



USAID
FROM THE AMERICAN PEOPLE

RWANDA

SOMA UMENYE IMPACT EVALUATION

AID-696-TO-16-00002

EVALUATION - YEAR I BASELINE REPORT

April 26, 2018

This Year I Baseline Report was prepared for USAID/Rwanda by International Business & Technical Consultants, Inc. (IBTCI) under Task Order AID-696-TO-16-00002. The authors are Edward Jay Allan, Magda Raupp, Ali Protik, Kenn Ndirangu, and Nikita Ramchandani. The views expressed in this document do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

SOMA UMENYE IMPACT EVALUATION

AID-696-TO-16-00002

Emmanuel Gasana, COR

EVALUATION - YEAR I BASELINE REPORT

APRIL 26, 2018

Edward Jay Allan, IBTCI, Project Director

Magda Raupp, IBTCI, Team Leader

Ali Protik, Mathematica Policy Research, Methodologist

Kenn Ndirangu, Incisive Africa, Data Manager

Nikita Ramchandani, Mathematica Policy Research, Research Analyst

International Business & Technical Consultants, Inc. (IBTCI)

8618 Westwood Center Drive

Suite 400

Vienna, VA 22182

USA

DISCLAIMER

This Year I Report is made possible by the support of the American people through the United States Agency for International Development (USAID). The contents are the sole responsibility of IBTCI and do not necessarily reflect the views of USAID or the United States Government.

SOMA UMENYE IMPACT EVALUATION

Year I Baseline Report

TABLE OF CONTENTS

MAP OF RWANDA.....	v
ACRONYMS	vi
EXECUTIVE SUMMARY	i
Introduction and Background.....	1
Methodology and Impact Evaluation Design	2
Findings and Conclusions	3
1. Impact of Soma Umenye on student reading competency (EGRA timed sub-tests).....	3
2. Impact of Soma Umenye on student reading competency (EGRA untimed sub-tests) ..	4
3. EGRA findings for male and female students.....	5
4. EGRA findings for Repeaters and Non-Repeaters	5
5. EGRA findings for Students in Urban and Rural Sectors.....	5
Summary Conclusion.....	6
1. INTRODUCTION: EARLY GRADE READING CONTEXT IN RWANDA	7
2. OVERVIEW OF THE USAID/RWANDA SOMA UMENYE EARLY GRADE READING ACTIVITY	9
3. METHODOLOGY	10
3.1 Evaluation Questions	10
3.2 Conceptual Framework for the Impact Evaluation.....	10
3.3 Impact Evaluation RCT Design.....	11
3.4 Random Assignment of Sectors and Stratification.....	11
4. SAMPLING AND DATA COLLECTION	14
4.1 Sample Size Determination	14
4.2 Sampling Strategy.....	14
Random Selection of Schools.	14
Random Selection of Students.....	14
Selection of Teachers.	15
4.3 Data Collection Instruments and Administration	16
Brief Explanation of EGRA and of its Sub-tests.	16
Teacher Interview.	18
Head Teacher Interview.	18
Kinyarwanda Lesson Observation.	18
4.4 Enumerator Training and Data Quality Assurance.....	18

4.5	Data Overview and Data Cleaning Procedures	19
5.	ANALYSIS OF STUDENT AND TEACHER OUTCOMES.....	21
5.1	Student-level Analysis	21
	Outcomes.....	21
	Empirical Model	21
5.2	Teacher-level Analysis	22
	Outcomes.....	22
	Empirical Model.	23
6.	DESCRIPTION OF THE SAMPLE	24
6.1	Characteristics of the Students Interviewed and Assessed	24
6.2	Characteristics of the Teachers Included in the Sample	25
6.3	School and Classroom Characteristics.....	26
7.	FINDINGS	28
7.1	EGRA Findings.....	28
7.2	Lesson Observation Findings	32
7.3	Findings for Additional Student Outcomes	33
7.4	Sub-Populations of Interest	34
	What are the impacts of Soma Umenye on females and males?	34
	To what extent do the impacts of Soma Umenye differ across female and male students?	
		35
	Does Soma Umenye impact repeaters and non-repeaters differently?	36
	Does Soma Umenye impact students in Urban and Rural sectors differently?.....	38
	Does Soma Umenye impact students in different provinces differently?	40
8.	CONCLUSIONS	41
9.	RECOMMENDATIONS	44
	SELECTED REFERENCES	46

ANNEXES

Annex A. Proposed Revised Scope of Work	A1-A20
Annex B. Regulatory Certifications	B1-B5
Annex C. Evaluation Design and Year I Workplan	C1-C78
Annex D. Conflict of Interest Certifications	D1-D5
Annex E. Districts and Sectors by Phase	E1-E9
Annex F. Supervisor Manual	F1-F33
Annex G. The EGRA Instrument.....	G1-G18
Annex H. Student Context Form	H1-5
Annex I. Lesson Observation Form	I1-3
Annex J. Additional Results.....	J1-18
Annex K. Statement of Objections – Does not apply	

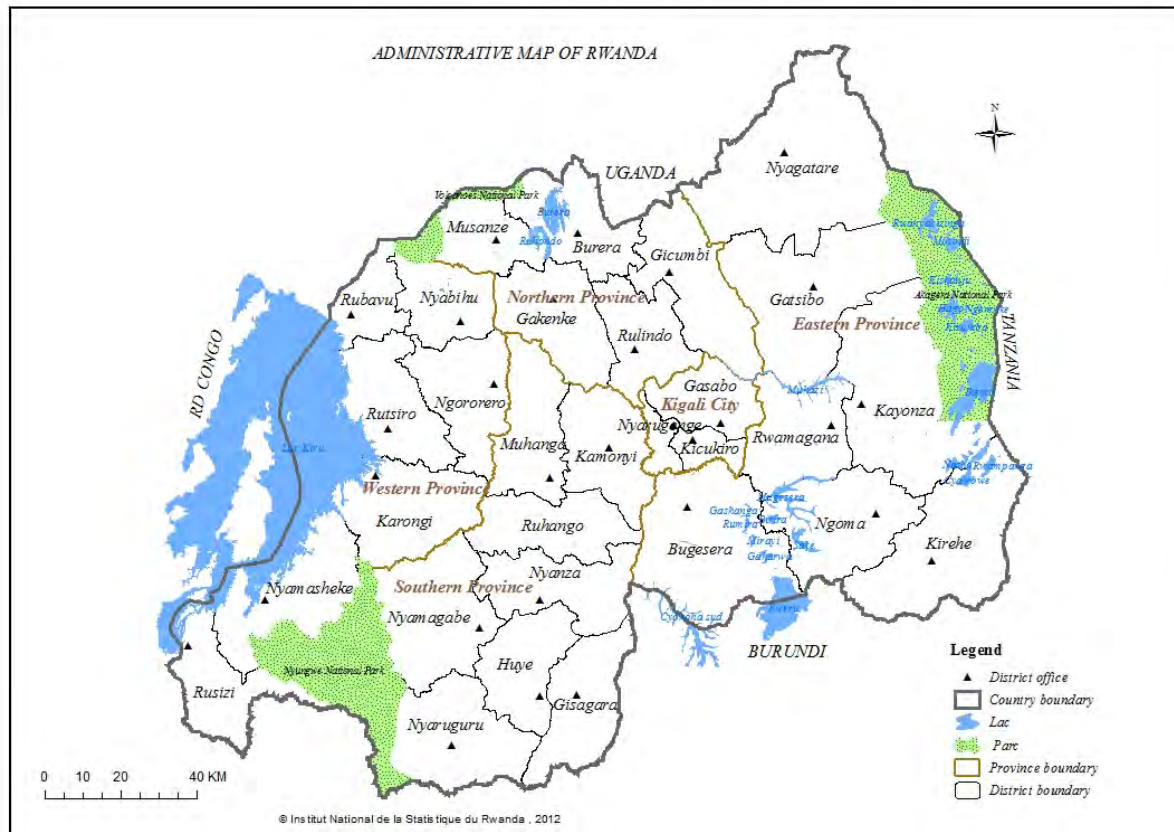
LIST OF TABLES

Table ES.1	Scores on the EGRA Timed Sub-tests by Treatment and Control Group	4
Table ES.2	Scores on the EGRA untimed sub-tests for Treatment and Control Group	4
Table 3.1	Pre-intervention Characteristics of Sectors and Schools	13
Table 4.1	Sample Size of Schools.....	14
Table 4.2	Data Collection Sample Sizes.....	16
Table 6.1	Age of Teachers Interviewed/Observed.....	25
Table 6.2	Teacher Years of Experience.....	26
Table 6.3	Average School Characteristics in the Sample.....	27
Table 7.1	Scores on the EGRA Timed Sub-tests by Treatment and Control Group	28
Table 7.2	Scores on the EGRA un-timed sub-tests for Treatment and Control Group.....	31
Table 7.3	Impacts of Soma Umenye on Classroom Instruction - Raw Scores	33
Table 7.4	Impacts of Soma Umenye on Classroom Instruction – Composite Score.....	33
Table 7.5	Impacts of Soma Umenye on Self-Reported Class Attendance	34
Table 7.6	Differential Impacts of Soma Umenye on Student Outcomes, by Gender.....	36
Table 7.7	Differential Impacts of Soma Umenye on Student Outcomes, by Grade Repetition Status	38
Table 7.8	Differential Impacts of Soma Umenye on Student Outcomes, by Urban-Rural Status	40

LIST OF FIGURES

Figure ES.1	The school-based package implemented in 2017 vis-à-vis the full package to be implemented in years 2 – 4	2
Figure 2.1	The school-based package implemented in 2017 vis-à-vis the full package to be implemented in years 2 – 4	9
Figure 6.1	Sampled Student Demographics (percentages of students except age)	24
Figure 6.2	Household Assets Reported by Students Who Took the EGRA	25
Figure 7.1	Score Distribution on Letter Identification Sub-Test	29
Figure 7.2	Score Distribution on Familiar Word Reading Sub-Test	30
Figure 7.3	Students Able to Read with Comprehension vs. Non-Readers	32
Figure 7.4	Scores on Timed EGRA Sub-Tests for Female and Male Students by Treatment Group	35
Figure 7.5	Scores on Untimed EGRA Sub-Tests for Repeaters and Non-Repeaters by Treatment Group	37
Figure 7.6	Scores on Timed Sub-tests for Students in Rural Sectors by Treatment Group ..	39

MAP OF RWANDA



ACRONYMS

COR	Contracting Officer's Representative
CLPM	Correct Letters per Minute
CPD	Continuing Professional Development
CPMD	Curriculum and Pedagogical Materials Development Department
CWPM	Correct Words per Minute
EDC	Education Development Center, Inc.
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
FBO	Faith-based organization
GoR	Government of Rwanda
IBTCI	International Business & Technical Consultants, Inc.
IE	Impact Evaluation
IP	Implementing partner
L3	EDC's Literacy, Language, and Learning project
LARS	Learning Achievement in Rwandan Schools
LOI	Language of Instruction
M&E	Monitoring and Evaluation
MDI	Minimum Detectable Impact
MINEDUC	Ministry of Education
NGO	Non-governmental Organization
NISR	National Institute of Statistics of Rwanda
ORF	Oral Reading Fluency
P1	Grade Primary 1
P2	Grade Primary 2
P3	Grade Primary 3
RCT	Randomized Control Trial
REB	Rwanda Education Board
RNEC	Rwanda National Ethics Committee
SD	Standard Deviation
SU	Soma Umenye
TBD	To Be Determined
TLM	Teaching and Learning Materials
TTC	Teacher Training College
URCE	University of Rwanda College of Education
USAID	United States Agency for International Development

EXECUTIVE SUMMARY

Introduction and Background

Soma Umenye (SU), “Read and Know,” is a five-year program implemented by Chemonics International to develop and improve Kinyarwanda reading skills of students in the first three years of primary school. SU is intended to improve quality of early literacy instruction and strengthen the capacity of the education system in Rwanda to implement and sustain these improvements. USAID awarded International Business & Technical Consultants, Inc. (IBTCI) a parallel five-year task order to conduct an Impact Evaluation (IE) of SU, including other studies. The IE planned to incorporate the 2017, 2018, and 2019 school years. Because of acceleration of implementation of SU to include all schools, including the control schools, in 2018, the IE task order was modified to cover only the 2017 school year. This report describes the planning and implementation of the Impact Evaluation starting in October 2016 and presents the impact of Soma Umenye at the end of the first school year (October 2017).

To assess the impact of Soma Umenye, the Early Grade Reading Assessment (EGRA) was adapted for Rwanda and Kinyarwanda, Rwanda’s major language. Field tests and revision took place in August 2017, and in September/October 2017 the EGRA was administered to a random sample of 5,457 PI students in 304 Treatment and Control schools throughout Rwanda. During these school visits, a student context questionnaire and teacher and head teacher surveys were administered. Chemonics International, the implementer, led the EGRA adaptation effort, the development of the additional data collection instruments, and the field-testing of all instruments. IBTCI, the Impact Evaluator, provided feedback to the process and to the instruments as needed. Data collection responsibilities for piloting and full collection were shared by the two organizations: Chemonics collected the data on 2,454 students and on teachers in randomly selected classes at the randomly selected treatment schools while IBTCI did the same on the 3,003 students and on teachers at the control schools. The two organizations shared the data sets to conduct the analyses. However, the focus of the analyses was different—IBTCI examined the impact of SU on student reading competency and developed a baseline for study of change in teacher instructional behavior. The focus of the implementer was primarily on the achievement of SU performance indicators.

It is important to note that during the 2017 school year only an abridged version of the full intervention package was implemented. SU provided eight days of training to PI Kinyarwanda teachers in treatment schools on the principles of effective reading instruction. Head Teachers, Directors of Studies, and District and Sector Education Officers received a two-day orientation on evidence-based reading instruction. Figure ESI (Figure 2.1 in Section 2) shows the activities implemented in 2017 vis-à-vis the full package of intervention to be implemented at the school level in years 2 – 4.

Figure ES1. The school-based package implemented in 2017 vis-à-vis the full package to be implemented in years 2 - 4

Implemented in Year I (2017)
• Eight days of face-to-face training on the principles of effective reading instruction; • Two-day orientation to evidence-based reading instruction for Head teachers, Directors of Studies, and District and Sector Education Officers; • Distribution of government-selected P1 Kinyarwanda textbooks to all schools. Completed in August 2017, to reduce the overall student-textbook ratio to 3:1. Note: Because all schools received these materials, this component of the package was not included in the Impact Evaluation.
Full package to be implemented in Years 2 – 4
• Distribution of essential core of Kinyarwanda teaching and learning materials. For P1, this includes the SU-prepared Kinyarwanda reading textbook (1 textbook per child), teacher guide (1 guide per teacher), read aloud story book (1 book per teacher), and a set of decodable readers (24 titles, one per student in class) and leveled readers (1 set of 76 titles per classroom). For P2, the essential core consists of the student textbook (1 textbook per child), teacher guide (1 guide per teacher), read aloud story book (1 book per teacher), and a set of leveled readers (1 set of 150 titles per classroom). The P3 essential core includes the student textbook (1 textbook per child), teacher guide (1 guide per teacher), read aloud story book (1 book per teacher), and a set of leveled readers (1 set of 150 titles per classroom); • Teacher continuing professional development (CPD)—ten days of face-to-face training focused on use of the materials, plus coaching/mentoring, self-learning videos, and termly community of practice meetings; • Ten days of face-to-face training for Kinyarwanda mentors (directors of studies, head teachers, or Kinyarwanda school subject leaders) to conduct structured classroom observations and mentoring; • Two days of face-to-face training for head teachers to enable them to support reading instruction and become instructional leaders in their schools; • Print-rich classrooms—alphabet charts, pocket boards and flash cards.

Methodology and Impact Evaluation Design

The Soma Umenye Impact Evaluation was designed as a randomized controlled trial (RCT), where a randomly assigned group that does not receive the intervention—the control group—is used as the counterfactual. The difference in outcomes between the beneficiaries of the intervention (the treatment group) and the control group is the measure of impact. By constructing a counterfactual, an impact evaluation is able to assign observed changes in outcomes to the intervention and clarify what the outcomes would have been in the absence of the intervention. The Methodology and the design of the Impact Evaluation are summarized in Section 3 and detailed in Annex C, Evaluation Design and Year I Workplan.

The EGRA was administered to individual students in Kinyarwanda and consisted of six sub-tests: (1) Letter Recognition; (2) Syllable Recognition; (3) Familiar Word Reading; (4) Oral Comprehension; (5) Fluency (reading connected text); and, (6) Reading Comprehension. Naming letters, enunciating the sounds of syllables, and reading isolated words are foundational skills needed for reading fluency. These timed sub-tests allow us to assess whether students are achieving the desired level of automaticity in these skill areas. Oral Comprehension and Reading Comprehension required students either to listen or to read a story and answer five questions to assess their understanding of what they heard or read. Students who were unable to perform a single item on a sub-test received a zero score on that sub-test.

