually reinforcing.

Table **1** describes the sub-tests that are included in English and mother tongue EGRA.

Table 1. Early Literacy Skills, EGRA Subtasks

Early Literacy Skill	Sub-test	Measurement	English	Mother -Tongue
Alphabetic Knowledge	Letter Sound Knowledge	Number of letter sounds correctly identified out of 100 in 60 seconds	X	X
Phonemic Awareness	Segmenting	Number of words correctly segmented out of 10 words	X	X
Phonics/Alphabetic Principles	Nonword decoding	Number of nonwords correctly decoded out of 50 in 60 seconds	X	X
Fluency	Oral passage reading	Number of words in a reading passage of approximately 50-60 words read fluently (with accuracy) in 60 seconds	X	X
Reading Comprehension	Oral recall	Number of questions (out of five) about a reading passage (read by student) answered correctly	X	X
Listening Comprehension	Oral recall	Number of questions (out of three) about an passage read aloud (by facilitator) answered correctly	--	X
Receptive Vocabulary	Oral identification of common objects	Number of common objects correctly identified	X	--

In addition to the EGRA, the following data collection instruments were developed and administered:

- Learner Context interview: to determine learner attendance to pre-school, socio-economic status and home literacy environment
- Teacher interview: to determine language of instruction in the school, teacher qualifications (including attendance to SHRP trainings) and the amount of support received from head teachers and CCTs
- Head Teacher interview: to obtain school enrollment information and determine participation in SHRP Training and amount of support provided to teachers
- School inventory: to determine quality of school environment (access to electricity, water, functioning toilets/latrines, availability of a school library)
- Classroom observation: to determine the extent to which teachers are applying SHRP teaching practices

B.3 SAMPLING

The standard approach to determining sample size for analytical surveys is to estimate the sample size required to achieve a specified level of power (probability), such as 90 percent, for detecting a change of a specific magnitude. This sample size depends on a number of factors including the evaluation design, the impact estimate, design of the sample survey used to collect data, the statistical test, and the population under investigation.

An initial group of districts located in the 4 different language areas for Cluster 2 was selected by the IP and MoES to participate in the intervention. NORC calculated the number of schools needed in each language group and within each group, RTI selected the CCTs to be assigned to treatment and control

groups and the requisite sample of 56 treatment schools and 56 control schools using random assignment.

For the impact evaluation of the Reading Program, we estimated the sample required to detect a double-difference measure of impact of magnitude D = 0.20 with a power of 90%. These calculations determined that 14 schools were necessary in the treatment group and in each control group for each language, and that 30 students would be sampled for each school, for a total of 420 treatment and 420 control students per language. With 4 languages, the total sample size required amounted to 3,360 students from 12 schools per data collection round. Of the 112 schools, half (56 schools) would constitute the treatment group, and half the control group. Table 2 shows the target sample size for each round of data collection.

Table 2. Target Sample Size for Each Round of Data Collection

Language Group	Treatment	Control	Total
Runyoro-Rutooro	14 schools 420 students	14 schools 420 students	28 schools 840 students
Lumasaaba	14 schools 420 students	14 schools 420 students	28 schools 840 students
Lugbarati	14 schools 420 students	14 schools 420 students	28 schools 840 students
Acoli	14 schools 420 students	14 schools 420 students	28 schools 840 students
Total	**56 schools** **1,680 students**	**56 schools** **1,680 students**	**112 schools** **3,360 students**

C. DATA COLLECTION AND FINAL SAMPLE

Data collection was conducted at the beginning and at the end of the 2014 school year by the IP and its local partner Center for Social Research (CSR). The data collection was conducted using tablets in Tangerine software and consisted of EGRAs in English and 4 local languages (each student was assessed in his/her local language and in English), followed by a learner context interview. At each school, field teams also administered a teacher and head teacher interview and a school inventory. Finally classroom observations were also conducted for a subset of schools (~10%).

C.1 ASSESSOR TRAINING[4]

In preparation for data collection, RTI engaged 68 assessors for the Cluster 2 baseline training from February 3 to February 11, 2014 and over 100 assessors for the Cluster 1 and 2 follow-up training conducted from September 29 to October 1, 2014 and which was preceded by an additional training for new assessors from September 24 to September 26, 2014The best assessors were then selected based on inter-rater reliability tests (IRR) that were given throughout the week as well as interpersonal and leadership skills. Technical training, which was undertaken by School Health and Reading Program staff (RTI and CSR), included hands-on practice, where assessors spent one day in a school administering the tool to learners and teachers. Trainees were first trained to administer the tools on paper and then

[4] From RTI EGRA Baseline Report

introduced to electronic data collection on Nexus tablets so that they could be prepared for both circumstances. Four Data Quality Assurance (DQA) Officers who also acted as assessor trainers, as well as new assessors were also given an extra day of training.

C.2 DATA COLLECTION

Data were collected from February 13 through March 28, 2014 for the baseline round and from October 6 through October 24, 2014 for the endline round. Forty-eight assessors in teams of four (each of these teams included one supervisor) were deployed to the four language areas. Each of the four language area teams was supported by a DQA Officer. This DQA Officer was responsible for overseeing all aspects of data collection deployment, observing assessors and providing feedback and support, ensuring data were uploaded from the electronic tablets every evening, and observing the reading classes. Besides the team supervisors and the DQA Officers, data collection was overseen by MoES staff (from the Directorate of Education Standards and from Guidance and Counseling), Uganda National Examination Board (UNEB) staff, School Health and Reading Program staff, and staff from the external evaluation team from NORC.