Findings and Conclusions

This report addresses specifically Evaluation Questions 1 (student reading competency in Kinyarwanda) and 2 (the quality of literacy instruction in early grade reading classes). Our analysis provides initial information on these evaluation questions, but more work would be needed to assess the impact of Soma Umenye's full package on students' reading ability as well as on the other expected effects of the program. Evaluation Questions 3 and 4 were scheduled to be addressed in subsequent reports using data collected in the 2018 and 2019 school year.¹ The key findings are summarized below and detailed in Section 7.

- **Evaluation Question # 1: To what extent are changes in Kinyarwanda reading outcomes for students in Grades 1-3 attributable to the Soma Umenye activity (as a package)?**

1. Impact of Soma Umenye on student reading competency (EGRA timed sub-tests)

The average scores obtained by PI students in treatment schools were compared to students' average scores in schools that received no intervention (control schools). Table ES.1 highlights where the differences between the two groups are significant. Note that the number of items per sub-test differs.

Our analyses show very low overall performance on the four timed sub-tests. On the average, students identified only 20 percent of the 100 (19 to 23 letters); they recognized a little over 10 percent of the 100 syllables; and were able to read fewer than 5 of the 50 familiar words presented to them. In two of the four sub-tests, the differences noted between students in treatment and in control schools were statistically significant, showing the impact of Soma Umenye. Students in the treatment group had an average score of 4.51 on the reading of familiar words sub-test while students' in the control groups had an average score of 3.72 words per minute. Even though the difference is statistically significant on the two sub-tests, our findings show that overall, PI students in both groups are not acquiring the lower-order reading skills necessary to become fluent readers in later grades.

¹ In February 2018 USAID communicated to IBTCL that as the result of the accelerated expansion of Soma Umenye, the Impact Evaluation would be discontinued.

Table ES.1 Scores on the EGRA Timed Sub-Tests by Treatment and Control Group

EGRA timed sub-test scores	Treatment	Control	Difference between groups
Letter Identification (MAX 100 clpm)	22.19	19.41	2.77*
Syllable identification (MAX 100 csilpm)	12.11	11.19	0.93
Familiar Word Reading (MAX 50 cwpm)	4.51	3.72	0.79*
Reading Fluency (MAX 38 orf)	4.98	4.27	0.72

N=5,466

* Differences between treatment and control statistically significant at .01/.05 level. clpm: correct letters per minute; csilpm: correct syllable per minute; cwpm: correct words per minute; orf: oral reading fluency (number of words of connected text read correctly in one minute)

Nevertheless, students in the treatment schools achieved gains in reading performance that were not achieved in control schools. Considering the difference between the intended package of intervention planned to be implemented in years 2 – 4 and the abridged package that was possible to implement in 2017, we hypothesize that in 2018, and even more so in 2019, the effects of the project could be larger and that the gap between students in the treatment and in control schools could widen.²

2. Impact of Soma Umenye on student reading competency (EGRA untimed sub-tests)

Table ES.2 shows the average scores for students in the treatment and control groups, the differences observed between the two groups for the sub-tests Listening Comprehension and Reading Comprehension.

Table ES.2 Scores on the EGRA Untimed Sub-Tests for Treatment and Control Group

EGRA untimed sub-test scores	Treatment	Control	Difference between groups
Listening Comprehension (MAX 5 questions)	3.91	3.78	0.13*
Reading Comprehension (MAX 5 questions)	0.69	0.60	0.09

N=5,457

* Differences between treatment and control statistically significant at .05 level.

Both groups scored high on the **Listening Comprehension** sub-test—3.84 on average out of a possible score of 5, showing students' ability to listen to a story, understand what was read to them, and correctly answer oral comprehension questions. Overall students scored 78 percent, close to the proficiency level of 80 percent, answering four questions out of five correctly. It is also possible that the sub-test was too easy for most students and this caused the scores to concentrate heavily at the upper end of the allowable range, restricting the variation in scores (ceiling effect).

² Since the IE was interrupted, the effective impact of SU may not be knowable.

Average scores on the **Reading Comprehension** sub-test were extremely low for both groups, 0.64 on average out of a possible 5.0. Overall, 3.09 percent students (N=169) scored 80 percent or higher and 0.26 percent (N=14) scored 100 percent. **These 183 students total are considered able to read grade-level text with comprehension.** In the treatment group, 3.18 percent of the students (N=78) scored 80 percent or higher and in the control group, 3.02 percent (N=91) scored 80 percent or higher. The difference between treatment and control is not statistically significant. ORF scores for proficient readers—students who scored 80 percent or higher in reading comprehension—was 27.8 correct words per minute. Because of the low performance in words per minute (Table ES.1), it is not surprising that scores concentrated heavily at the lower end of the score distribution. EGRA studies have shown the high correlation (0.91) between correct words read per minute (fluency) and reading comprehension.³ We also observed this high correlation between oral reading fluency and reading comprehension in our sample, which was 0.92. Other findings that resulted from the analyses of the EGRA data are summarized below. The findings are detailed in Section 7.

3. EGRA findings for Male and Female Students

Soma Umenye had statistically significant positive impacts on female students in the treatment group compared to female students in the control group on three sub-tests: **letter identification, syllable identification, and familiar word reading.** The intervention also had statistically significant positive impacts on male students in the treatment group compared to male students in the control group on the **familiar word reading** sub-test. Given that the impacts on male students were not different from the impacts on female students, we conclude that, on average, SU benefitted female and male students equally.

4. EGRA findings for Repeaters and Non-Repeaters

Soma Umenye had statistically significant positive impacts on both groups of students. Repeaters and non-repeaters in the treatment group scored better in the **letter identification** sub-test compared to repeaters and non-repeaters in the control group. These findings suggest that while repeaters may have derived some benefits from their previous academic experience, even an intensive reading intervention such as Soma Umenye did not benefit them any more than it did the non-repeaters.

5. EGRA findings for Students in Urban and Rural Sectors

Our sample was overwhelmingly rural, 90.4 percent or 4,932 students were from 275 schools in rural sectors and only 525 (9.6 percent) were from 29 schools in urban sectors.⁴ Because rural students represented such a high proportion of the total sample, the impacts of SU on them closely resembles the impacts on the total sample and we are not able to conclude that SU benefitted rural students more or less than their urban counterparts.

³ Fuchs, L., Fuchs, D., Hosp, M.K. & Jenkins, J. (2001) Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis.

⁴ We used the official definition of rural-urban status of sectors as provided by the Rwanda Education Board.

- **Evaluation Question # 2: To what extent has classroom instruction changed as a result of Soma Umenye?**

Soma Umenye and IBTCI enumerators conducted a total of 281 observations of the Kinyarwanda reading lesson at the class randomly selected in each school for EGRA participation, using a lesson observation protocol developed by SU. Teachers were scored on five skills: (i) lesson preparation, (ii) teaching methodology, (iii) literacy technical skills, (iv) assessment techniques, and (v) use of teaching materials. Teachers in treatment schools, who participated in SU training received an average composite score that was half a standard deviation better compared to the score obtained by control group teachers and the difference is statistically significant. Because Evaluation Questions 3 and 4 call for data that would have required that the IE be continued, it is not possible to answer them in this Year I report.

Summary Conclusion

The main focus of the impact evaluation is the comparison between the EGRA scores obtained by students in intervention schools and their counterparts in control schools. These comparisons show an initial impact of Soma Umenye on two sub-tests: Letter Identification and Familiar Word Reading. Given the limited intervention implemented in 2017, these findings are encouraging. However, the great majority of students did not demonstrate that they were acquiring the lower-order reading skills that we can expect at P1—they identified only 20 percent of the 100 letters presented to them, recognized a little over 10 percent of the 100 syllables, and were able to read fewer than 5 of the 50 familiar words. Higher-order reading skills (fluency and comprehension) build on lower order skills (phonemic awareness, letter sound knowledge and, decoding). EGRA tests were developed to be predictive of later reading achievements, and numerous administrations of EGRA in multiple countries and languages have confirmed the correlations between lower- and higher-order reading skills.⁵ Overall, findings allow us to conclude the great majority of students are not learning the foundational skills that will allow them to become fluent readers by the end of P3. Improving the teaching—and consequently the learning—of foundational reading skills in P1 and P2, using EGRA findings to establish standards and benchmarks, and training reading teachers and headmasters on how to enhance students' reading outcomes should therefore be an urgent priority.

⁵ Early Grade Reading Assessment (EGRA) Toolkit. Page 18, 2nd Edition, March 2016.

I. INTRODUCTION: EARLY GRADE READING CONTEXT IN RWANDA

Soma Umenye (SU), “Read and Know,” is a five-year program implemented by Chemonics International to develop and improve Kinyarwanda reading skills of students in the first three years of primary school. SU is intended to improve quality of early literacy instruction and strengthen the capacity of the education system in Rwanda to implement and sustain these improvements. USAID awarded International Business & Technical Consultants, Inc. (IBTCI) a parallel five-year task order to conduct an Impact Evaluation (IE) of SU, including other studies. The IE planned to incorporate the 2017, 2018, and 2019 school years. Because of acceleration of implementation of SU to include all schools, including the control schools, in 2018, the IE task order was modified to cover only the 2017 school year. This report describes the planning and implementation of the Impact Evaluation starting in October 2016 and presents the impact of Soma Umenye at the end of the first school year (October 2017).

Over the past two decades, Rwanda has made progress towards its goal of becoming a middle-income, knowledge-based economy. Rwanda Vision 2020 emphasizes strengthening human capital formation, particularly through education and emphasizes formation of human capital through education service delivery of two key components: access and quality. In terms of access, Rwanda has achieved near universal access to primary education. According to the 2013/4 Global Monitoring Report, Rwanda ranked as one of the top three performers in the last five years by reducing the out-of-school population by at least 85 percent. While progress in terms of providing access to basic education services is an achievement, the 2013/4 Global Monitoring Report also noted the need for increased investments in the basic education sector in order to realize comparable improvements in the provision of quality basic education service delivery. The students in the three initial grades represent a wide range of reading ability levels and their ages may range from 5- to 15-years-old or older. The P1 repetition rate is 26 percent, and the enrollment in P3 is approximately 45 percent of that in P1. Additionally, absenteeism affects the already limited opportunities that students have to learn—a 2012 Early Grade Reading Assessment (EGRA) noted that 20 percent of students reported having been absent the previous week.

National assessments in Rwanda have documented the generally poor reading skills of primary school students. According to a national Early Grade Reading Assessment (EGRA) conducted in 2012, 13 percent of students in P4 could not read a single word of a P2/3 level text in Kinyarwanda. Another 13 percent were reading fewer than 15 correct words per minute, far below what is necessary for comprehension.⁶ Similarly, a 2014 reading assessment conducted by the USAID-funded *Literacy, Language, and Learning* (L3) activity found that 60 percent of P1 students, 33 percent of P2 students, and 21 percent of P3 students were unable to read a single word of grade-level text.⁷

Comprehension, or the ability of students to understand what they have read, is the ultimate goal of basic literacy instruction. Research and results from EGRA assessments worldwide have

⁶ De Stefano, J. et al. (2012) Early Grade Reading and Mathematics in Rwanda: Final Report. Research Triangle Park, ND: RTI International page 3.

⁷ Education Development Center. (November 2014). USAID L3: National Fluency and Mathematics Assessment Baseline Report

consistently demonstrated a causal linkage between oral reading fluency and reading comprehension. Oral reading fluency, defined as the ability to read with sufficient accuracy, speed, and expression, is prerequisite to comprehension. If students are reading too slowly, they forget what they have read and therefore cannot understand what they read.

Low reading comprehension levels follow the extremely low reading fluency levels in P1-P3. In the 2012 EGRA, almost 40 percent of P4 students were unable to correctly answer half of the comprehension questions posed (the generally accepted standard on the reading comprehension sub-test of EGRA is 80 percent). Twenty-seven percent could not answer a single question correctly.⁸ In L3's 2014 assessment, 93 percent of P3 students were unable to correctly answer 4 out of 5 (or 80 percent of) comprehension questions.⁹

Early-grade reading is the foundation for all other learning and is a major determinant of whether a child stays in school. In their foundational work, “Early Grade Reading: Igniting Education for All,” Gove and Cvelich argue that *“the point of reading is comprehension, and the point of comprehension is learning. Children who fail to learn to read in the early grades of school are handicapped because they must absorb increasing amounts of instructional content in print form.”*¹⁰ Gove and Cvelich further observe that *“a nation’s economic prospects follow the learning curve of its children,”* and that *“a 10 percent increase in the share of students reaching basic literacy translates into a 0.3 percentage point higher annual growth rate of that country.”* Improving early grade reading skills is essential to develop the human capital that Rwanda needs to meet its development goals and compete in the regional and global knowledge economy.

⁸ DeStefano, J. et al. (2012) Early Grade Reading and Mathematics in Rwanda: Final report. Research Triangle Park, ND: RTI International, page 5.

⁹ Education Development Center. (2014 November). USAID/L3. National Fluency and Mathematics Assessment Baseline Report.

¹⁰ Gove, A. and P. Cvelich. 2011. Early Reading: Igniting Education for All. A report by the Early Grade Learning Community of Practice. Revised Edition. Research Triangle Park, NC: Research Triangle Institute

2. OVERVIEW OF THE USAID/RWANDA SOMA UMENYE EARLY GRADE READING ACTIVITY

The Soma Umenye activity was to be implemented in three phases, starting with grade P1 in 2017 and adding one grade with each subsequent year—approximately $\frac{1}{3}$ of the schools each year starting in 2017. Each implementation phase would have been one year long and aligned with Rwanda's school year, which begins in January/February and continues until October/November. The phased implementation would allow time for the IP team to prepare materials and to adjust activities as needed. It would also make the Impact Evaluation possible.¹¹ In this report, the IE estimates the average impact of the Soma Umenye package “as implemented” or the average impact of SU in terms of the difference between the treatment group that received the intervention package as it was implemented in 2017 and the control group that did not receive the intervention. Our findings relate only to the intervention package implemented in Year 1 (2017).

Figure 2.1 compares the activities implemented in Year 1 (2017) to the complete package of intervention to be implemented in years 2 - 4.¹²

Figure 2.1 The school-based package implemented in 2017 vis-à-vis the full package to be implemented in years 2 - 4

Implemented in Year 1 (2017)
• Eight days of face-to-face training on the principles of effective reading instruction; • Two-day orientation to evidence-based reading instruction for Head teachers, Directors of Studies, and District and Sector Education Officers; • Distribution of government-selected P1 Kinyarwanda textbooks to all schools. Completed in August 2017, to reduce the overall student-textbook ratio to 3:1. Note: Because all schools received these materials, this component of the package was not included in the Impact Evaluation.
Full package to be implemented in Years 2 – 4
• Distribution of essential core of Kinyarwanda teaching and learning materials. For P1, this includes the SU-prepared Kinyarwanda reading textbook (1 textbook per child), teacher guide (1 guide per teacher), read aloud story book (1 book per teacher), and a set of decodable readers (24 titles, one per student in class) and leveled readers (1 set of 76 titles per classroom). For P2, the essential core consists of the student textbook (1 textbook per child), teacher guide (1 guide per teacher), read aloud story book (1 book per teacher), and a set of leveled readers (1 set of 150 titles per classroom). The P3 essential core includes the student textbook (1 textbook per child), teacher guide (1 guide per teacher), read aloud story book (1 book per teacher), and a set of leveled readers (1 set of 150 titles per classroom); • Teacher continuing professional development (CPD)—ten days of face-to-face training focused on use of the materials, plus coaching/mentoring, self-learning videos, and termly community of practice meetings; • Ten days of face-to-face training for Kinyarwanda mentors (directors of studies, head teachers, or Kinyarwanda school subject leaders) to conduct structured classroom observations and mentoring; • Two days of face-to-face training for head teachers to enable them to support reading instruction and become instructional leaders in their schools; • Print-rich classrooms—alphabet charts, pocket boards and flash cards.

¹¹ As the implementation schedule changed and all schools started receiving the intervention in 2018, the IE was left without its counterfactual, making the assessment of Soma Umenye's impacts impossible.

¹² Soma Umenye Annual Report Year 1: July 20, 2016 to September 30, 2017.

3. METHODOLOGY

In this section we state and discuss the original research questions that guided the impact evaluation and provided the conceptual framework on which we based the evaluation design and summarize the methodology for the impact evaluation conducted in 2017. Details of the methodology as designed prior to the modification of the IE scope (January 2018) are included in Annex C, the Evaluation Design and Year 1 Work Plan.

3.1 Evaluation Questions

The set of four evaluation questions identified by USAID Rwanda provided the guidance for the Impact Evaluation design and sampling. Because the expansion of Soma Umenye to all Rwandan schools for the 2018 school year brought the Impact Evaluation to an end after its first school year, we could only provide some initial answers to Questions 1 and 2. Question 1 refers to grades P1 – P3, but the IE findings reported in this document only refer to grade P1. Questions 3 and 4 would have been answered in subsequent years had the IE continued to its term. In addition, questions that relate to specific small sub-populations of interest to USAID—for example, the apparent impact of the school feeding program on student learning, studies on students with disabilities—that could not be part of the IE would have been the object of special studies in the later years of the task order. The four evaluation questions are presented below.

1. To what extent are changes in Kinyarwanda reading outcomes for students in Grades 1-3 attributable to the Soma Umenye activity (as a package)?
2. To what extent has classroom instruction changed as a result of Soma Umenye?
3. What is the evidence that improved student reading outcomes in Kinyarwanda lead to improved opportunities for students to succeed in schooling?
4. To what extent are the improvements in the systemic capacity for early grade reading instruction a result of the Soma Umenye activity?

Question 1 required that we compare student reading scores on the EGRA sub-tests with and without the Soma Umenye intervention. Question 2 required that we compare early reading instructional practices exhibited by teachers who have received the limited training implemented in 2017 by Soma Umenye training to the practices of teachers who did not receive any Soma Umenye training. Please note that the scope of the Impact Evaluation was significantly reduced with the acceleration of the implementation of Soma Umenye. Questions 3 and 4 would be answered through the collection and analyses of data in 2018, 2019, and 2020. Other than self-reported student data on absenteeism, there were no 2017 data to address Questions 3 and 4.

3.2 Conceptual Framework for the Impact Evaluation

An Impact evaluation (IE) is an assessment of how the intervention being evaluated affects outcomes. The proper analysis of impact requires a counterfactual of what outcomes would have been in the absence of the intervention. The Soma Umenye evaluation was designed as a randomized controlled trial (RCT), where a randomly assigned group that does not receive the intervention—the control group—is used as the counterfactual. By constructing a counterfactual—a control group—an IE is able to assign observed changes in outputs and

outcomes to the intervention. The difference in outcomes between the beneficiaries of the intervention (the treatment group) and the control group is the measure of impact.

3.3 Impact Evaluation RCT Design

Considering that the Soma Umenye IE was designed ex-ante, randomization was possible. Soma Umenye's initial plan was to implement the activity in groups of schools over three years, with roughly 30 percent starting the activity in 2017 (Phase 1), another roughly 30 percent in 2018 (Phase 2), and the remaining 40 percent in 2019 (Phase 3). We randomly assigned sectors to these three phases before implementation began in 2017. Because schools in sectors selected for implementation in Phase 3 will not have received the package of interventions for two years, they were chosen to represent the counterfactual condition—what would have happened without the intervention. Schools in Phase 2 sectors were not included in the IE sample to avoid any risk of contamination between Phase 1 and Phase 2 schools. With random assignment of sectors into phases, schools in sectors selected for Phase 1, and Phase 3 are likely to be the same, on average, in terms of their characteristics. Thus, comparing outcomes at schools in Phase 1 sectors to those of schools in Phase 3 sectors will allow us to measure the impacts of the Soma Umenye package of interventions on student reading ability and on other variables suggested by the IE evaluation questions.

Note that a decision was made to roll out Soma Umenye to all schools and students in Rwanda starting in 2018, after the data for this report was collected. As a result, SU will be implemented in schools in all three phases starting in the 2018 school year. As a result, the IE for future years will not be possible. Below, we briefly summarize the RCT design that forms the basis of the findings presented in this report. The details of the original longitudinal RCT evaluation design are provided in Annex C.

3.4 Random Assignment of Sectors and Stratification

The IE team guided the IP in the random assignment of sectors into phases conducted in mid-December 2016 so that the implementation of Soma Umenye could be planned accordingly, before the beginning of the intervention in January of the 2017 school year. The implementation schedule needed for the IE was fully congruent with the IP's own schedule, at the time, for phasing in of its own implementation. It was important that random assignment of all three phases took place prior to the implementation of the project so that it became clear from the start which sectors would receive the intervention and which sectors would be part of the control group.

The random assignment was carried out according to the following stratification:

- **Stratification by district.** We stratified the random assignment of sectors by district for two reasons. First, there are geographic variations in education outcomes in Rwanda, which could confound the effectiveness of Soma Umenye if it is implemented in one region of the country and not in the others. By assigning sectors within each district, we ensure that the intervention and the control group sectors are evenly spread across the country. Second, because all districts, and thus all provinces, are represented in both the intervention and

the control groups, we will be able to explore any variation in impacts by province.¹³ Third, this strategy is helpful for the IP as it will not have to exclude any district from their implementation plan in any of the phases.

- **Stratification by sector size.** Within each district, we further stratified the random assignment by the size of the sectors in terms of the number of schools. The IP recommended that we follow this stratification process because sectors with a large number of schools are likely to operate differently than sectors with a small number of schools. For example, larger sectors may be better able to attract better quality teachers or they may have more resources to monitor student progress. We stratified each sector into two pots – one containing the larger sectors and one the smaller sectors. The randomization of sectors into phases was carried out in the following manner:

The random assignment was carried out publicly in each district per guidance provided by the IE to the IP. Because implementation of Soma Umenye was going to be accomplished in phases, transparency was felt to be particularly important to avoid potential perceptions that some schools might be the beneficiaries of favoritism by having their schools to be among the first for implementation. To assure both fairness and the appearance of fairness, representatives of a district's sectors were convened to a meeting in December 2016 at which they pulled sector names from a container and assigned randomly to an implementation phase.

In preparation for this public random assignment, we compiled a list of sectors by district and, within each district, grouped them into two pots according to their sizes in terms of the number of public and government-aided schools. The first pot in each district contained sectors that have more public and government-aided schools than the average number of these schools in the district. The second pot contained sectors that have less than or equal to the average number of these schools in the district. In addition, we calculated the number of sectors in each pot that should be assigned to Phase 1, Phase 2, and Phase 3 so that the final distribution would be roughly 30 percent, 30 percent, and 40 percent across the three phases. In case where the percent distribution resulted in a fraction, we always rounded up in favor of Phase 1 and rounded down for Phase 2. This resulted in a slightly higher percentage of Phase 1 sectors (33 percent) and lower percentage of Phase 2 sectors (27 percent) than was originally proposed. The IP used the fraction of sectors for each district and publicly conducted lotteries in the presence of sector officials to assign sectors to Phases 1, 2, and 3 for each pot. The random assignment by pot ensured a more even distribution of large and small sectors, in terms of number of public and government-aided schools, into the three phases. The list of sectors by districts selected for the roll out implementation phases is included in Annex E. Once a sector was assigned to a phase, all schools in that sector are to receive the intervention in that phase.

We examined the balance between the treatment and control groups using pre-intervention administrative data. Ideally, we would have collected pre-intervention baseline student-level data, immediately following the random assignment in December 2016. However, we were not able to administer a baseline student-level survey before the intervention was implemented or immediately following the start of the intervention in January 2017, when P1 students began the

¹³ Because there are five provinces in Rwanda (East, Kigali City, North, West, South), the sample size for each province will be about one-fifth of the entire sample. This may not allow us enough statistical power to examine impacts separately for each province. However, we will still be able to test whether the impacts of Soma Umenye varied across provinces.

new school year. Nevertheless, the pre-intervention administrative data indicates that there were no statistically significant differences between the sectors or schools in treatment and control groups on most of the indicators available to us (Table 3.1). There was only one significant difference between the two groups, in terms of the percentage of schools having a teacher training curriculum, which was lower among the treatment schools. It is possible that this difference has appeared simply by chance. Nonetheless, we account for this difference in our analysis.

Table 3.1. Pre-intervention Characteristics of Sectors and Schools

Characteristic	Treatment	Control	Difference
<i>Panel A: Sector-level characteristics</i>			
Number of schools in sector	6.08	5.98	0.10
Has HGSF program (%)	3.02	4.82	-1.81
Rural (%)	91.28	90.01	1.28
<i>Panel B: School-level characteristics</i>			
Public school (%)	36.95	34.05	2.91
Government-aided school (%)	63.05	64.97	-1.92
Number of PI classrooms	5.10	5.03	0.08
Has library (%)	35.14	36.99	-1.85
Has nursery (%)	76.40	76.47	-0.07
Has teacher training curriculum (%)	1.64	5.26	-3.62*
Number of Kinyarwanda teachers	2.35	2.38	-0.03
Number of PI students	261.60	241.13	20.46
Number of male PI students	130.28	120.68	9.61
Number of female PI students	131.31	120.46	10.86
Number of students with disabilities	1.58	1.20	.38
<i>School Sample size</i>	<i>136</i>	<i>168</i>	<i>304</i>

N = 304

Sources: Administrative data from Rwanda Education Board, 2016.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. Sample sizes for some characteristics may be smaller due to missing data.

* Difference between treatment and control group means is statistically significant at the .05 level.

4. SAMPLING AND DATA COLLECTION

In this section we describe the data collection procedures used, provide an overview of the data, and describe the methods of analyses conducted.

4.1 Sample Size Determination

For this impact evaluation, the major consideration for sample size determination at the design stage was the statistical power. Ideally, impact evaluations are powered to detect, measure and validate smaller, rather than larger minimum detectable impacts (MDIs)¹⁴, so that at least small yet meaningful impacts—which are common in educational interventions—can be documented and lessons regarding what works can be learned. Overall dropout rates can be as high as 15 percent, depending on the year reported, as estimated by the 2015 Rwanda Education Statistical Yearbook. Therefore, we took a relatively large sample size of 18 students per school. We also decided to sample one school in every sector in our study.

4.2 Sampling Strategy

Random Selection of Schools.

All sectors that were part of original Phase 1 (sectors where Soma Umenye was implemented during 2017) and Phase 3 (sectors where Soma Umenye was to be implemented during 2019) were automatically part of our study. As the first step for selecting the sample for analysis, we randomly selected one school within each of the treatment and control sectors to be visited to maximize statistical power. Because the random assignment is conducted at the sector level, statistical power for our IE was extremely sensitive to the number of sectors. This is a common feature in cluster random assignment studies where a relatively large sample size of the unit of assignment (sector in this case) is required to gain a targeted level of precision compared to non-cluster random assignment studies (e.g. at the student level) where the unit of assignment and analysis are the same (Schochet, 2008). By selecting one school per sector, we ensured that each intervention and control sector was represented in the sample. The resulting sample size of intervention and control schools are presented in Table 4.1.

Table 4.1. Sample Size of Schools

	Sample size
Number of intervention (Phase 1) schools	136
Number of control (Phase 3) schools	168
Total number of schools for the IE	304

Random Selection of Students.

In many Rwandan schools, PI is taught in two shifts (morning and afternoon) and the average enrollment of PI students in public and government-aided schools is over 500 per school. Selecting students and collecting data across both shifts would have posed a significant challenge

¹⁴ The MDIs show the smallest effects that the evaluation can reliably detect, measured in the units of the outcome variable. We also report MDIs in terms of the percent change in standard deviation units.

both to the IE, the IP, and the participating schools themselves as this would have required asking morning shift students to come in the afternoon or afternoon shift students to come in the morning for the assessment. This would cause great disturbance to the school day and to parents. Therefore, the decision was made to involve only morning shift students. As delineated in Annex F, the Supervisor Manual, the steps below were followed:

1. First, we randomly selected one classroom from the total number of P1 classes in the morning shift.
2. At each school in our sample, we randomly selected nine boys and nine girls from the randomly selected classroom. This was carried out according to a pre-set protocol in the Supervisor Manual detailing the rules of selection that the data collection teams adhered to. The data collection team executed the random selection in the presence of the headmaster and/or a representative from the school staff and/or a representative from the community.
3. For the initial assessment, we obtained assent from each student in accordance with RNEC and NISR procedures. If a student did not give consent, the student was replaced by another randomly-selected student. No individually identifiable information was used in preparation or dissemination of findings from this IE.

Selection of Teachers.

In each intervention and control school, the Kinyarwanda teacher of the randomly selected class was automatically selected to be interviewed and observed as he or she conducted the reading lesson.¹⁵ This resulted in an intended teacher sample of 304—136 in the treatment group and 168 in the control group. Out of the 304 classrooms selected, only 287 Kinyarwanda teachers were interviewed due to teacher absenteeism on the day of data collection. Lesson observation was carried out for 281 of the 287 teachers present. There were six cases where the teacher was interviewed but the lesson did not take place during the time of the survey and the lesson could not be observed. Nonetheless, the response rates were quite high—97 percent in the treatment and 92 percent in the control group. Only 268 head teachers/director of studies were interviewed as the enumerators found that many head teachers were absent when visiting the schools. A common reason given for their absence was sector or other meetings outside of the school premise.

Table 4.2 below shows the actual data collection sample sizes for schools, students, and teachers for classroom observations. All 304 schools in the sample were visited by either the IBTCI or the Soma Umenye data collection teams. The overall sample size has six fewer students than the planned sample size of 5,472 that should result from randomly sampling 18 students from each of the 304 schools. This is because in some cases the class sampled had fewer than nine males and/or females, not allowing us to sample 18 students. Also, in some cases, our sample consists of more than nine male and/or female students per school. This was done due to enumerator error when sampling by gender. The evaluation team received full explanations from the data collection teams on these issues.

¹⁵ Teachers frequently teach multiple grades in Rwanda. So, a teacher assigned to P1 may also be assigned to teach P2 and/or P3 classrooms.

Table 4.2. Data Collection Sample Sizes

Characteristic	Treatment	Control	Total
Schools	136	168	304
Students	2,455	3,011	5,466
Males	1,229	1,523	2,752
Females	1,226	1,488	2,714
Head Teachers	122	146	268
Teachers	132	155	287
Reading Lesson Observations	129	152	281

4.3 Data Collection Instruments and Administration

Based on the data requirements for the IE and for Soma Umenye, five instruments were adapted or developed by the IP subcontractor EdIntersect. EdIntersect and the IP conducted a series of workshops with REB to ensure that the adaptation of instruments matched REB's guidelines and expectations. The Chemonics and IBTCI teams worked closely with EdIntersect and provided the necessary inputs to improve the quality of the instruments. In July/August 2017 REB reviewed and validated all the instruments.

- 1. Early Grade Reading Assessment (EGRA)**—The EGRA instrument used by the IP in October 2016 was revised and adapted by EdIntersect with input from REB and from the IP the IE teams. The EGRA was administered to the 5,466 PI students included in the sample.
- 2. Student Context Questionnaire**—Prior to the EGRA assessment, each of the 18 randomly selected participating students responded to a brief, orally administered questionnaire. The purpose of the questionnaire was to gather information about the home and school contexts that might explain students' reading performance.
- 3. Teacher Interview Protocol**—287 teachers whose classes were randomly selected for assessment were interviewed. The Interview protocol focused on teacher variables such as sex, age, length of time working as a teacher, pre-service and in-service or previous training received.
- 4. Lesson Observation Instrument**—The purpose of the instrument was to capture instances of teacher instructional behavior (when teaching reading) that could be seen as caused by the intervention. The observation was conducted in 281 classes.
- 5. Head Teacher Interview Protocol**—The interview was conducted with 268 head teachers of the schools randomly selected for assessment. A semi-structured interview protocol was developed to gather the perception of school principals and to explore the human, technical, and financial resources existing at the schools as well as the factors that facilitate or limit the implementation of Soma Umenye.

Brief Explanation of EGRA and of its Sub-tests.

The ability to read and understand a simple text is one of the most fundamental skills a child can learn. Yet, measuring early reading can be challenging since most tests are administered in higher

grades and they are not likely to capture the fundamental or emerging skills that students need to become fluent readers.¹⁶ Early assessment of the pre-reading and foundational skills required for fluency allows the implementation of measures to correct deficiencies where they exist.¹⁷

The EGRA tool offers an opportunity to determine whether students in the early grades are developing the fundamental reading skills, and, if not, where efforts might be best directed. The EGRA has been adapted and used in over 50 countries and can capture more subtle impacts from specific teaching approaches than pencil-and-paper tests, as it incorporates sub-tests that measure pre-reading skills.¹⁸ The assessments are administered orally and individually. Administering the EGRA Instrument to PI students in Rwanda took about 45 minutes per child and the students' responses were recorded on tablets. Data collected at the treatment and control schools were transmitted to Chemonics, the Soma Umenye implementer, on a daily basis and uploaded into the Tangerine database maintained by RTI.