C.3 SCHOOLS AND LEARNERS SAMPLED

The final sample included in the analysis consists of a panel of 56 schools. In those schools, a total of 3,240 students were tested at baseline (96.4% of target of 3,360 students) and 3,178 students at endline (94.6% of target of 3,360 students). Table 3 presents the number of students included in the analysis; the number of schools is indicated in parenthesis.

Table 3. Final Sample Used in the Analysis

	Treatment		Control	
	Baseline	**Endline**	**Baseline**	**Endline**
Acoli	390	380	400	375
	(14)		(14)	
Lugbarati	413	408	408	415
	(14)		(14)	
Lumasaaba	416	423	399	400
	(14)		(14)	
Runyoro-Rutooro	402	408	404	377
	(14)		(14)	
Total	**1,621**	**1,619**	**1,611**	**1,567**
	(56)		**(56)**	

D. BALANCE AT BASELINE

In order to explore comparability of pre-treatment characteristics of the treatment and control groups, we conducted mean equality tests to test for balance. Because comparisons between languages cannot be made given that no benchmarking and psychometric analysis has been performed to assess comparability between languages, mean equality tests of local language literacy scores are conducted within language subgroup. On the other hand, tests for English literacy scores are conducted on the entire sample.

Table 4. English literacy scores, demographics and other characteristics at baseline for treatment and control groups. All language groups.

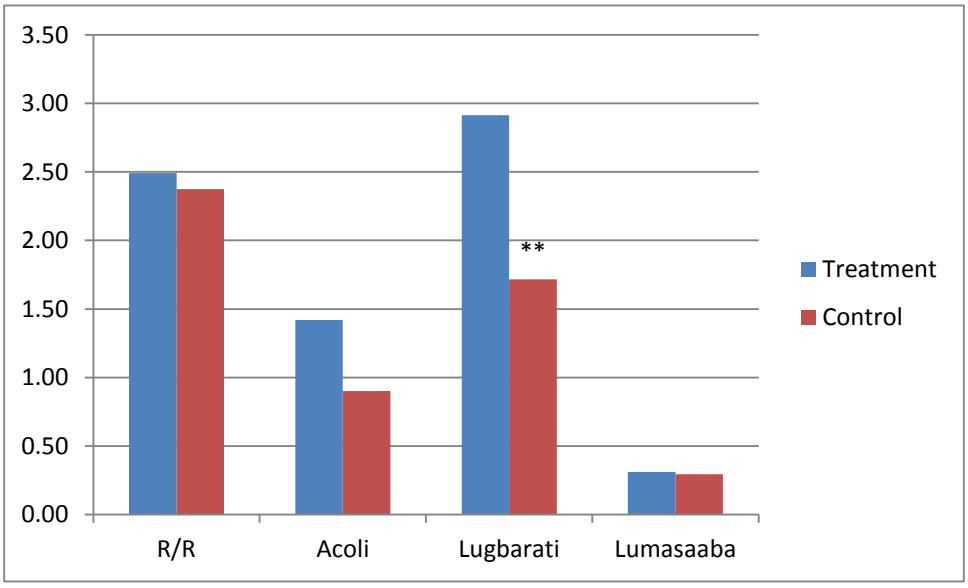
Variable	Treatment Mean (SE)	Control Mean (SE)	Difference (C – T)
Age of student	7.645 (0.051)	7.704 (0.051)	0.059
% of cases with missing age information	0.183 (0.010)	0.199 (0.010)	0.016
Gender of student (female = 1)	0.497 (0.012)	0.490 (0.012)	-0.007
Number of assets (max = 8)	2.464 (0.034)	2.448 (0.032)	-0.016
Lives with both parents (yes = 1)	0.579 (0.012)	0.644 (0.012)	0.065**
Does not live with mother (yes = 1)	0.187 (0.010)	0.173 (0.009)	-0.014
Reads at home (yes = 1)	0.422 (0.012)	0.417 (0.012)	-0.005
Attended preschool (yes = 1)	0.418 (0.012)	0.386 (0.012)	-0.032
Student absent any day in the week prior to assessment (yes = 1)	0.443 (0.012)	0.446 (0.013)	0.003
Teacher absent any day in the week prior to assessment (yes = 1)	0.455 (0.013)	0.491 (0.013)	0.036
Gender of teacher (female = 1)	0.700 (0.012)	0.679 (0.013)	-0.021
Letter sound score (max = 100)	0.913 (0.066)	0.585 (0.044)	-0.328*
Word segmenting score (max = 10)	4.349 (0.041)	4.565 (0.033)	0.216**
Nonword decoding score (max = 50)	0.035 (0.014)	0.009 (0.005)	-0.026
Oral reading fluency (max = 68)	0.071 (0.017)	0.045 (0.009)	-0.026
Reading comprehension score (max = 5)	0.001 (0.001)	0.000 (0.000)	-0.001
Receptive vocabulary score (max=20)	4.807 (0.061)	4.431 (0.058)	-0.376

Overall, our analysis indicates that pre-treatment literacy scores between treatment and control groups are balanced. **Error! Reference source not found.** shows that the treatment and control groups are similar in sociodemographic and other characteristics with the exception of the indicator "lives with both parents" which is slightly higher among control learners. The baseline scores in English literacy

differs between treatment and control for the letter sound and the word segmenting subtask. In the letter sound subtask the treatment group performs better than the control group while for the word segmenting subtasks the inverse is true.

Literacy scores in local languages, as well as demographics and other characteristics, are similarly balanced between treatment and control groups within each language subgroup. Figure 2 provides the only example where we find one difference; for the Lugbarati group, treatment students scored significantly better than control students on letter sound knowledge. Differences between treatment and control groups for R/R, Acoli and Lamsaaba are not statistically significant. None of the other literacy scores showed statistically significant differences between treatment and control groups. Annex A shows the results of all variables for all language subgroups.

Figure 2. Local language letter sound knowledge score, by language subgroup



** significant at 5% level

Overall, the sample at baseline is quite well-balanced between treatment and control, although we have found some differences, particularly in letter sound and segmentation in English and in letter sound in Lugbarati. In our estimation of the program impact we, therefore, control for these baseline differences.

E. IMPACT ANALYSIS RESULTS

This section summarizes the results of the impact analysis after one year of intervention for Cluster 2 schools. Our analysis is conducted for each language subgroup for each of the outcome scores on subtasks conducted in the local language. On the other hand, for outcome scores on English subtasks, our analysis is applied to the entire sample (pooled sample from all four language subgroups). It is important to keep in mind that this evaluation does not allow for comparison of progress and impact between languages as the EGRA tools were developed independently for each local language and no psychometric analysis has been conducted to determine whether EGRA scores in one language can be compared to EGRA scores in another language. Therefore, while we can determine whether SHRP has had an impact for a particular language subgroup vs. another language subgroup, we cannot comment on the relative magnitude of these impacts between language subgroups, except for the English scores.

First, we study the effect of the school-level intervention following the model in equation (1). We present findings for scores in local language within each language subgroup and in English for the pooled sample. In the tables below, we report the average treatment effect on each outcome of interest and for each language subgroup using different model specifications to check for robustness. All regression models include individual controls (sex and age of the student, a dummy for whether age is missing, household asset index, dummies for living with both parents, having someone read to the student at home, and language spoken at home)[5] and district dummies. For the regressions on English scores using the pooled sample, we include dummies for local language subgroups to control for potential differences between them. We also test results with and without school fixed effects. All standard errors are robust standard errors and allow for correlation in the unobservables between learners in the same class. All regressions are Ordinary Least Squares (OLS) regressions.

E.1 IMPACT ON LETTER SOUND KNOWLEDGE

Impact on Letter Sound scores

Each cell in Table 5 shows the average effect of the treatment (SHRP intervention) on Letter Sound Score for each language subgroup using 5 different model specifications. Since schools were randomly assigned to the treatment and control groups, and mean equality tests show that both groups are well-balanced at baseline, we start with the simplest model (model (1)) which does not include any controls. To reach more precision and to check for robustness, we add controls in models (2) through (5).

- Model (1) includes no controls.
- Model (2) is similar to model (1) but includes districts fixed effects.
- Model (3) is similar to model (2) but adds individual controls.
- Model (4) is similar to model (3) and also includes school fixed effects.
- Model (5) is similar to model (4) but only includes the subsample of students 6 years and older. We do this because a somewhat considerable proportion of students interviewed at baseline and endline reported being 5 years of age even though the official age for starting P1 in Uganda is 6 years old. At baseline, more than 9 percent of the students in our sample reported being 5 years of age and at follow-up, 4.3 percent reported being 5 years old. Although it may be possible that some students officially enroll in P1 before the age of 6, it is also possible that field teams sampled children who were too young to attend P1 but they attend the class. For this

[5] With the exception of models (1) and (2) in **Error! Reference source not found.**.

analysis we decided to include a model that only analyzes the subsample of students who reported being 6 years and older, i.e. students officially enrolled in P1.

With the simplest model (column (1)), we find that, in its first year, the intervention did not have an effect on letter sound scores in any local language with the exception of Runyoro-Rutooro. We find that results are very similar across all other models; the intervention for this cluster, in its first year, did not have an impact on letter sound scores except for the Runyoro-Rutooro subgroup for which treatment students scored on average around 2.5 (0.35 standard deviation) letter sounds higher than control students. Both groups started with a baseline score of around 2.4-2.5 letter sounds per minute. At the end of the academic year, while the Runyoro-Rutooro control group scored only 5 correct letter sounds, the Runyoro-Rutooro treatment group averaged 7.5 correct sounds. This difference is significant as a percentage however the improvement of both groups during the academic year is small in absolute terms.

In English, we observe a small difference, 0.7 words (0.13 standard deviation), in favor of the treatment group which is significant at the 10 percent level. Again the only group that shows a positive effect of the program, on the English letter sound subtask, is the Runyoro-Rutooro language subgroup.