The adaptation of the EGRA used in Rwanda in September/October 2017 assessed students' competency in five sub-tests:

1. Letter recognition assessed ability to provide the names of the letters of the alphabet naturally and without hesitation. This is a timed test that assesses automaticity and fluency of letter recognition and is measured in number of letter names correct per minute. Students were shown a chart containing 10 rows of 10 random letters (in uppercase and lowercase) and asked to name as many letters as they could within one minute yielding a score of correct letters read per minute (clpm). Letters were presented to the students in either block or cursive formats and lower- and upper-case format (each type on one side of a large plasticized card) as familiarity with the two formats was found to vary during field-testing of the instrument.

2. Syllable recognition assessed ability to read high-frequency syllables. Recognizing and reading high frequency syllables is critical for a speaker of Kinyarwanda since, as in many other countries, early grade reading in Rwanda is taught by combining syllables to form words. In this timed sub-test, students were asked to sound out as many syllables (in a list of 100) as they could within one minute, yielding a score of correct syllables per minute (csilpm).

3. Familiar word reading assessed students' skills at reading high-frequency words. Recognizing familiar words is critical for developing reading fluency. In this timed sub-test, students were asked to sound out as many words (in a list of 50) as they could within one minute, yielding a score of correct words per minute (cwpm).¹⁹

¹⁶ Emergent reading skills are "skills, knowledge, and attitudes that are developmental precursors to conventional forms of reading and writing. These skills are the basic building blocks for how students learn to read and write." (Connor et al., 2006, p. 665).

¹⁷ Abadzi, Helen. (2009). "Instructional Time Loss in Developing Countries: Concepts, Measurement, and Implications." World Bank Research Observer. 24 (2): 267-290.

¹⁸ The Early Grade Reading Assessment (EGRA) was originally developed by the Research Triangle Institute (RTI) administered orally to students in the early grades of primary school. As pointed out by RTI, the EGRA evaluates students' foundational reading skills, including pre-reading skills like phonemic awareness and listening comprehension, which have been shown to predict later reading abilities. Research Triangle Institute (RTI), www.rti.org

¹⁹ To facilitate recognition, a large plasticized chart of 30 words of 1-3 syllables was presented to the student.

4. Listening comprehension assessed students' ability to listen to a story and answer five comprehension questions about the story without the challenge of decoding the words. This sub-test allows us to assess if difficulties stem primarily from low reading skills or from low overall language comprehension. The sub-test yielded a maximum score of 5.

5. Oral passage reading assessed students' fluency in reading a passage of grade-level aloud and their ability to understand what they read. There are two parts to this sub-test:

- a. Oral reading fluency (ORF): As described above, the ability to read passages fluently is considered a necessary component of reading comprehension. In this sub-test, students were given a 38-word story and asked to read aloud in one minute. The oral reading fluency score was the number of correct words in a text read by minute (ORF).
- b. Reading comprehension: After the students finished reading the passage, or at the end of one minute, the text was removed. Students were orally asked five questions that required them to recall basic facts from the passage. The reading comprehension score was the number of correct answers with a maximum possible score of 5.²⁰

Teacher Interview.

One of the enumerators administered a face-to-face interview with the teacher whose class had been selected for EGRA administration.

Head Teacher Interview.

This interview was conducted face-to-face with the head teacher. If he/she was not present, the director of studies was interviewed. Some items required head teachers to present proof of the answers they gave. For example, when asked whether they recorded teacher or student attendance or tardiness, they were asked to show the logs or forms used for this purpose.

Kinyarwanda Lesson Observation.

The focus of the observation was on the instructional behaviors exhibited by the teacher when teaching reading for one full Kinyarwanda lesson period. The enumerator used a structured observation protocol that listed reading instruction behaviors promoted by SU in the teacher training sessions. Instrument administration required that the enumerator arrive at the scheduled time for the class to start, to record the actual time the class started, and to stay until the end of the class—40 minutes.

The IE Team Leader observed some of the pilot training and pilot visits, plus some of the early data collection. Overlapping with her, the IE Research Analyst observed some of the final training and the beginnings of the data collection and conferred with Soma Umenye and USAID on setting up the datasets.

4.4 Enumerator Training and Data Quality Assurance

In January of 2017, IBTCI and Chemonics began planning for the data collection to take place in September/October of 2017. To assure consistency in data collection, it was agreed that

²⁰ Lookbacks—i.e., referencing the passage for the answer—may be permitted to reduce the memory load but are not typically used in the core EGRA instrument. (EGRA Toolkit, 2nd edition, 2016, page 27). However, assessors were instructed to allow students quick lookbacks.

enumerators from the IP and the IE would be trained jointly with the participation of staff from both teams.

Training was led by senior staff of the IP's subcontractor, EdIntersect, with the participation of the IP staff, the IBTCI IE Team Leader and IBTCI Field Director (Incisive Africa, a local research firm), and observation by the IE Research Analyst. The training provided to IBTCI's 75 enumerators and to Soma Umenye's personnel focused on preparing them to collect reliable data in the most effective and accurate way possible. The teams of supervisors were trained to support the effort of the data collection teams, advise the enumerators, clarify doubts and review all completed forms entered on tablets to identify missing data or incorrect entries.

A total of 60 field personnel and five people from the Quality Control team were trained on the EGRA tool and survey administration as well as on the use of tablets for data collection. Each instrument required a different procedure for administration. We should note that while IBTCI did have some opportunity to offer input, Soma Umenye was responsible for developing all the instruments and getting approval for their use from REB. (IBTCI was responsible for getting the requisite approvals from the Rwanda National Ethics Committee, the National Institute of Statistics of Rwanda, and the MINEDUC Directorate of Science, Technology, and Research, with Soma Umenye responsible for helping IBTCI to obtain the necessary letters of support from REB.) The first training was conducted August 22 to 24 to field test the instruments and procedures and to identify potential supervisors. Following the training, a pilot test was carried out twice in eight schools.

The second round of training took place from September 12 to September 22, prior to data collection. Data collection started on September 25 and was finalized by October 11. To guide the implementation of the assessment visits, a very detailed manual for supervisors was developed by the IE team with input from the IP and its subcontractor and revised as needed following the assessments at the pilot schools. Please see Annex F, the Supervisor Manual.

A team of four (one supervisor and three enumerators) was assigned to each school. Besides introducing the team and explaining the purpose of the visit, the supervisor interviewed the head teacher and the teachers, administered the reading lesson observation protocol and observed the work of the different enumerator teams calling attention to incorrect procedures, if any. Data collectors were able to complete a successful EGRA Assessment in an average of 45 minutes.

4.5 Data Overview and Data Cleaning Procedures

Upon the completion of assessments, observations and interviews, each supervisor was responsible for verifying the number of interviews done and uploading them (network coverage permitting) to the server after confirming the total number of instruments administered. A team of five data collection protocol adherence supervisors randomly accompanied different five data collection teams on each day to observe compliance and adherence to the data collection protocol.

The IE team developed a data cleaning protocol before data collection was completed. The cleaning protocol applied to the full set of data. The goal of the data cleaning process was to investigate the consistency and quality of the data. Some examples of the domains investigated

include whether the dates and times of the various tools in the dataset were consistent with the field work and school times, whether the rules set regarding number of male and female students were adhered to, whether the maximum times per question for the individual sub-tests were followed, whether the sub-test scores were in range, and, whether the responses regarding teacher and head teacher years' of experience were consistent with their reported ages. The evaluation team investigated the data and made inquiries to the field supervisors from the treatment and control data collection teams. The field supervisors looked through their internal field logs and data as well as consulted with their teams to provide responses to the questions posed. The IE team then corrected or discarded irrelevant data and prepared all data sets for analysis. Data cleaning and analysis was done using the statistical software STATA.

5. ANALYSIS OF STUDENT AND TEACHER OUTCOMES

5.1 Student-level Analysis

Outcomes

The student-level analysis utilizes student scores from all the sub-tests of the EGRA administered to students, as described in Section 4.2. For the four timed sub-tests, we selected the fluency score or the correct items identified/read per minute as the primary outcome for each sub-test, as adequate competency in these skills is a prerequisite for competency in the successively more advanced skills. For untimed sub-tests, we selected the percentage of correctly answered questions as the primary outcome.²¹

As part of the student-level analysis, we also collected data for subsequent analysis of the impact of Soma Umenye on additional student outcomes that may change as a result of students' improved performance in terms of their reading abilities. Specifically, we examined impacts of Soma Umenye on student's lateness to school and absenteeism, the only topic in which any credible impact of Soma Umenye could plausibly be expected to possibly appear at the end of partial implementation of PI. We utilized self-reported information from the student context survey as to whether the student had been late for school the day prior to the survey or absent from school during the week prior to the survey.

Empirical Model

To produce unbiased impact estimates, the impact analysis relies on a multivariate linear regression model. We utilize a statistical model to estimate program impacts for two reasons:

1. **Correctly specify the precision of impact estimates.** Although random assignment is conducted at the sector level and implementation is carried out at the school level, outcomes are measured at the student level. This situation can cause us to overstate the precision of impact estimates if it is not addressed statistically. We use a regression specification to correctly estimate standard errors of impact estimates.
2. **Improve the precision of impact estimates.** Despite random assignment, it is possible that small differences in observable pre-intervention characteristics across research groups may arise by chance or from missing data. The statistical model controls for any such differences in observable characteristics. Estimating program impacts in this way will improve the precision of the impact estimates (Raudenbush 1997).

The statistical model for the impact analysis relies on a linear regression model. This model can be used for both continuous (such as correct words per minute) and binary outcomes (such as whether the student scored zero or not). It can be expressed as follows:²²

$$(5.1) \quad y_{ist} = \alpha + \beta x_{is0} + \gamma z_{s0} + \delta T_s + B_s + \varepsilon_{ist}$$

where y_{ist} is the outcome of interest (such as correct words per minute) for student i in school s , and at time t . The vector x_{is0} represents the time-invariant characteristics of student i in school

²¹ We choose one outcome for each subtask to prevent multiple comparison problem. We discuss this in detail in Annex J.

²² For binary outcomes, this specification is referred to as a linear probability model. The linear probability model is our preferred specification because of its ease of interpretation.

s, such as age and gender measured at follow-up (we do not have baseline data).²³ A complete list of these control variables is in Table Annex J.II. The vector z_{s0} represents the pre-intervention school and sector characteristics of school s, such as enrollment and rural/urban status. A complete list of these control variables is in Table Annex J.II. The variable T_s is an indicator equal to one for students in treatment group schools and zero for those in control group schools. The term B_s represents a set of dummy variables indicating the randomization block²⁴, while ε_{ist} is a random error term for student i in school s observed at time t.

The parameter δ is the regression-adjusted difference in mean outcomes between the treatment and the control groups, which gives the impact of SU on the outcome of interest. The program effect estimates provide what is known as the “intent-to-treat” (ITT) effect. It estimates the average difference in outcomes between all students in target grades at time t in treatment and control schools, regardless of whether the students in the treatment schools participated in program activities (by attending class everyday) and regardless of whether the school implemented the intervention. Such an estimate can be interpreted as the average impact on the population of students in schools being given the option to implement the intervention. This is the key parameter of interest, since from a policy perspective this shows the impact of introducing the intervention into schools. This will not necessarily involve take-up by all students or full implementation by the school.

In addition, we adjust the standard errors of the impact estimates student observations clustering at the school level using the Huber-White “sandwich” estimator of variance. Because all students from a school are sampled from the same classroom, this is essentially an adjustment of student observation clustering at the classroom or teacher level as well. We also apply design weights. Because different randomization blocks had different number of sectors to be randomized, the probability of any one sector to be assigned to the treatment group varied by block. The design weights are thus calculated as the inverse of the probability of being assigned to treatment or control, calculated within each block. Finally, for missing values of control variables from non-responses or missing data we imputed values according to the following rules. For variables from the student context survey, we imputed the values using the average values for the school. For missing values at the sector level for pre-intervention data, we imputed using average values for the district. None of the outcome variables are imputed.

5.2 Teacher-level Analysis

Outcomes

The teacher-level analysis focused on examining the impact of Soma Umenye on teachers’ instructional practices when teaching reading to early graders. We assessed teachers’ performance using the lesson observation protocol, which included five domains: (1) lesson preparation, (2) teaching methodology, (3) literacy technical skills, (4) assessment skills, and (5) use of teaching materials. Within each of these five domains, teachers were scored on a number

²³ Ideally, we will include baseline measures of outcome y for each student i. However, we do not have baseline data for any of the outcomes on PI students related to academic performance.

²⁴ Recall that randomization was carried out within two blocks in each district, one that contained sectors with more than the district average number of schools (larger sectors) and one that contained less than or equal to the average number of these schools in the district. Because the blocks were constructed using pre-existing sector characteristics to match them as closely as possible, controlling for block fixed effects are appropriate for this stratified random assignment design (Duflo et al., 2008).

of items; for each item they received a score of 1 if the observer recorded “yes” for the item and zero if the observer recorded “no”. For example, in the lesson preparation domain, one of the items was “whether register is available and completed daily”. Teachers received a score of 1 if the observer recorded “yes” and 0 if “no”. For each teacher, we then calculated the total score for each of the domains and then the total raw classroom observation score by adding scores across all domains.

To facilitate comparison across the treatment and control group teachers, we constructed a composite standardized classroom observation score for each teacher. Following King, Liberman, and Katz (2007), we created each teacher’s composite score by first subtracting the control group mean score and then dividing by the control group standard deviation.

As in the student-level analysis, we selected this single composite standardized classroom score as our primary outcome to prevent a multiple comparison problem.²⁵

Empirical Model.

The teacher-level model mirrors the student-level model and is as follows

$$(5.2) \quad y_{ist} = \alpha + \beta x_{is0} + \gamma z_{s0} + \delta T_s + B_s + \varepsilon_{ist}$$

At the teacher-level, standard errors were adjusted for clustering at the district level (as opposed to the school level with the student-level analysis), also using the Huber-White “sandwich” estimator of variance. Design weights were applied to the model. We controlled for several student and teacher characteristics which are presented in Annex I. Missing values from the teacher survey and pre-intervention data were imputed using average values from the district.

²⁵ We discuss the multiple comparison problem in detail in Annex J.

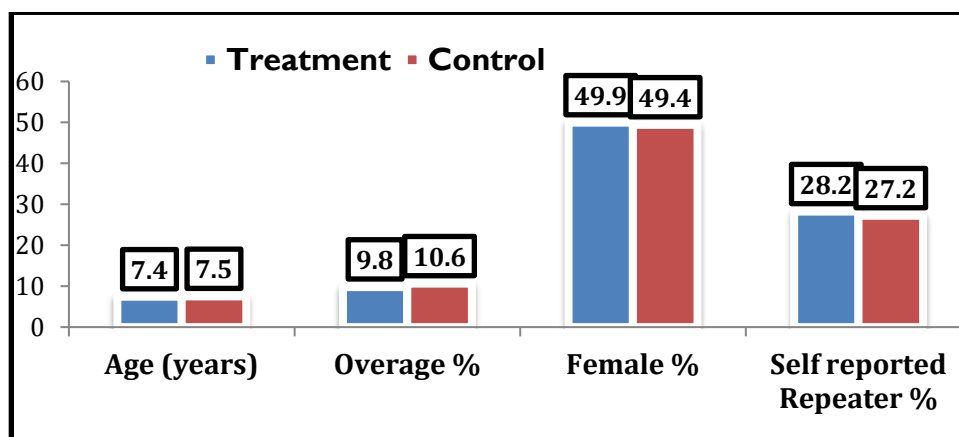
6. DESCRIPTION OF THE SAMPLE

Section 6 describes the characteristics of the student population assessed and interviewed and of their teachers interviewed and observed as well as characteristics of the schools and of the headmasters included in the sample. These characteristics provide a useful picture of the context learners and teachers face daily. The 5,456 students assessed and interviewed are roughly 0.9 percent of the 606,712 students enrolled in all PI schools (including private schools) in 2016.

6.1 Characteristics of the Students Interviewed and Assessed

Student demographic characteristics can be a factor related to their performance. Students in our sample were split evenly between females and males because we deliberately sampled by gender as explained in Section 4. The expected age for a PI student in the Rwandan system is seven years of age and students were, on average, seven and a half years old. About 10 percent were overage for their grade as defined by being more than three years older than they should be at PI. We also asked students participating in the EGRA if they were repeating PI. Repetition rates are extremely high in PI with over 25 percent of students sampled reporting that they were repeating each year. Figure 6.1 displays the self-reported age of students, the proportion of overage students, and the percentage of female students in the sample.