Table 5. Cluster 2 SHRP School Level Effect - Letter Sound Score

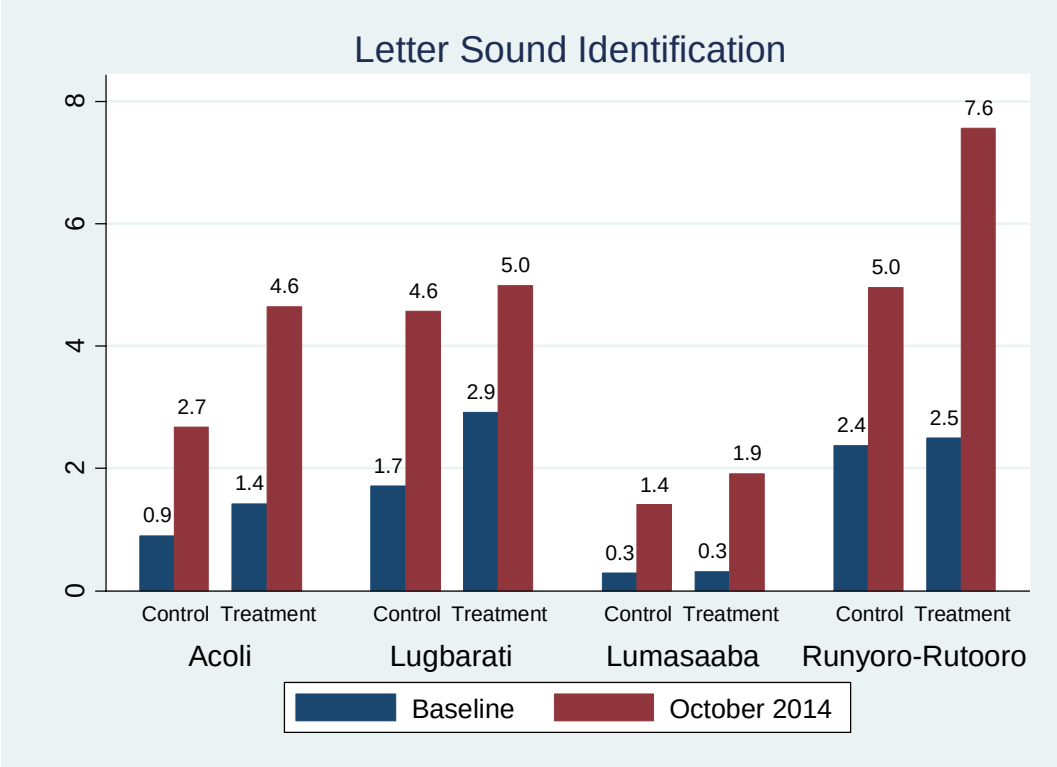
Language	(1)	(2)	(3)	(4)	(5) 6 years and older
Runyoro-Rutooro	2.484**	2.474**	2.485**	2.403**	2.871***
	(0.931)	(0.929)	(0.985)	(0.967)	(1.005)
Lumasaaba	0.491	0.487	0.507	0.660*	0.700
	(0.405)	(0.404)	(0.398)	(0.384)	(0.516)
Lugbarati	-0.783	-0.783	-0.943	-0.849	-1.147
	(0.616)	(0.616)	(0.637)	(0.622)	(0.675)
Acoli	1.450	1.464	1.518	1.641*	1.641
	(0.937)	(0.938)	(0.934)	(0.938)	(1.025)
English	0.728*	0.727*	0.720*	0.774*	0.722
	(0.421)	(0.421)	(0.428)	(0.423)	(0.469)
Districts Fixed Effects	no	yes	yes	yes	yes
Individual Controls	no	no	yes	yes	yes
School Fixed Effects	no	no	no	yes	yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Individual controls include age and sex of learner, dummy for age missing, score on household assets index, dummies for student living with both parents, someone at home reading to the student and language spoken at home. Regressions for English scores include controls for local language.

Figure 3 shows the average number of letter sounds identified correctly by treatment and control learners at baseline and in October 2014 by local language. As explained above, the only statistically significant difference in knowledge change between the treatment and control groups is found in the Runyoro-Rutooro language subgroup. The Runyoro-Rutooro is also the subgroup that shows the

highest average level of letter sound knowledge among learners. In general, we find that the absolute performance scores of students for this task is very low, even at the end of P1, with treatment students only able to identify 1.9 letter sounds for the Lumasaaba language group, 4.6 letter sounds for Acoli, 5 for Lugbarati, and 7.6 for R/R after one year of the SHRP school-level interventions.

Figure 3. Letter Sound scores by language sub-group



Impact on zero scores for Letter Sound Knowledge

Figure 4 shows that the percentage of learners that are not able to identify a single letter sound correctly remains very high at the end of P1. The language subgroup that shows the worse performance is Lumasaaba, were 70% and 80% of the learners in treatment and control groups respectively were not able to identify a single letter sound. In the Acoli and Lugbarati subgroups the percentages are somewhat lower but still more than 50% of the learners cannot identify any letter sound. The lowest percentages were found in the Runyo-Rutooro subgroup where 40% and 30% of the students in control and treatment respectively scored zero in this EGRA subtask.

Figure 4. Proportion of Students with Zero Scores on Letter Sound Subtask

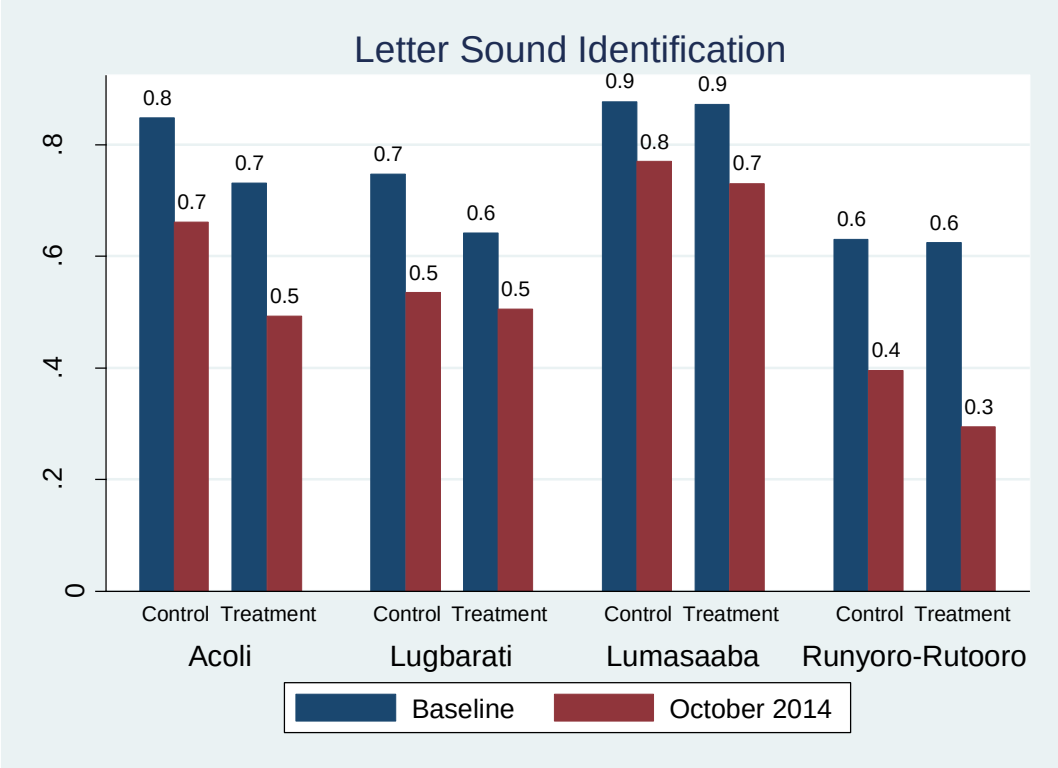


Table **6** shows the average treatment effect on the percentage of students who could not identify any letters in the letter sound knowledge subtask (percentage of zero scores). The regressions all include district fixed effects, individual controls and a constant term; columns (1) and (2) show the results of regressions without and with school fixed effects respectively. The intervention had no impact on the percentage of students who scored zero in any language. This suggests that the positive effect that we observed in the number of correctly identified sounds in Runyoro-Rutooro is probably the result of increasing knowledge among those learners that already had some familiarity with letter sounds rather than a large reduction in the number of those that do not.

Table 6. SHRP School Level Effect - Percentage of students who scored zero on letter sound subtask

Language	(1)	(2) With school fixed effects
Runyoro-Rutooro	-0.086	-0.078
	(0.063)	(0.062)
Lumasaaba	-0.045	-0.054
	(0.050)	(0.050)
Lugbarati	0.080*	0.076
	(0.046)	(0.045)
Acoli	-0.066	-0.070
	(0.074)	(0.075)
English	-0.018	-0.031
	(0.033)	(0.032)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include district fixed effects, individual controls and a constant term. Individual controls include age and sex of learner, dummy for age missing, score on household assets index, dummies for student living with both parents, someone at home reading to the student and language spoken at home. Regressions for English language include control for local language.

E.2 IMPACT ON WORD SEGMENTING SCORE

Table 7 shows the average treatment effect on word segmenting scores (regressions include individual controls, district fixed effects and a constant term). Students received one point for every word that they could segment correctly for a maximum of 10 points. We find a slightly significant small impact for the Lumasaaba subgroup for which treatment students scored approximately 0.5 words (0.28 standard deviation) better than the control students, although this effect vanishes when using the school fixed effect model. No impact is found for any other local languages, nor English word segmenting.

Table 7. SHRP School Level Effect - Word Segmenting Score

Language	(1)	(2) With school fixed effects
Runyoro-Rutooro	0.243	0.224
	(0.433)	(0.436)
Lumasaaba	0.469*	0.349
	(0.276)	(0.245)
Lugbarati	-0.285	-0.247
	(0.445)	(0.452)
Acoli	0.297	0.426
	(0.290)	(0.278)
English	-0.006	-0.013
	(0.121)	(0.120)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include district fixed effects, individual controls and a constant term. Individual controls include age and sex of learner, dummy for age missing, score on household assets index, dummies for student living with both parents, someone at home reading to the student and language spoken at home. Regressions for English language include control for local language.

E.3 IMPACT ON NON-WORD DECODING, ORAL READING FLUENCY, READING COMPREHENSION AND ENGLISH RECEPTIVE VOCABULARY

In terms of the impact of the intervention on higher level literacy skills, namely non-word decoding, oral reading fluency, reading comprehension and English receptive vocabulary, we find no impact whatsoever for any of the languages, with the exception of non-word decoding in English for which the positive effect is very small (0.15 standard deviation) and only borderline statistically significant. In Table 8, we show our findings using the school fixed effect models[6].

Table 8. SHRP School Level Effect - Non-Word Decoding, Oral Reading Fluency, Reading Comprehension and English Receptive Vocabulary

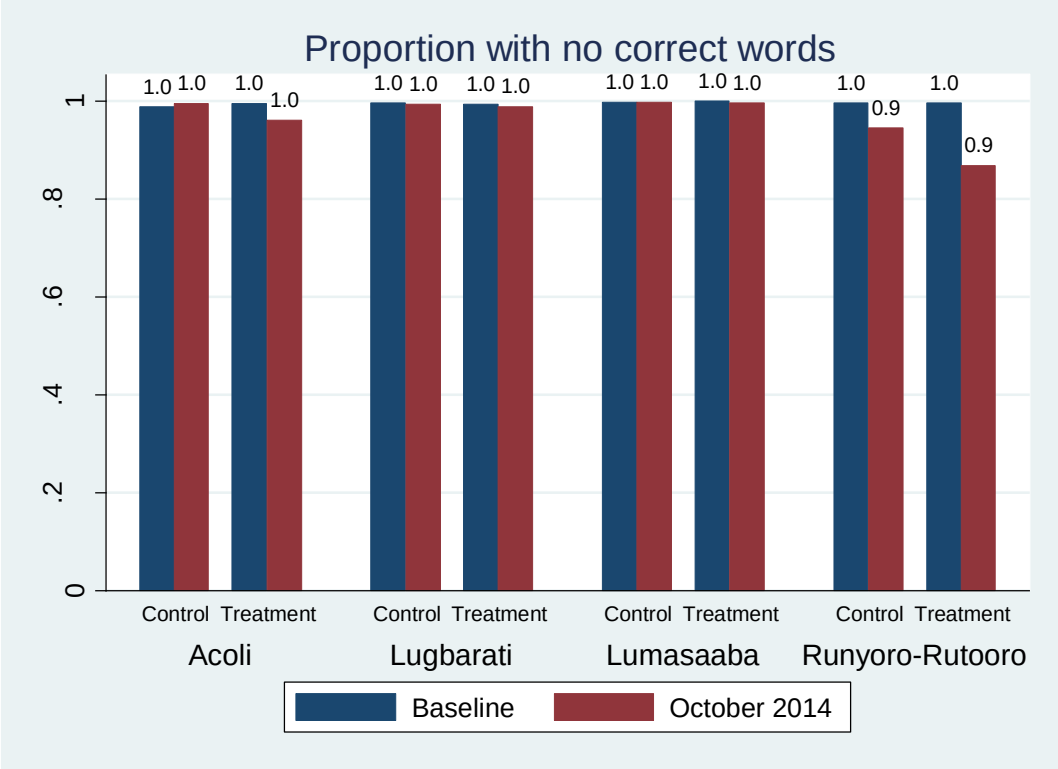
Dependent Variable	(1) Non-Word Decoding	(2) Oral Reading Fluency	(3) Reading Comprehension	(4) English Receptive Vocabulary
Runyoro-Rutooro	0.899	0.818	0.045	
	(0.564)	(0.501)	(0.033)	
Lumasaaba	0.002	0.042	**(a)**	
	(0.004)	(0.046)		
Lugbarati	0.001	0.022	**(a)**	
	(0.055)	(0.034)		
Acoli	0.090	0.169	0.003	
	(0.060)	(0.108)	(0.003)	
English	0.207*	0.137	0.003	-0.110
	(0.111)	(0.102)	(0.004)	(0.181)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include district and school fixed effects, individual controls and a constant term. Individual controls include age and sex of learner, dummy for age missing, score on household assets index, dummies for student living with both parents, someone at home reading to the student and language spoken at home. Regressions for English language include controls for local language.
(a) no positive scores reported.