Figure 6.1. Sampled Students Demographics (percentages of students except age)

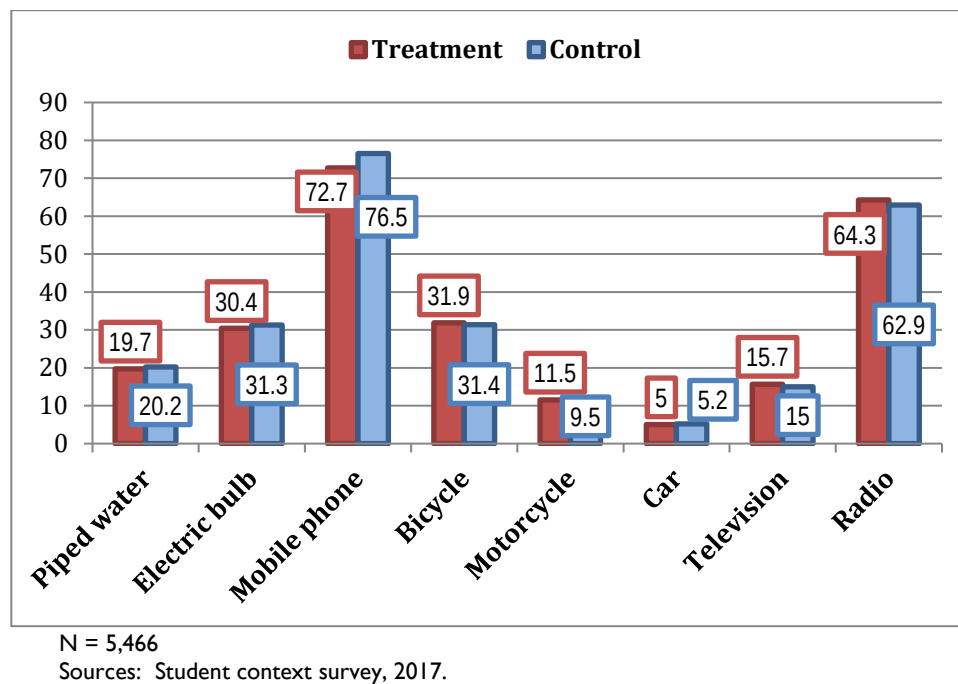


N=5,466

Repeaters include those who were in P1, P2, or P3 in the year prior to the survey and were re-enrolled in P1 during the time of the survey.

The Student Context Questionnaire included questions related to the household assets. Figure 6.2 displays the information provided by students.

Figure 6.2. Household Assets Reported by Students Who Took the EGRA



6.2 Characteristics of the Teachers Included in the Sample

A structured interview protocol was administered to 287 teachers of classes randomly selected to participate in the EGRA. The tables that follow describe the characteristics of the teachers interviewed. The information reported here refers only to the IE sample.

Kinyarwanda primary school teachers are overwhelming female—81.8 percent in the treatment schools and 82.6 percent in control schools. Intervention schools had slightly fewer younger (≤ 25 years old) and slightly fewer older teachers (≥ 65 years old) than did the Control schools (Table 6.1).

Table 6.1. Age of Teachers Interviewed/Observed

Teacher Age	Intervention	Control	Overall
≤ 25 years old	1.5%	3.2%	2.4%
26 - 45 years old	72%	74.9%	73.4%
46 and older	26.5%	21.9%	24.0%
Sample size	132	155	287

Of the 287 teachers interviewed, 23.7 percent have fewer than five years of experience. The majority (65.8 percent) have between five and twenty years of experience, and only 11.5 percent reported having been teaching over twenty years (Table 6.2). This has implications for in-service training and also for the hiring and deployment of teachers.

Table 6.2 Teacher Years of Experience

Teacher Years of Experience	Treatment	Control	Total
Less than 5	18.9%	27.7%	23.7%
5-10	34.1%	29.7%	31.7%
11-20	34.8%	31.6%	33.1%
21 +	12.1%	10.9%	7.0%

N=287

6.3 School and Classroom Characteristics

Head teachers were asked to use school records to answer questions related to school resources and enrollment. A small percentage of the schools, 1.6 percent, in our sample did not have any desks for PI students; on average, three students had to share a desk. We also asked head teachers if students had basic school supplies such as pens, pencils, and notebooks. In 10 percent of the schools, more than half of the PI students did not have these basic school supplies. This information is important since like better instructional methods or curricula, basic school resources are important complements for learning. The treatment and control schools are similar on indicators related to resources, with slightly more control schools not having desks for PI students (1.8 percent in control schools compared to 1.5 percent in treatment schools) and slightly more treatment students not having sufficient supplies than control students (11 percent treatment students compared to 8.9 percent control students).

The number of students enrolled in a grade is a proxy for school size, which can be an important factor in student learning. The average number of PI students enrolled per school is 232. Treatment schools have slightly more PI students enrolled (240.46) than control schools (225.88). On average, there are two PI Kinyarwanda teachers per school.

Inherently a child who repeats a grade increases the class size, and a child who repeats and still makes limited progress is more likely to drop out than a child who repeats and succeeds. Therefore, it is important to know the proportion of repeaters in both groups. Overall, there were 19.85 percent PI repeaters per school. The number of repeaters is slightly less in the treatment group (18.81 percent) compared to the control group (20.71 percent). The level of students who dropped out of school is an important measure to calculate, as it can be hypothesized that those who drop out are the poorest performers, which could lead to overestimates in learning outcomes of those sampled. Both groups have similar levels of dropouts on average (approximately 3.28 percent), with only a .02 percent difference between them. Table 6.3 displays the average school characteristics in the sample gathered through the Head Teacher Survey.

Table 6.3 Average School Characteristics in the Sample (percentages unless otherwise noted)

Characteristic	Treatment	Control	Overall
<i>Panel A: School status</i>			
Primary only school	41.9%	45.2%	43.8%
Public school	33.1%	29.2%	30.9%
Government-aided school	18.4%	20.8%	19.7%
<i>Panel B: School Resources</i>			
Schools without desks for PI students	1.5%	1.8%	1.6%
Number of PI students sharing desk	3.38	3.41	3.40
More than half the students do not have basic school supplies (pen/pencil/notebook/uniform)	11.0%	8.9%	9.9%
Number of PI classrooms in morning shift	3.83	3.41	3.41
<i>Panel C: Enrollment^b</i>			
Number of PI students	240.46	225.88	232.52
Number of repeaters	55.99	57.79	56.97
Number of drop outs	25.34	21.01	22.98
Number of PI Kinyarwanda teachers	2.09	2.02	2.05
Student-Kinyarwanda teacher ratio in PI	114.21	126.68	121.03
<i>Sample size</i>	136	168	304

Sources: Head teacher survey, 2017 (information in Panel C are directly captured from the school records).

Note: Standard errors are presented in parentheses, clustered at the school level. Sample sizes for some characteristics may be smaller due to missing data.

7. FINDINGS

In this section we present and discuss findings resulting from the analyses of the quantitative EGRA data, of the interviews conducted with students and their teachers and head teachers, and of the lesson observations conducted at the classes randomly selected for assessment. This report primarily addresses Evaluation Questions 1 and 2 and presents starting-point information on absenteeism for Evaluation Question 3. Research on other questions was to have begun during the 2018 school year. The main focus of Evaluation Question 1 is the impact of SU on student reading skills. Whenever there was the possibility of detecting impact data were disaggregated to show impact on sub-groups of interest: male and female students, repeaters and non-repeaters, students in urban and rural settings. Class observations and interviews conducted with students, teachers, and head teachers were utilized to provide a context for the findings and an answer to Evaluation Question 2. The findings directly address the impact of the processes and strategies implemented by Soma Umenye in 2017. SU monitoring data were used to provide context for the findings (for example, the level of implementation of SU's activities), but were not used to determine the impact of the intervention on students and teachers.

7.1 EGRA Findings

Our key findings are related to the impact of Soma Umenye on student EGRA scores. Table 7.1 shows regression-adjusted scores by treatment group for the four timed EGRA sub-tests and points out where the differences between the two groups are statistically significant. Note that the number of items per sub-test differs. For example, the Letter identification and syllable identification sub-tests included 100 items each; Familiar Word Reading included 50 words/items; and, Reading Fluency required that students read as many words out of 38 (number of words in the total text) as they could in one minute. Figure 7.1 displays the distribution of scores by treatment group.²⁶

Table 7.1. Scores on the EGRA Timed Sub-tests by Treatment and Control Group

Sub-group	Treatment	Control	Difference between groups
Size of Sample	2,454 (44.9%)	3,003 (55.1%)	557 (10.2%)
EGRA timed sub-tests and maximum scores possible	Mean Score	Mean score	Score Difference
Letter Identification (MAX 100)	22.19	19.41	2.77*
Syllable identification (MAX 100)	12.11	11.19	0.93
Familiar Word Reading (MAX 50)	4.51	3.72	0.79*
Reading Fluency (MAX 38)	4.98	4.27	0.72

N=5,457

* Differences between treatment and control statistically significant at .01 and .05 level.

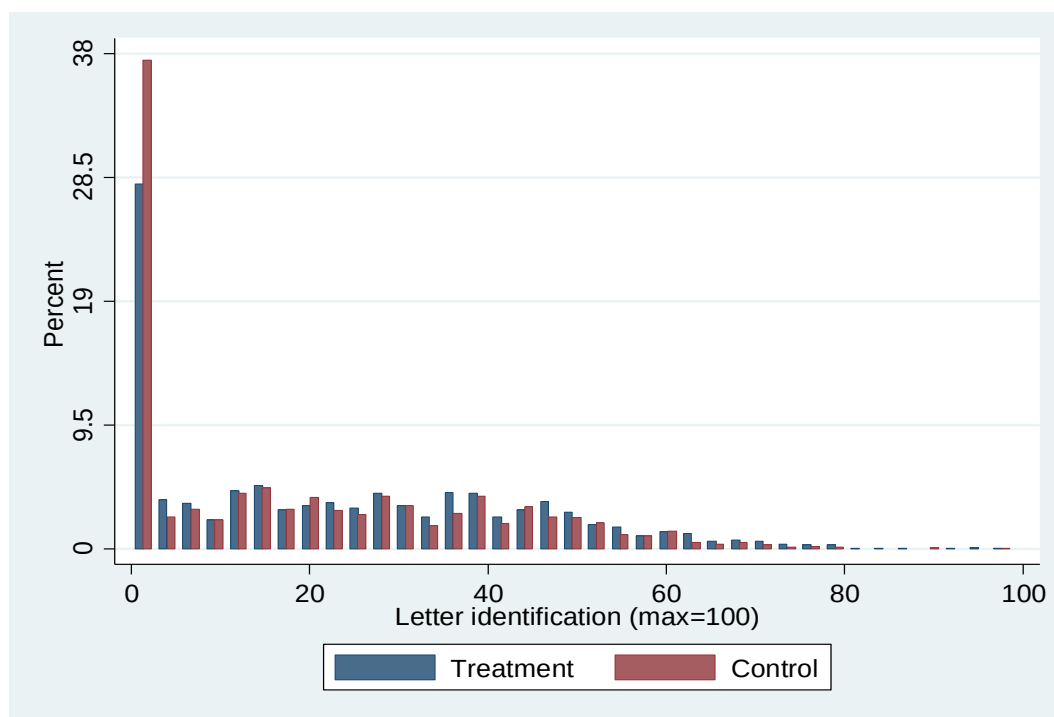
dpm: correct letters per minute; csilpm: correct syllable per minute; cfwpm: correct familiar words per minute; ORF: oral reading fluency

²⁶ We present impacts on secondary outcomes in Tables Annex J.3-Annex J.8.

Soma Umenye had statistically significant positive impacts on students' timed test scores on the letter identification and familiar words reading EGRA sub-tests. On average, students in the treatment group identified 22 letters correctly per minute compared to 19 identified by control group students, on average. Students in the treatment group were also able to read, on average, about 5 familiar words per minute compared to 4 in the control group. The differences between the two groups on these two sub-tests were statistically significant at the 1 percent and 5 percent level respectively.

Upon examining the score distributions, it appears that the differences came from a higher percentage of students in the treatment group scoring non-zero scores. Figure 7.1 shows the distribution for the letter identification sub-test scores by treatment and control groups. The scores are skewed to the left due to a large number of students in both treatment and control groups scoring zero. However, a larger fraction of students in the control group had a score of zero (32 percent) compared to the treatment group (24 percent). For the rest of the distribution there was little visible difference.

Figure 7.1. Score Distribution on Letter Identification Sub-Test by Treatment and Control

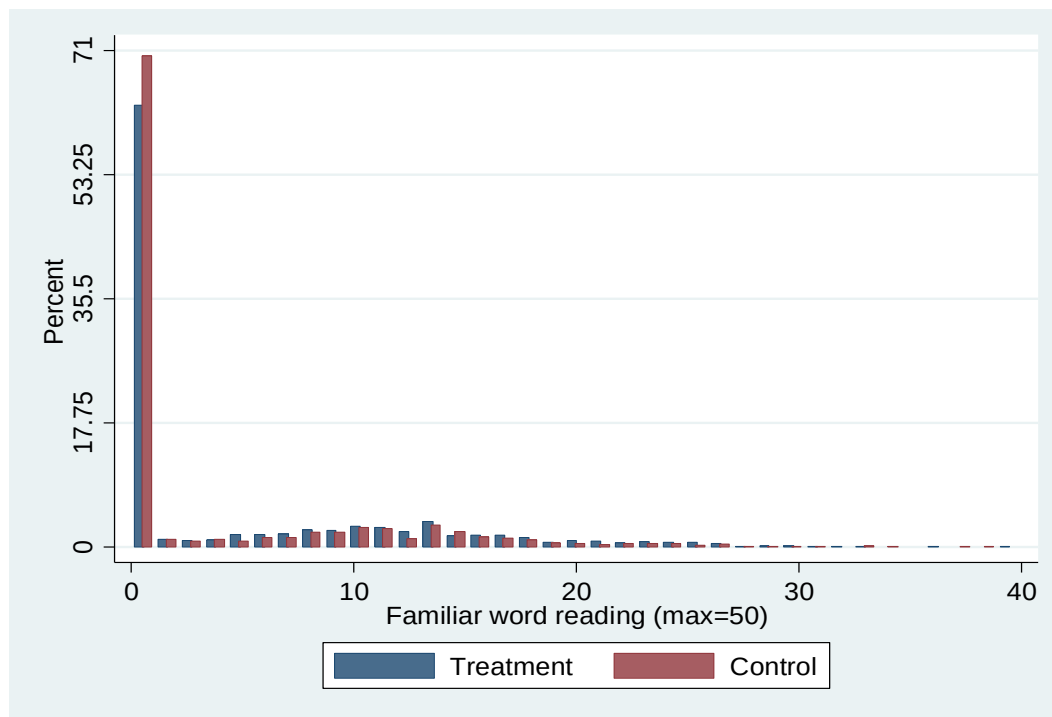


We observe a similar trend for the familiar word reading sub-test. As shown in Figure 7.2, the fraction of students in the control group scoring a zero is much larger (69 percent) than the fraction of students scoring a zero in the treatment group (62 percent). There is no visible difference in the rest of the distribution.²⁷

²⁷ Impact on zero scores in familiar word reading subtask is also presented in Annex J Table J.5.

We conclude that Soma Umenye was able to affect the number of students that scored a non-zero or got at least one item correct. This is in line with research in other contexts that found that improvements in reading abilities for early grades usually begin to appear in lower tail of the score distribution.²⁸

Figure 7.2 Score Distribution on Familiar Word Reading Sub-Test by Treatment and Control



Soma Umenye was successful in impacting student scores in two of the EGRA timed sub-tests, but the findings still show very low overall performance in the four timed sub-tests. On average, students identified only 19 to 23 **letters** out of the 100 presented to them and were able to read less than one-sixth of the 50 **familiar words**. For the two timed subtasks for which we do not observe any impacts—syllable identification and oral reading fluency—students in both groups had very low average scores. As shown in Table 7.1, student identified an average of 12 or 11 syllables in the treatment and control group, respectively, out of a maximum of 100. For oral reading fluency, students in both groups could read between 4-5 words in one minute on average out of a maximum of 38 (Table 7.1). A very high percentage of students could not read any words, 63 percent in the treatment group and 70 percent in the control group.²⁹ Considering that 2017 was Soma Umenye’s first year of implementation, that teacher training began in March (the Rwanda school year runs from January/February to October/November), that evidence-based teaching and learning materials have not yet been distributed, and that generally speaking projects become more efficient as implementation proceeds, we hypothesize that in 2018

²⁸ <http://documents.worldbank.org/curated/en/319981468060904888/pdf/ACSI3261-REVISED-PUBLIC-Myanmar-Early-Grade-Reading-Assessment-IDU.pdf>

²⁹ Percent of students with zero scores in oral reading fluency subtask is also presented in Annex J Table J.6.

and even more so in 2019 the effects of the project will be larger and that the gap between students in the treatment and in control schools will widen. However, without an Impact Evaluation, it will not be possible to confirm this hypothesis.