In Figure 5 below we show the fraction of learners that cannot read a single word in a reading passage in their local language at the beginning and end of P1, after one year of the intervention, by language subgroup. For all language groups, except Runyoro-Rutooro, close to 100 percent of students cannot read even one word in their local language. Even in the Runyoro-Rutooro subgroup, only a few students - approximately 10 percent - are able to read at least one word.

[6] Results are very similar when we do not include school fixed effects.

Figure 5. Proportion of Students with Zero Scores on Oral Reading Fluency



E.4 IMPACT ON BOYS VS. GIRLS

Finally, in general we do not find differences between the performance of boys and girls. However, when focusing on letter sound identification, girls show an advantage over boys in Runyoro-Rutooro. This advantage is around 0.9 letter sounds over boys on average. We do not observe differences in any of the other languages and we do not observe differences in the impact of SHRP by gender.

E.5 DISCUSSION OF IMPACT OF SCHOOL-LEVEL INTERVENTION

Cluster 2. Impact of school-level intervention:

- Positive impact on letter sound in Runyoro-Rutooro subgroup, in local language and English.
- No impact on any other EGRA subtasks in Runyoro-Rutooro subgroup
- No impact for any EGRA subtask in any other local language (Lumasaaba, Lugbarati, Acoli)

In summary, these analyses suggest that the intervention among Cluster 2 learners, after one year of implementation, has only had an impact on students from one language subgroup (Runyoro-Rutooro students) and only on letter sound knowledge but not on phonemic awareness or the higher level skills – decoding non-words, oral reading fluency and reading comprehension.

The effect is also low in absolute terms. On average, after one school year, despite the presence of the SHRP literacy project, students still only know few letter sounds: 7.5 letter sounds for the Runyoro-Rutooro treatment group (up from 2.5 letter sounds), for example, while the proportion of students who cannot identify a

single letter sound at the end of the school year is 39 percent for the same group (down from 57 percent at the beginning of the year).

Another alarming finding is the fact that practically no students can read any words in a connected text in any of the language groups, except a small number in Runyoro-Rutooro. This means that for practically 100 percent of students, their oral reading fluency scores was zero word per minute at the end of one year of primary schooling.

As implementation continues to be rolled out, and students move up in grades, it is possible that impacts on higher level literacy skills could manifest themselves. Additionally, with time, we may see a gradual transfer of these skills from the local language to English, if the hypothesis that mother-tongue based early grade literacy with transition to English is the most effective bilingual literacy acquisition model holds. However, such higher level impacts must likely be preceded by substantial improvements in students' foundational reading fluency and comprehension skills.

Overall, the results from this first impact evaluation of SHRP on Cluster 2 students folllow a similar trend to the results that we found for the first impact evaluation of SHRP on Cluster 1 students. For Cluster 1, we did not find much impact of the intervention in year 1 on any of the subtasks except in Luganda but did see more positive impacts in Year 2, although absolute levels in literacy skills remain low. It is therefore possible that impacts of SHRP for Cluster 2 students will follow similar trends.

CONCLUSION

This report describes the impact of SHRP on Cluster 2 students after one year of implementation, focusing on Result 1 activities, i.e. those related to Early Grade Reading. In general, there has been little impact of the SHRP program on early reading skills. We found:

- Positive impact on letter sound in Runyoro-Rutooro subgroup, in local language and English.
- No impact on any other EGRA subtaks in in Runyoro-Rutooro subgroup.
- No impact for any EGRA subtask in any other local language (Lumasaaba, Lugbarati, Acoli).

Discussions with stakeholders, including USAID and the implementing agency may help uncover possible reasons regarding why positive impacts were only found in the Runyoro-Rutooro language subgroup, and why, in general, there have been little to no discernible impacts on reading skills of the SHRP Result 1 interventions during the first year of their implementation. It is possible that the intervention was implemented differently in Runyoro-Rutooro speaking districts than in other areas (in terms of quality and effectiveness of teacher trainings and/or instructional materials for instance) which may explain the fact that the intervention was effective in these districts but not in other districts. We intend to delve deeper into these inquiries during upcoming stakeholder meetings and results dissemination workshops.

ANNEX A. BALANCE AT BASELINE – TREATMENT VS. CONTROL

Table A.1. Treatment vs. Control, R/R Language

Variable	Treatment Mean (SE)	Control Mean (SE)	Difference (C - T)
Age of student	7.300	7.410	0.110
	(0.108)	(0.092)	
% of cases with missing age information	0.197	0.149	-0.048
	(0.020)	(0.018)	
Gender of student (female = 1)	0.478	0.490	0.012
	(0.025)	(0.025)	
Number of assets	2.505	2.604	0.099
	(0.067)	(0.058)	
Lives with both parents (yes = 1)	0.458	0.555	0.097*
	(0.025)	(0.025)	
Does not live with mother (yes = 1)	0.214	0.209	-0.005
	(0.020)	(0.020)	
Someone reads to student at home (yes = 1)	0.539	0.559	0.020
	(0.025)	(0.025)	
Attended preschool (yes = 1)	0.613	0.653	0.040
	(0.024)	(0.024)	
Student absent any day in the week prior to assessment (yes = 1)	0.451	0.462	0.011
	(0.025)	(0.025)	
Teacher absent any day in the week prior to assessment (yes = 1)	0.499	0.537	0.038
	(0.026)	(0.026)	
Gender of teacher (female = 1)	0.712	0.806	0.094
	(0.026)	(0.022)	
Teacher teaches in R/R (yes = 1)	0.871	0.861	-0.010
	(0.017)	(0.017)	
Speaks R/R at home (yes = 1)	0.614	0.681	0.067
	(0.024)	(0.023)	
R/R Letter sound score (max = 100)	2.493	2.375	-0.118
	(0.252)	(0.208)	
R/R Word segmenting score (max = 10)	5.570	5.355	-0.215
	(0.125)	(0.105)	
R/R Nonword decoding score (max = 50)	0.087	0.084	-0.003
	(0.047)	(0.049)	
R/R Oral reading fluency (max = 68)	0.037	0.070	0.033
	(0.035)	(0.061)	
R/R Reading comprehension score (max = 5)	0.005	0.005	-.000
	(0.005)	(0.005)	
R/R Listening comprehension	1.373	1.284	-0.089
	(0.059)	(0.057)	
English Letter sound score (max = 100)	1.465	1.243	-0.222
	(0.171)	(0.125)	
English Word segmenting score (max = 10)	4.306	4.509	0.203
	(0.086)	(0.070)	
English Nonword decoding score (max = 50)	0.124	0.020	-0.104
	(0.055)	(0.020)	
English Oral reading fluency (max = 68)	0.183	0.084	-0.099
	(0.056)	(0.022)	
English Reading comprehension score (max = 5)	0.000	0.000	0.000
	(0.000)	(0.000)	
English Receptive vocabulary score (max=20)	5.619	5.350	-0.269
	(0.132)	(0.115)	

*p<0.10, **p<0.05, ***p<0.001

Table A.2. Treatment vs Control, Lugbarati Language

Variable	Treatment Mean (SE)	Control Mean (SE)	Difference (C - T)
Age of student	8.668	8.764	0.096
	(0.095)	(0.097)	
% of cases with missing age information	0.140	0.098	-0.042
	(0.017)	(0.015)	
Gender of student (female = 1)	0.506	0.498	-0.008
	(0.025)	(0.025)	
Number of assets	2.483	2.561	0.078
	(0.064)	(0.065)	
Lives with both parents (yes = 1)	0.610	0.664	0.054
	(0.024)	(0.023)	
Does not live with mother (yes = 1)	0.215	0.174	-0.041
	(0.020)	(0.019)	
Someone reads to student at home (yes = 1)	0.334	0.323	-0.011
	(0.023)	(0.023)	
Attended preschool (yes = 1)	0.200	0.167	-0.033
	(0.020)	(0.019)	
Student absent any day in the week prior assessment (yes = 1)	0.432	0.455	0.023
	(0.024)	(0.025)	
Teacher absent any day in the week prior to assessment (yes = 1)	0.312	0.369	0.057
	(0.024)	(0.025)	
Gender of teacher (female = 1)	0.661	0.521	-0.140
	(0.023)	(0.026)	
Teacher teaches in Luganda (yes = 1)	0.816	0.819	0.003
	(0.019)	(0.019)	
Speaks Luganda at home (yes = 1)	0.966	0.828	-0.138***
	(0.009)	(0.019)	
Luganda Letter sound score (max = 100)	2.915	1.716	-1.199**
	(0.240)	(0.182)	
Luganda Word segmenting score (max = 10)	4.864	4.662	-0.202
	(0.138)	(0.128)	
Luganda Nonword decoding score (max = 50)	0.010	0.000	-0.010
	(0.010)	(0.000)	
Luganda Oral reading fluency (max = 68)	0.012	0.009	-0.003
	(0.008)	(0.006)	
Luganda Reading comprehension score (max = 5)	0.000	0.000	0.000
	(0.000)	(0.000)	
Luganda Listening comprehension	2.085	2.007	-0.078
	(0.049)	(0.053)	
English Letter sound score (max = 100)	1.576	0.713	-0.863***
	(0.178)	(0.099)	
English Word segmenting score (max = 10)	1.576	0.713	-0.863***
	(0.178)	(0.099)	
English Nonword decoding score (max = 50)	4.024	4.243	0.219
	(0.094)	(0.085)	
English Oral reading fluency (max = 68)	0.012	0.010	-0.002
	(0.006)	(0.008)	
English Reading comprehension score (max = 5)	0.046	0.015	-0.031
	(0.030)	(0.011)	
English Receptive vocabulary score (max=20)	0.002	0.000	-0.002
	(0.002)	(0.000)	