We examined student scores on untimed sub-tests³⁰—**listening comprehension** and **reading comprehension**. Students in the treatment group scored statistically significantly higher on listening comprehension compared to their counterparts in the control group. However, both groups scored high on this sub-test, above 3.78 on average out of a possible score of 5. Given this very high score in the first year of the project, there is little room for Soma Umenye to improve the score of this sub-test in the future.

Table 7.2. Scores on the EGRA Untimed Sub-tests for Treatment and Control Group

Sub-group	Treatment	Control	Difference between groups
Size of Sample	2,454 (44.9%)	3,011 (55.1%)	557 (10.2%)
EGRA timed sub-tests and maximum scores possible	Mean Score	Mean score	Score Difference
Listening Comprehension (MAX 5 questions)	3.91	3.78	0.13*
Reading Comprehension (MAX 5 questions)	0.69	0.60	0.09

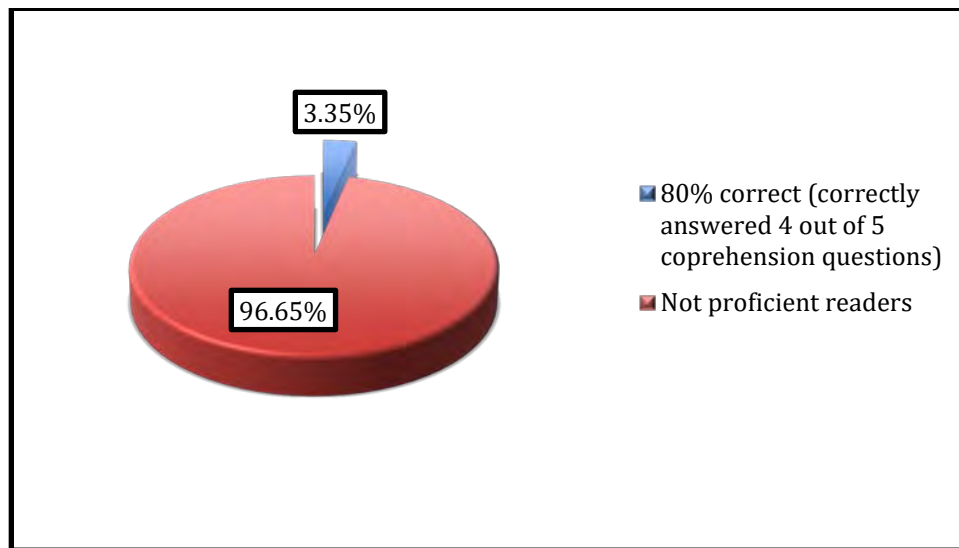
N=5,457

* Differences between treatment and control statistically significant at .05 level.

Average scores on the **reading comprehension** sub-test were extremely low for both groups—0.69 for students in the treatment group and 0.60 for their counterparts in the control group out of a possible score of 5.0. The difference between the average score of students in the treatment and control groups is not statistically significant. More than seventy percent of the students assessed received a score of zero, meaning that they were not able to answer any of the five comprehension questions. Figure 7.3 visually depicts the findings.

³⁰ Though these sub-tests did not have a sixty-second total time period, enumerators were instructed to move to the next question if the student could not answer the question in a specified period of time.

Figure 7.3. Students Able to Read with Comprehension



N=5,466

Overall, 3.09 percent students (N=169) scored 80 percent or higher and 0.26 percent scored 100 percent. In the treatment group 3.18 percent of the students (N=78) scored 80 percent or higher; in the control group 3.02 percent (N=91) scored 80 percent or higher. Differences between the groups were not statistically significant. The average ORF score for proficient readers—students who scored 80 percent or higher in reading comprehension—was 27.8 correct words per minute. Because of the low performance in words per minute (Table 7.1) it is not surprising that reading comprehension scores concentrated heavily at the lower end of the allowable range. EGRA studies have shown the high correlation (0.91) between correct word read by minute (fluency) and reading comprehension.³¹ We observe almost similar correlations in our data as well. Correlations between oral reading fluency (orf) scores and reading comprehension scores were 0.92 in the overall sample, 0.91 in the treatment group, and 0.93 in the control group.³²

7.2 Lesson Observation Findings

A total of 281 observations of the Kinyarwanda reading lesson were conducted in September/October 2017 at the class randomly selected for EGRA participation in each school. The protocol scored teachers on lesson preparation, teaching methodology, literacy technical skills, assessment techniques, and use of teaching materials. Our analyses show that teachers in treatment schools (teachers who received Soma Umenye training) were more effective in their classroom instructional methods than their counterparts at schools that did not receive training.

³² Of the students who scored more than 27.8 in orf, 53.5 percent scored at or above 80 percent in reading comprehension. On the other hand, only 1.9 percent of the students who scored less than 27.8 in orf scored at or above 80 percent in reading comprehension.

Table 7.3. Impacts of Soma Umenye on Classroom Instruction – Raw Scores

Dependent Variables	Treatment	Control	Impact
Sample size	129	152	
Total lesson observation raw score	20.07	17.88	.45**
Lesson preparation raw score	3.01	3.02	.19
Teaching methodology raw score	5.5	5.34	.27
Literacy technical skills raw scores	4.28	4.23	.06
Assessment raw score	3.81	2.85	.96*
Use of teaching materials raw score	3.47	2.54	.92**

N=281

Source: Lesson Observation Protocol 2017.

* Impact (difference between treatment and control group means) is statistically significant at the .05 level.

From the results of the analyses displayed on Table 7.3, we created a composite standardized score. An explanation of how we created this score is described in Section 4.

Table 7.4. Impacts of Soma Umenye on Classroom Instruction – Composite Score

Dependent Variables	Treatment	Control	Impact
Composite classroom observation standardized score	0.45	0.00	.45*
Sample size	129	152	

Sources: Classroom Observation 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated.

Teachers in the treatment schools received an average composite score that was half a standard deviation better compared to teachers in the control group and the difference is statistically significant. This finding is promising because one of Soma Umenye's major intervention components is teacher training. The positive impact of Soma Umenye on teacher technical skills is likely to impact student reading skills in the longer run.

7.3 Findings for Additional Student Outcomes

Improvements in reading skills can enhance student performance on other academic areas, and Evaluation Question 3 called on the Impact Evaluation to answer the question, *What is the evidence that improved student reading outcomes in Kinyarwanda lead to improved opportunities for students to succeed in schooling?*

To begin to address this, we estimated impacts of Soma Umenye on student self-reported absenteeism and tardiness as indicated on their Student Context Forms. More than half of the students in both treatment and control groups reported being absent at least once in the week preceding the assessment. Because the data were self-reported and the differences in attendance rates were not statistically significant, we conclude that, while

participation in Soma Umenye may contribute to improved student academic achievement and/or improved efficiency down the road, it was not possible to observe any impact during the first year of implementation.³³

Table 7.5. Impacts of Soma Umenye on Self-Reported Class attendance

Dependent Variables	Treatment	Control	Impact
Absent from school last week (%)	55.6	53.8	1.8
Late to school yesterday (%)	14.5	13.9	0.6
Sample size	2,419	2,975	5,394

Source: Student Survey 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools.

7.4 Sub-Populations of Interest

In addition to analyzing differences between treatment and control, IBTCI was called upon to identify whether or not the Soma Umenye activity differentially affected the reading outcomes of girls and boys, children living in urban and rural areas, and children representing different levels of socio-economic status, inter alia.

What are the impacts of Soma Umenye on females and males?

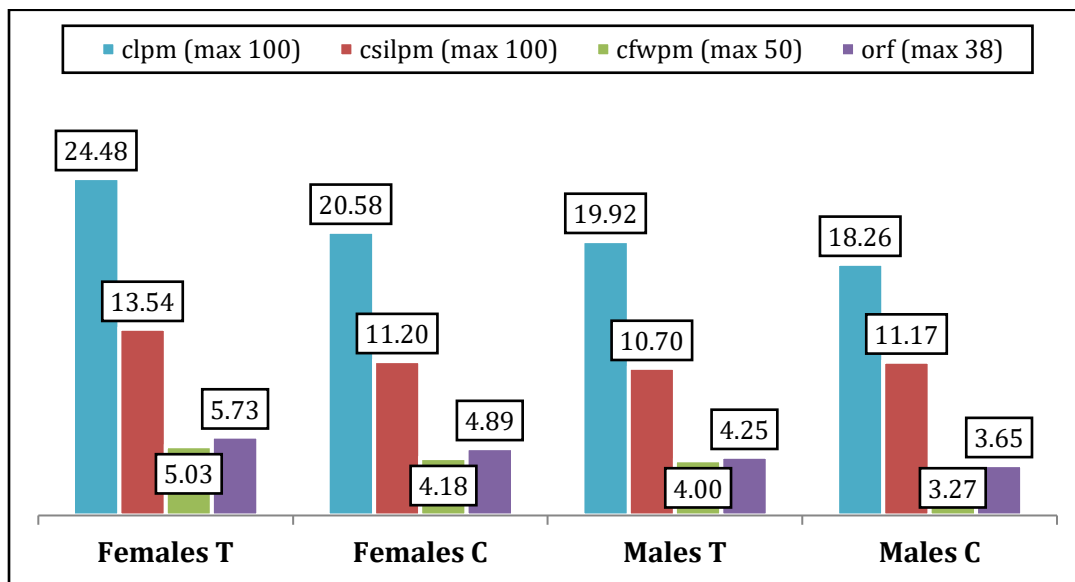
To provide answers to this question, we examined the subgroups that make up the sample to determine whether SU benefitted different groups of students differently. We estimate the impacts of Soma Umenye separately for the female students and then for the male students. Our approach to estimating these impacts is the same as for the full sample: comparing scores of students in the treatment group to scores of students in the control group.

Female students in the treatment group outperformed female students in the control group on the **letter identification**, **syllable identification**, and **familiar word reading** sub-tests at a statistically significant level. However, the magnitude of the impacts in practical educational terms is not substantial; on average, female students in the treatment group identified about 4 more letters, 2 more syllables, and 1 more familiar words. Male students in the treatment group outperformed male students in the control group only in the **familiar words reading** and **listening comprehension** sub-tests at a statistically significant level. Once again, however, the magnitude of the impacts was not substantial; male students in the treatment group identified 1 more familiar word and provided 2 percentage points more correct answers for the listening comprehension sub-test. The findings are displayed on Figure 7.4.

³³ Findings for the impact of Soma Umenye on other student outcomes, such as mathematics, reading at home, drop-out, and repetition were to have been examined for Midpoint in 2018 and Endpoint in 2019.

Next, we compared scores of female students in the treatment group for the EGRA untimed sub-tests with the scores of female students in the control group. For the untimed sub-tests there was no statistically significant differences between the performances of female students from either group. We conducted the same analysis for male students and found no statistically significant differences between the male students, either.

Figure 7.4. Scores on timed EGRA sub-tests for female and male students by treatment group



clpm=correct letters per minute; csilpm=correct syllables per minute; cfwpm=correct familiar words per minute; orf=oral reading fluency or correct words connected text per minute

To what extent do the impacts of Soma Umenye differ across female and male students?

The findings presented above show that Soma Umenye was successful in having small but positive impacts on selected EGRA sub-tests on both female students and male students in the treatment schools compared to respective subgroups of students in the control group. In addition, we also examined whether the impacts Soma Umenye had on female students are statistically different from the impacts the program had on male students. The primary research question we are addressing here is, “whether and to what extent the impacts of Soma Umenye differed across female and male students?” For this analysis, we compare the impacts of Soma Umenye on female students (difference between female treatment students and female control students) with the impacts of Soma Umenye on male students (difference between male treatment students and male control students).

Table 7.6 displays the differential impacts on student EGRA scores for female and male students for each EGRA sub-test. The impacts on female students for all but one of the sub-tests (listening comprehension) were higher than the impacts on male students. However, none of the differences between males and females were statistically significant, meaning that, on average, Soma Umenye benefitted both groups equally.

Table 7.6. Differential Impacts of Soma Umenye on Student Outcomes, by Gender

<i>Dependent Variables</i>	Impacts on Females	Impacts on Males	Differential Impact (Impacts on female-impacts on male)
Letter identification, items correct per minute	3.90*	1.66	2.23
Syllable identification, items correct per minute	2.33*	-0.46	2.80
Familiar words, words correct per minute	0.85*	0.73*	0.12
Oral reading fluency, words correct per minute	0.84	0.59	0.25
Listening comprehension, percent of items correct	2.05	2.26*	-0.20
Reading Comprehension, percent of items correct	1.08	1.39	-0.31
Sample Size	2,707	2,750	

Sources: EGRA administration data, 2017.

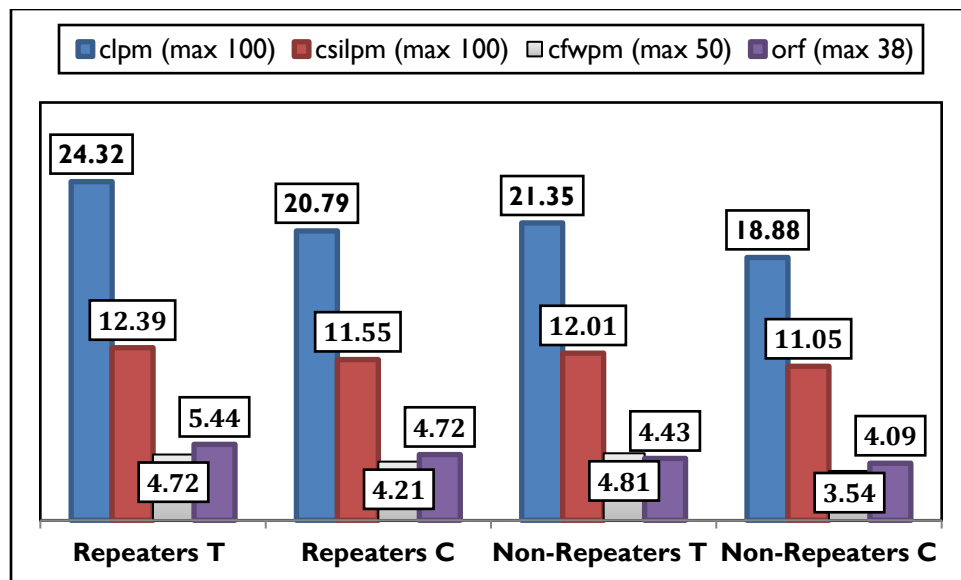
Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools.

* Impact (difference between treatment and control group means) is statistically significant at the .01/.05 level.

Does Soma Umenye impact repeaters and non-repeaters differently?

We estimated impacts for students who were repeaters and students who were attending PI grade for the first time (non-repeaters) to examine whether Soma Umenye had different impacts on these subgroups of students. Similar to our analysis for the gender subgroups, we first examine impacts for the repeater and non-repeater group separately. Figure 7.5 displays scores of repeaters and of non-repeaters by treatment group for the EGRA timed sub-tests, respectively Soma Umenye had statistically significant positive impacts on repeating treatment students in the letter identification sub-test. Repeaters in the treatment group identified approximately 4 more words on average than repeaters in the control group. There were no significant impacts on any of the other timed or untimed sub-tests.

Figure 7.5. Scores on Untimed EGRA Sub-tests for Repeaters and Non-repeaters by Treatment Group



N=5,466

lpm=correct letters per minute; csilpm=correct syllables per minute; cfwpm=correct familiar words per minute; orf=oral reading fluency or correct words connected text per minute

* Differences between treatment and control statistically significant at .01 level.

Non-repeaters in the treatment group outperformed non-repeaters in the control group in the letter identification and the familiar words reading sub-tests at a statistically significant level (Figure 7.5). They identified approximately 2.5 more letters and almost one more familiar word compared to the repeating students in the control group.

On untimed sub-tests, non-repeaters from the treatment group scored slightly higher on listening comprehension, about 3 percentage points, compared to their control group counterparts at a statistically significant level.

Similar to our analysis for the gender subgroups, we also examined whether the impacts Soma Umenye had on repeater students are statistically different from the impacts the program had on non-repeater students. For this analysis, we compare the impacts of Soma Umenye on repeater students (difference between repeater treatment students and repeater control students) with the impacts of Soma Umenye on non-repeater students (difference between non-repeater treatment students and non-repeater control students).

Table 7.6 presents the impacts on student EGRA scores by repetition status for each EGRA sub-test. There is no clear pattern in terms of the differences in impacts between the two groups. However, none of the differences in impacts between the two groups are statistically significant. This implies that Soma Umenye affected both groups of students, repeaters and non-repeaters, equally so far.

Table 7.7. Differential Impacts of Soma Umenye on Student Outcomes, by Grade Repetition Status

Dependent Variables	Impact on repeaters	Impact on non-repeaters	Differential Impact (Impact on repeaters-impact on non-repeaters)
Letter identification, items correct per minute	3.53*	2.47*	1.06
Syllable identification, items correct per minute	0.83	0.96	-0.13
Familiar words, words correct per minute	0.51	0.90*	-0.39
Oral reading fluency, words correct per minute	0.72	0.71	0.00
Listening comprehension, percent of items correct	0.16	2.93*	-2.77
Reading Comprehension, percent of items correct	0.78	1.41	-0.63
Sample Size	1,510	3,947	

Sources: EGRA, 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools.

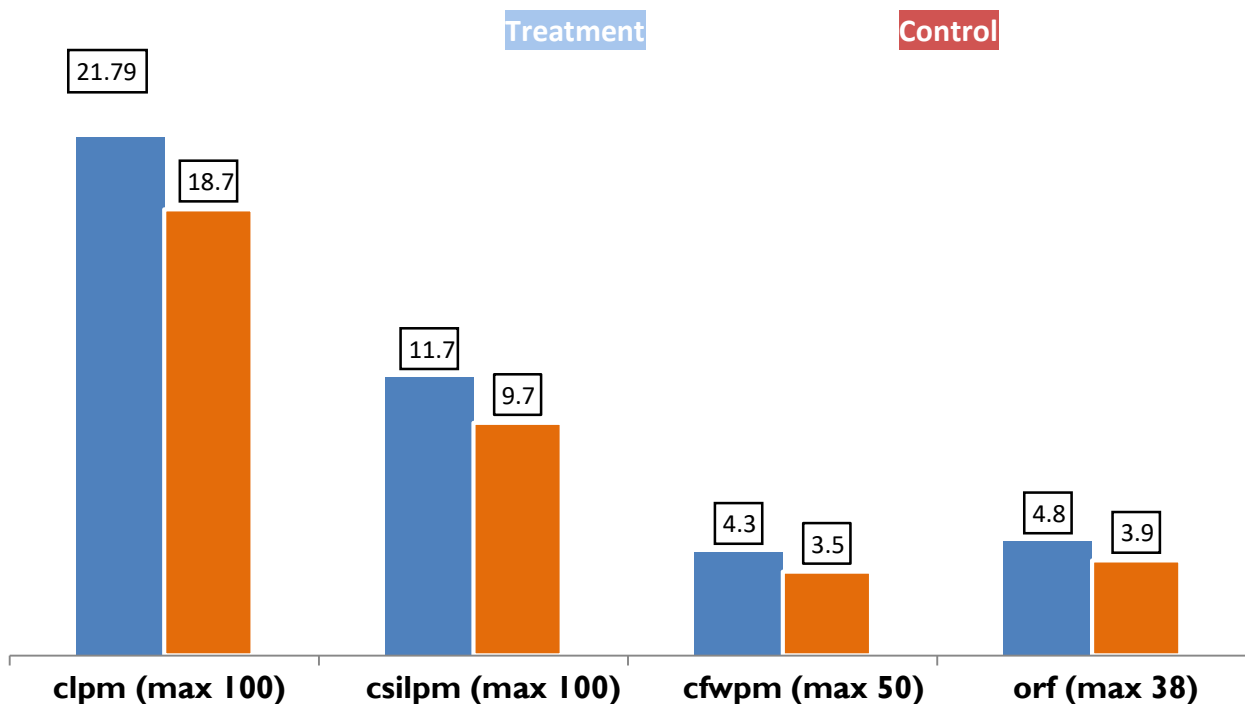
* Impact (difference between treatment and control group means) is statistically significant at the .01/.05 level.

Does Soma Umenye impact students in Urban and Rural sectors differently?

Finally, we also examined impacts for students in rural and urban sectors separately. We used the official definition of rural-urban status of sectors as provided by the Rwanda Education Board. Our sample, which was randomly selected as described in Section 4.2 above, was overwhelmingly rural: 4,932 (90.4 percent) students were from 275 schools in rural sectors and only 525 (9.6 percent) were from 29 schools in urban sectors, closely tracking the actual ratio of students and schools in Rwanda. Because of the high proportion of rural schools and students to the total populations, impacts on students from rural sectors are very likely to resemble the impacts for the full sample of students.

Figure 7.6 shows the comparison between the scores of students in the treatment group in rural and in the urban sectors to scores of students in the control group in rural and control sectors. For rural students, Soma Umenye had statistically significant positive impacts for all timed sub-tests (Figure 7.3) and also on the listening comprehension sub-test. On average, students in the treatment group in rural sectors identified 3 more letters, 2 more syllables, and 1 more familiar word. Although these impacts are very small and the overall level of fluency in the sample is very low, the positive impacts of Soma Umenye on students in rural sectors indicate that the intervention will likely succeed in improving reading fluency in rural areas as the program continues.

Figure 7.6. Scores on Timed Sub-tests for Students in Rural Sectors by Treatment Group



N=4,932

* Differences between treatment and control statistically significant at .05 level.

We did not observe any impacts of Soma Umenye on urban sector students on any of the EGRA sub-tests. Because the sample size of students from urban sector is very small, 525 students, it is very likely that we did not have the statistical power to detect any impacts. In other words, the “no impacts” for students in urban sectors could either be true results because there were in fact no impacts or it could be the result of the fact that the sample size is not large enough to detect statistically significant impacts.

We examined whether the impacts of Soma Umenye on students from rural sectors are statistically different from the impacts the program had on students from urban sectors. For this analysis, we compared the impacts of Soma Umenye on rural students (difference between rural treatment students and rural control students) with the impacts of Soma Umenye on urban students (difference between urban treatment students and urban control students). Although the impacts of Soma Umenye on rural students were higher than the impacts on urban students (last column of Table 7.7), none of the differences is statistically significant, implying that the program affected rural and urban students equally. This result also lends further support to the fact that the no impacts for urban students is more likely to be the result of small sample size rather than there being no true impacts.

Table 7.8. Differential Impacts of Soma Umenye on Student Outcomes, by Urban-Rural Status

Dependent Variables	Impact for rural sectors ^a	Impact for urban sectors	Differential Impact (Impacts on rural – impacts on urban sectors)
Letter identification, items correct per minute	3.05*	0.11	2.95
Syllable identification, items correct per minute	1.93*	-8.50	10.43
Familiar words, words correct per minute	0.85*	0.21	0.64
Oral reading fluency, words correct per minute	0.82*	-0.27	1.08
Listening comprehension, percent of items correct	2.39*	-0.04	2.43
Reading Comprehension, percent of items correct	1.29	0.72	0.57
Sample Size	4,932	525	

Sources: EGRA, 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools.

* Impact (difference between treatment and control group means) statistically significant at the .01/.05 level.

^a Urban and rural have been categorized based on information compiled by Soma Umenye before implementation using administrative records.

Does Soma Umenye impact students in different provinces differently?

Finally, we attempted to examine the impacts of Soma Umenye disaggregating the sample by Province. Similar to the urban subgroup above, disaggregating the sample into five province subgroups resulted in small samples for each of these subgroups. We were unable to estimate impacts for the provinces separately because the small sample sizes did not allow us sufficient statistical power. However, we estimated whether the impacts on student EGRA scores on different sub-tests differed across provinces and did not find any evidence that it did.³⁴ This result suggests that at least in its first year of implementation, Soma Umenye benefitted all provinces equally in terms of having impacts on students' reading outcomes.

³⁴ Statistical power was not an issue for this test as we could use the entire sample to estimate whether impacts differed across region.

8. CONCLUSIONS

Because we followed a randomized controlled trial (RCT) framework, we can be certain that our comparison group is an accurate representation of the counterfactual. Nonetheless, in interpreting the findings reported in Section 7, we should bear in mind that during 2017 Soma Umenye did not provide its complete package of reading interventions as described in Figure 2.1. This said, the findings presented in Section 7 allow the following conclusions:

- Participation in the Soma Umenye activity caused students to improve their basic reading skills, in particular in the letter and syllable recognition and familiar word reading sub-tests. However, the results of the EGRA reveal that by the end of PI, the great majority of the 5,466 students assessed in the 304 randomly selected schools have not yet acquired the lower-order reading skills (knowledge of letters, syllables, and reading of familiar words) on which higher-order reading skills such as fluency and reading comprehension are based. This is confirmed by the finding that only 3 percent (169) of the 5,457 students randomly selected for assessment showed ability to read the 38-word story and score 80 percent or higher, i.e., able to answer correctly four of the five questions related to the story.
- Given students' difficulties in identifying letters and syllable sounds, the low scores observed in the decoding words sub-test are not surprising. Presented with a chart of 50 familiar words, on the average, students in both treatment and control groups could read fewer than five words per minute.
- Evidence from EGRA studies worldwide confirms that there is a strong predictive relationship between the early mastery of pre- or foundational reading skills, reading fluency, and reading comprehension. The average score obtained on the Reading Comprehension sub-test was 0.69 for the treatment and 0.60 for the control group.³⁵ The extremely low scores can be directly linked to the lack of oral fluency.
- Male and female students in treatment schools benefitted equally from the Soma Umenye activity and both groups scored higher than their counterparts in the control schools.
- The Listening sub-test provides information about what students are able to comprehend without the challenge of decoding a text. The large concentration of scores near the upper limit in the listening comprehension sub-test (average score of 3.91 out of a possible 4.0) show that PI students who have not yet learned to decode have oral language, vocabulary, and comprehension skills that allow them to answer the questions correctly. At the same time, an examination of the score distribution suggests that the **Listening Comprehension** sub-test may have been too easy. If an EGRA sub-test is much too easy for most students, the scores

³⁵ Difference not statistically significant

will concentrate heavily at the upper end of the allowable range and produce ceiling effects, restricting the variation of scores and negatively impacting the validity of the tool itself.³⁶

- Soma Umenye's impact on students who self-reported repeating PI were not statistically different from impacts on non-repeaters on any of the EGRA sub-tests. This suggests that Soma Umenye does have positive additive value even for students who have already had prior experience with PI. The finding that both repeaters and non-repeaters in the treatment group scored better in the **Letter Identification** sub-test compared to repeaters and non-repeaters in the control group, confirms that the intervention affected both groups equally so far.
- The low scores on the EGRA sub-tests that measure pre-reading or emergent reading skills (on the average 20 letters recognized out of 100; 16 syllables out of 100; and fewer than five familiar words out of 50) suggest that students are not receiving sufficient focused classroom instruction and practice to master these skills and become fluent readers by the end of grade P3.³⁷ Taken together, the findings suggest that reading instruction in the early grades is not focusing on helping students to learn the sounds associated with each letter and applying this knowledge to sound out familiar (or unfamiliar) words.
- Given the novelty of the instructional methodology, the relatively short period of training received by teachers, the absence of training materials, and the significance of the statistical difference between teachers in treatment and in control schools, it is encouraging that the instructional behavior of the 129 teachers observed who received Soma Umenye training was half a standard deviation better compared to the 152 control group teachers.³⁸
- The high rate of absenteeism is cause for concern. Over half of the PI students assessed self-reported having been absent for at least one day the previous week – and by actual count against the class rosters, over 20 percent of PI students enrolled were absent the day that their class was assessed.
- Our findings—differences in performance observed between students in treatment and control groups—in 2017 suggest that students who participate in Soma Umenye have a better chance to acquire the reading skills essential for learning to read with comprehension. However, the low intensity of inputs (approximately seven months of exposure to SU, in the start-up year of the activity and without supporting materials) was not sufficient to raise overall means to near acceptable levels. As the Soma Umenye intervention becomes more intensive in

³⁶ EGRA Toolkit (2nd edition, March 2016)

³⁷ Fluent readers as defined by the EGRA Toolkit (2nd edition, March 2016) “Students are considered fluent readers if they read the entire passage in one minute or less and can answer 80 percent of the reading comprehension questions correctly.” (3.2 out of 4 questions in this particular case).

³⁸ Measured by the composite score on five competencies: lesson preparation, teaching methodology, literacy technical skills, assessment techniques, and use of teaching materials.

years 2 – 4, we can hypothesize that more students will improve their reading skills. However, without an impact evaluation and given all the different activities implemented by a number of donors, it will not be possible to say that Soma Umenye caused the improvement.

- Finally, impact evaluations are usually of specific interventions in a specific context. It is not necessarily the case that the findings of the USAID Soma Umenye program in Rwanda reported here can be generalized to a replicated intervention in different contexts.

9. RECOMMENDATIONS

In Section 9 we follow the conclusions presented in the previous section with a set of recommendations aimed at enhancing the impact of Soma Umenye.

- **In the early grades (P1 and P2), focus on pre-reading skills.** Although reading comprehension strategies are important at all ages, the instructional focus, especially at P1, must be on the foundational skills of phonological awareness and alphabetic principles, coupled with practice in applying these skills to the decoding of syllables and words. Regular and ample practice will increase the recognition of new words and this will translate into increased fluency, which in turn, will increase students' ability to comprehend what they read. This focus has the potential to reduce the high levels of repetition, especially in P1 (between 25 and 27 percent).
- **Revisit the EGRA tool used to assess students' Listening Comprehension skills.** It is possible that this EGRA tool has produced ceiling effects and if so, it should be revised to increase its validity.
- **Focus on early identification of slow or immature learners and provide extra assistance to avoid leaving students behind.** While important for all under-performing students, this can be especially important for children who are repeating in order to keep them from becoming discouraged to the point of dropping out.
- **Monitor the Fidelity of Implementation.** In subsequent years it is important that Soma Umenye implementation be carried out in substantially the same way across all schools to expand the Soma Umenye impact observed. The level of possible impacts may be affected if the planned implementation package is not fully deployed in subsequent years.
- **Provide incentives for teachers to disseminate the training received.** It is often the case that teachers receive training and do not continue teaching the same grade or that they transfer to another school. One way to ensure that training is not "wasted" is to provide teachers with opportunities to disseminate the training to their colleagues. This should be part of the training Soma Umenye will provide to Head Teachers.
- **Focus on pre-service training and on what future teachers are learning and practicing while still at teachers' preparation colleges.** Results from the EGRA can inform the design of both teacher training and especially of pre-service and professional development programs. In-service training alone, as provided by Soma Umenye, is only part of the answer to better teaching. In-service training is an expensive model and is not sustainable in the long term. Teacher preparation (pre-service) needs to focus on the development of early grade-specific skills in teaching reading.

- **The development of benchmarks and reading standards for grades P1 – P3 should be a priority.** Use EGRA findings to refine benchmarks or standards for reading for each grade. Once they are developed there should be ample dissemination so that teachers, head teachers, and parents can focus their efforts on making sure that P1 students are able to read X words per minute; P2 able to read X+ words; and, P3 X++ words. The standards could be instrumental in assisting educators to address the teaching of reading in a concrete manner and with a specific objective.

SELECTED REFERENCES

United States Agency for International Development

- Automated Directives System (ADS) 200. Development Policy
- Protection of Human Subjects in Research Supported by USAID: A Mandatory Reference for ADS Chapter 200, 200mbe_122606_cd46
- ADS 201. Program Cycle Operational Policy
- Conducting Mixed Methods Evaluations. Monitoring and Evaluation Series Technical Note. June 2013.
- Early Grade Reading Assessment (EGRA) Toolkit, Second Edition. 2015
- Impact Evaluations. Monitoring and Evaluation Series Technical Note. September 2013
- USAID Education Strategy 2011- 2015, Extended to December 2017
- Implementation Guidance, Revised April 2012
- Technical Notes
- Reference Materials
- Update to Education Strategy Reporting Guidance (Issued August 2014)
- USAID Evaluation Policy, Updated October 2016
- USAID Scientific Research Policy, December 2014

U.S. Office of Management and Budget. Monitoring and Evaluation Guidelines for Federal Departments and Agencies that Administer United States Foreign Assistance. January 11, 2018

Rwanda Ministry of Education. 2016 Education Statistical Yearbook, Kigali, February 2016

- 2015 Education Statistical Yearbook, Kigali, June 2015
- 2014 Education Statistical Yearbook, Kigali, March 2014
- A study on children with disabilities and their right to education: Republic of Rwanda. 2016
- Integanyanyigisho Y'imibare Y'ikiciro Cya Mbere Cy'amashuri Abanza P1-P3 [Lower Primary Mathematics Syllabus]. Kigali, 2015

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

- (2013). "Literacy for All in 100 days?" Global Partnership for Education. Series on learning No. 7 (2009). "Instructional Time Loss in Developing Countries: Concepts, Measurement, and Implications." World Bank Research Observer. 24 (2): 267-290

Aggarwala, N.K. (2004). Evaluation Report: Quality assessment of primary and middle education in mathematics and science. Retrieved from

http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/Aggarwala_UNDP_Evaluation_Report.pdf. Accessed 2013 June 15.