*p<0.10, **p<0.05, ***p<0.001

Table A.3. Treatment vs. Control, Lumasaaba Language

Variable	Treatment Mean (SE)	Control Mean (SE)	Difference (C - T)
Age of student	7.148	7.057	-0.091
	(0.095)	(0.100)	
% of cases with missing age information	0.317	0.381	0.064
	(0.023)	(0.024)	
Gender of student (female = 1)	0.502	0.459	-0.043
	(0.025)	(0.025)	
Number of assets	2.234	2.180	-0.054
	(0.074)	(0.068)	
Lives with both parents (yes = 1)	0.558	0.589	0.031
	(0.024)	(0.025)	
Does not live with mother (yes = 1)	0.214	0.233	0.019
	(0.020)	(0.021)	
Someone reads to student at home (yes = 1)	0.440	0.458	0.018
	(0.025)	(0.025)	
Attended preschool (yes = 1)	0.466	0.443	-0.023
	(0.025)	(0.025)	
Student absent any day in the week prior assessment (yes = 1)	0.546	0.527	-0.019
	(0.025)	(0.025)	
Teacher absent any day in the week prior to assessment (yes = 1)	0.614	0.717	0.103***
	(0.025)	(0.024)	
Gender of teacher (female = 1)	0.922	0.893	-0.029
	(0.014)	(0.018)	
Teacher teaches in Lango (yes = 1)	0.986	0.907	-0.079
	(0.006)	(0.015)	
Speaks Lango at home (yes = 1)	0.993	0.870	-0.123**
	(0.004)	(0.017)	
Lango Letter sound score (max = 100)	0.310	0.293	-0.017
	(0.046)	(0.054)	
Lango Word segmenting score (max = 10)	4.591	4.556	-0.035
	(0.078)	(0.072)	
Lango Nonword decoding score (max = 50)	0.000	0.000	0.000
	(0.000)	(0.000)	
Lango Oral reading fluency (max = 68)	0.000	0.043	0.043
	(0.000)	(0.043)	
Lango Reading comprehension score (max = 5)	0.000	0.000	0.000
	(0.000)	(0.000)	
Lango Listening comprehension	0.502	0.406	-0.096
	(0.035)	(0.035)	
English Letter sound score (max = 100)	0.195	0.120	-0.075
	(0.038)	(0.027)	
English Word segmenting score (max = 10)	4.500	4.744	0.244
	(0.071)	(0.050)	
English Nonword decoding score (max = 50)	0.000	0.005	0.005
	(0.000)	(0.005)	
English Oral reading fluency (max = 68)	0.012	0.055	0.043
	(0.008)	(0.026)	
English Reading comprehension score (max = 5)	0.000	0.000	0.000
	(0.000)	(0.000)	
English Receptive vocabulary score (max=20)	4.017	4.078	0.061
	(0.105)	(0.114)	

*p<0.10, **p<0.05, ***p<0.001

Table A.4. Treatment vs. Control, Acoli Language

Variable	Treatment Mean (SE)	Control Mean (SE)	Difference (C - T)
Age of student	7.339	7.316	
	(0.083)	(0.086)	-0.023
% of cases with missing age information	0.069	0.170	
	(0.013)	(0.019)	0.101***
Gender of student (female = 1)	0.503	0.513	
	(0.025)	(0.025)	0.010
Number of assets	2.646	2.443	
	(0.070)	(0.066)	-0.203
Lives with both parents (yes = 1)	0.692	0.770	
	(0.023)	(0.021)	0.078*
Does not live with mother (yes = 1)	0.100	0.075	
	(0.015)	(0.013)	-0.025
Someone reads to student at home (yes = 1)	0.378	0.326	
	(0.025)	(0.024)	-0.052
Attended preschool (yes = 1)	0.398	0.292	
	(0.025)	(0.023)	-0.106
Student absent any day in the week prior assessment (yes = 1)	0.337	0.338	
	(0.024)	(0.024)	0.001
Teacher absent any day in the week prior to assessment (yes = 1)	0.396	0.342	
	(0.025)	(0.025)	-0.054
Gender of teacher (female = 1)	0.479	0.545	
	(0.028)	(0.028)	0.066
Teacher teaches in Ateso (yes = 1)	0.918	0.890	
	(0.014)	(0.016)	-0.028
Speaks Ateso at home (yes = 1)	0.918	0.938	
	(0.014)	(0.012)	0.020
Ateso Letter sound score (max = 100)	1.421	0.900	
	(0.141)	(0.129)	-0.521
Ateso Word segmenting score (max = 10)	5.144	5.175	
	(0.091)	(0.102)	0.031
Ateso Nonword decoding score (max = 50)	0.000	0.020	
	(0.000)	(0.014)	0.020
Ateso Oral reading fluency (max = 68)	0.005	0.018	
	(0.004)	(0.008)	0.013
Ateso Reading comprehension score (max = 5)	0.000	0.003	
	(0.000)	(0.003)	0.003
Ateso Listening comprehension	1.841	1.405	
	(0.059)	(0.062)	-0.436***
English Letter sound score (max = 100)	0.408	0.253	
	(0.062)	(0.052)	-0.155
English Word segmenting score (max = 10)	4.577	4.770	
	(0.069)	(0.051)	0.193
English Nonword decoding score (max = 50)	0.003	0.000	
	(0.003)	(0.000)	-0.003
English Oral reading fluency (max = 68)	0.046	0.025	
	(0.024)	(0.011)	-0.021
English Reading comprehension score (max = 5)	0.000	0.000	
	(0.000)	(0.000)	0.000
English Receptive vocabulary score (max=20)	4.341	3.423	
	(0.133)	(0.117)	-0.918*

*p<0.10, **p<0.05, ***p<0.001