Bruns, B., Mingat, A., & Rakotomalala, R. (2003). Achieving universal primary education by 2015: A chance for every child. Washington, D.C: The World Bank. doi:10.1596/0-8213-5345-4

Chemonics (2017a). Soma Umenye Activity: Year I Work Plan.

- (2017b). USAID Soma Umenye Project: Annual Program Report. Year I: July 20, 2016-September 30, 2017

De Stefano, J. et al. (2012) Early Grade Reading and Mathematics in Rwanda: Final Report. RTI International

- Education Development Center. (November 2014). USAID Literacy, Language and Learning Initiative (L3): National Fluency and Mathematics Assessment Baseline Report
- (January 2017). USAID L3: National Fluency and Mathematics Assessment of Rwandan Schools: Endline Report.
- Evans, D (2016). That zero effect may not mean what you think it means, and other lessons from recent educational research. Impact Evaluations, 1/21/2016 (<http://blogs.worldbank.org/impactevaluations>)
- Friedlander, E. & Goldenberg, C. (eds.). (2016). Literacy Boost in Rwanda: Impact Evaluation of a 2-year Randomized Control Trial. Stanford, CA: Stanford University (with Save the Children)
- Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. 2016. Impact Evaluation in Practice, second edition. Washington, DC: Inter-American Development Bank and World Bank
- Glewwe, Paul and Karthik Muralidharan (2015). Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications. RISE (Research on Improving Systems of Education) Working Paper October 2015
- Gove, A. and P. Cvelich (2011). Early Reading: Igniting Education for All. A report by the Early Grade Learning Community of Practice. Revised Edition. Research Triangle Park, NC: Research Triangle Institute www.eddataglobal.org
- IPSOS. Baseline Study: Home Grown School Feeding Program, 2016-2020. July 22, 2016
- Johnson K, Hayes C, Center H, Daley C. (2004). Building capacity and sustainable prevention innovations: A sustainability planning model. Evaluation and Program Planning. 27:135–149.
- Lance, P. and A. Hattori. (2016). Sampling and evaluation: A guide to sampling for program impact evaluation. Chapel Hill, North Carolina: MEASURE Evaluation, University of North Carolina.
- Laterite. System Dynamics in Primary and Secondary Education in Rwanda: Draft Inception Report, August 17, 2016
- Malik, Saima S., Beatrice J.V. Balfour, Jean Providence Nzabonimpa, Sofia Cozzolino, Gabriela Dib, and Amy Jo Dowd. (2015). Endline Evaluation of Rwandan Children's Book Initiative (RCBI), Save the Children.
- National Council of Persons with Disabilities and National Commission for Children. (2016). Report on National Assessment of Centres Caring for Children with Disabilities in Rwanda. Kigali.
- Piper, Benjamin, Stephanie S. Zuilkowski, and Salome Ong'ele, (2016). "Implementing Mother Tongue Instruction in the Real World: Results from a Medium-Scale Randomized Controlled Trial in Kenya," Comparative Education Review 60, no. 4: 776-807.
- Schochet, Peter. (2009) "An Approach for Addressing the Multiple-Testing Problem in Social Policy Impact Evaluations." Evaluation Review, vol. 33, no. 6:539–567.
- Spaull, N. and Taylor, S. (2015) "Access to What? Creating a Composite Measure of Educational Quantity and Educational Quality for 11 African Countries," Comparative Education Review 59:133-165
- World Bank. Myanmar Early Grade Reading Assessment for the Yangon Region. June 4, 2015
- World Food Programme et al., (2016). Rwanda 2015: Comprehensive food security and vulnerability analysis.

ANNEXES

Annex A.	Current Scope of Work
Annex B.	Regulatory Certifications
Annex C.	Evaluation Design and Year I Workplan
Annex D.	Conflict of Interest Certifications
Annex E.	Districts and Sectors by Phase
Annex F.	Supervisor Manual
Annex G.	The EGRA Instrument
Annex H.	Student Context Form
Annex I.	Lesson Observation Form
Annex J.	Additional Results
Annex K.	Statement of Objections – Does not apply

ANNEX A. CURRENT SCOPE OF WORK

- 2) Section C., "DESCRIPTION/SPECIFICATIONS/STATEMENT OF WORK", Delete the section in its entirety and replace it with the following:

C.1 OBJECTIVES

The United States Agency for International Development in Rwanda (USAID/Rwanda) requires technical assistance to conceptualize and conduct an Impact Evaluation (IE) of the planned *Soma Umenye* activity. The IE will evaluate the impact of the *Soma Umenye* activity, specifically the extent to which *Soma Umenye's* interventions have improved classroom instruction, improved early grade reading outcomes and led to success in schooling, as well as to evaluate and provide recommendations to improve the success and sustainability of programming. The chosen name of the impact evaluation activity is *Soma Umenye* IE.

The results of the IE will be used by USAID, the Government of Rwanda, other development organizations, policy makers, and local stakeholders to inform and advocate for policy and future program design in the education sector, both in Rwanda and in other countries working to improve student reading outcomes, during and beyond the life of the *Soma Umenye* activity. The IE will also strengthen understanding and evidence of the link between improved reading and higher level goals of the USAID/Rwanda Country Development Cooperation Strategy (i.e., success in

schooling) in the Rwandan context to inform broader USAID development strategies. In addition, the IE contractor will provide recommendations to USAID/Rwanda to improve current programming based on the evaluation analysis, findings, and conclusions.

C.2 BACKGROUND

Description of the Problem, Theory of Change, and Development Hypothesis

Early grade reading is the foundation for all other learning and is a major determinant of whether a child stays in school. Children who do not acquire foundational reading skills in the first three grades are on what the USAID Global Education Strategy calls a “lifetime trajectory of limited education progress and therefore limited economic and developmental opportunity.”¹ Unfortunately, the majority of Rwandan students are not learning to read in the early grades. A 2015 reading assessment conducted by the USAID-funded *Literacy, Language, and Learning* (L3) activity found that 50% of students at the end of grade 1, 25% of students at the end of grade 2, and 18% of students at the end of grade 3 cannot read a single word of grade-level text in Kinyarwanda.²

Though the gap is large between the present early grade learning outcomes and the outcomes necessary for Rwanda to achieve its ambitious goals for human capital development, there is strong evidence for what works to improve early grade reading outcomes. While many student-level and home environment factors are strongly correlated with reading outcomes, reading research has concluded that many reading problems can be resolved by providing effective instruction and intervention in the early grades of school. The global evidence for what works in reading can be summarized by five pillars - Teaching, Time, Text, Tongue, and Testing, commonly referred to as the “Five Ts of Effective Reading Instruction.” Students have the best chance of learning how to read when, at a minimum, the following conditions are provided:

TIME: There must be sufficient time (at least an hour a day) in the school day for children to practice reading.

TONGUE: Reading instruction must take place in a language (tongue) that children already speak and understand.

TEXTS: Fiction and nonfiction texts designed for teaching reading must be available to children in that language.

TEACHING: Teachers must know the component skills of reading, how to teach them, and how to measure their students’ mastery of those skills; for this, they need training, coaching, and support.

TESTING: Actors at the central, decentralized, and school levels of the school systems must be able to measure, for both formative and summative purposes, children’s reading abilities using metrics that conform to international standards.

USAID/Rwanda’s development hypothesis is that if the *Soma Umenye* activity improves these conditions in non-private primary schools across

Rwanda, then reading outcomes will improve for students in these schools, Grades 1-3. Further, if literacy outcomes for children in the primary grades improve, then opportunities for Rwandan children to succeed in schooling will be increased.

Specifically, the *Soma Umenye* activity is expected to achieve the two following goals:

To improve reading outcomes in Kinyarwanda for at least 1 million children (unique direct beneficiaries) in public and government-aided schools in Rwanda by the end of Grade 3. One million children with improved reading skills is 60% of the children estimated to receive the *Soma Umenye* intervention during the life of the activity. (Attachment 8 provides more details about student enrollment in grades one through three.)⁴

To ensure that at least 65% of P1-P3 students are able to read grade-level text with fluency and comprehension.

USAID/Rwanda anticipates awarding the *Soma Umenye* activity before August 2016. It is anticipated that *Soma Umenye* will begin to conduct its baseline in January 2017, and it will begin its programming in Rwandan primary schools in the 2017 school year.⁵

Related Issues in the Education Sector

The following section highlights salient issues in the education sector that intersect with the *Soma Umenye* activity and are of interest to USAID, as well as issues that should be taken into account in the impact evaluation as key control variables.

Policy-level Constraints: In the Rwandan education system, Kinyarwanda is the language of instruction in Grades 1-3. From Grade 4, the language of instruction is English. The Government of Rwanda (GoR) and development partners have highlighted challenges with English as the language of instruction in Rwanda as the priority barrier to improved student learning outcomes. As a result, attention and continuous professional development for teachers has been primarily focused on the English-language skills of teachers.

However, USAID/Rwanda hypothesizes that it is a foundational problem that Rwanda's education system does not effectively ensure that all children learn to read in Kinyarwanda in the early grades. USAID/Rwanda and the *Soma Umenye* activity will advocate to the Government of Rwanda to prioritize attention, time, and investment to improve reading outcomes in Kinyarwanda in order to establish this cornerstone as a critical step toward improving other quality of education challenges, including the inefficiency of the system and performance in other subject areas. The IE will evaluate the extent to which improved reading in Kinyarwanda contributes to improved performance in other subjects in Rwanda. This information will be used to inform advocacy about the prioritization of investments in the education sector.

Academic Persistence - Dropout and Repetition: Reducing the high and rising rates of dropout and repetition is a priority for the GoR. In 2013, the primary school dropout rate was 14.3%, and the repetition rate was 18.3%. As a result, in 2014 the number of students enrolled in Primary 4 was only 52% of the number of students enrolled in Primary 1. In 2012 Rwanda Education Board (REB) personnel carried out a case study on the factors affecting dropout and repetition of students in four districts, and the study suggested that quality of education affected repetition. Among the proposed causes of repetition were unqualified or unmotivated teachers, large class sizes, and poor student performance. In turn, repetition was suggested to be a factor influencing dropout. (Other factors influencing dropout included poverty, pregnancy and early marriage, and parental attitudes.)⁷ To expand on these findings, UNICEF and MINEDUC have commissioned a study to identify age- and grade-specific dropout rates over the period 2009-2013. The study will assess the causes of dropout and repetition, focusing on push, pull, and contextual factors which influence the process of repetition and dropout by age, gender, location, socio-economic status, and district. It is expected that findings will be available in late 2016.

USAID/Rwanda hypothesizes that since the low quality of instruction and the resulting poor student performance drive dropout and repetition, improved instruction and improved reading outcomes will reduce dropout and repetition. The advocacy of USAID/Rwanda and the *Soma Umenye* activity to GoR will focus attention and investment on improving reading outcomes in Kinyarwanda as an effective strategy for reducing dropout and repetition.

Student Attendance: Absenteeism affects the already limited opportunities that students have to learn. A 2012 Early Grade Reading Assessment (EGRA) noted that 20 percent of students reported having been absent the previous week.⁸ USAID/Rwanda hypothesizes that improved student attendance will increase reading outcomes. At the same time, implementation of *Soma Umenye* (including the provision of materials and improved classroom instruction) is expected to increase attendance. In addition, improved student reading outcomes are expected to further increase student attendance. USAID/Rwanda is interested in analysis of the interrelationship between the implementation of *Soma Umenye*, improved classroom instruction, improved reading outcomes, and student attendance.

Nutrition: The health and nutritional status of children, as well as simply whether or not they are hungry during their lesson, affects their learning outcomes. USAID/Rwanda is interested in understanding how the cross-sectoral approach of a nutrition intervention at the primary school level affects learning outcomes. The relationship between school feeding and attendance is also critical to analyze because improved attendance is hypothesized to be a potential key driver of improved learning outcomes.

Students with Disabilities in Rwanda: As part of the national education data collection, schools are required to report annually on the number of enrolled students with disabilities. The guidance given to schools for the purpose of identifying children with disabilities is similar to the disability measure used in the 2012 Census. It is based on the International Classification of Functioning, Disability and Health (ICF), and it uses the concept of functional ability (such as difficulty seeing, hearing, speaking, walking/climbing and learning/concentrating). The Ministry of Education's Annual Statistical Yearbook 2014 indicates that students with disabilities represent only 0.8% of the total enrolled in primary education, and data about disability disaggregated below the national level is not available. This small percentage raises concerns about the under-identification of enrolled students with disabilities, as well as about the lack of enrollment of students with disabilities. Despite this complication, USAID/Rwanda is interested in understanding the effect of *Soma Umenye* on students with physical disabilities (such as hearing, visual, and speaking) as well as students with learning disabilities.

UNICEF has conducted a qualitative situational analysis of students with disabilities in Rwandan schools that includes an analysis of factors that promote and hinder the academic success of students with disabilities. It is expected that this UNICEF study will be disseminated during 2016.

Other Activities in the Education Sector

The following are some of *Soma Umenye*'s complementary activities and programs in the education sector. Awareness of these activities is important because the IE Contractor is required to determine the net effect of the *Soma Umenye* activity through the estimation of the counterfactual, controlling for the effects of other programs and other factors. In addition, USAID is interested in better understanding complementarity of programming with respect to improving early grade reading outcomes.

Literacy, Language and Learning (L3): The *Soma Umenye* activity will build on one of USAID/Rwanda's current investments, *Literacy, Language, and Learning (L3)*, implemented by Education Development Center from August 2011 to January 2017. L3 has aimed to improve early grade literacy and numeracy outcomes by supporting teacher training, developing and distributing instructional materials for Kinyarwanda and English (grades 1-4), and engaging communities. L3 is present in all public schools nationwide. Table 1 shows data from L3's January 2016 reading assessment midline report, which indicate that gains have been made to the following magnitudes:

Table 1 - L3 Baseline to Midline Gains

Grade	Subtest	BASELINE	MIDLINE	GAIN	EFFECT SIZE (Cohen's d)
P1	Percent	17.2%	25.9%	8.60 (± 3.48)	0.28 (± 0.11)
	Words Per Minute	4.76	7.48	2.70 (±1.06)	0.29 (± 0.11)
	Zero Score	60.3%	50.4%	-10.10 (±0.06)	0.20 (± 0.11)
P2	Percent	43.1%	50.5%	7.40 (± 4.41)	0.19 (± 0.11)
	Words Per Minute	19.2	21.5	2.30 (± 2.00)	0.13 (± 0.11)
	Zero Score	32.7%	25.5%	-7.20 (±0.16)	0.11 (± 0.11)
P3	Percent	37.5%	44.7%	7.20%	0.26 (± 0.11)
	Words Per Minute	22.1	25.1	2.99 (±1.79)	0.19 (± 0.11)
	Zero Score	21.3%	18.6%	-2.71 (±0.05)	0.07 (± 0.11)

Other results indicate that there was an increase by 10% in non-zero scores on the oral reading fluency assessment in P1. Notable results for P2 showed an increase of 6.3% on the proportion of students reading at least 20 correct words per minute. In P3, there was an increase of 7.2% in the proportion of students reading at least 33 words correct per minute.

Mureke Dusome Activity: *Mureke Dusome* ("Let's Read") is a USAID-funded activity implemented by Save the Children running from January 2016 to January 2020. The objective of the activity is to strengthen partnerships between schools and the broader community to support the development of student reading skills in the early grades throughout Rwanda. *Mureke Dusome's* approach is: 1) to strengthen the capacity of school leadership to improve student literacy through school-community partnerships, 2) to increase effective community and parental involvement and improve literacy skills, and 3) to foster a culture of reading. *Mureke Dusome* will phase in some activities by district, but will have nationwide reach. Activities include the training of district and sector education officials, the capacity-building of Head Teachers and School General Assembly Committees, the provision of literacy promotion and practice activities outside of school, as well as national sensitization and mobilization to support children's reading. USAID is interested in whether participation in *Mureke Dusome* activities is associated with higher reading outcomes.

Mureke Dusome will be nationwide by its final year of implementation, but it will be phased in by district.

McGovern-Dole International Food for Education Child Nutrition Program: The McGovern-Dole Food for Education program is a 5-year award by the United States Department of Agriculture (USDA) McGovern-Dole International Food for Education and Child Nutrition Program to the World Food Program (WFP) Rwanda to support the school feeding program in four districts: Karongi, Nyamagabe, Nyaruguru and Rutsiro. The dates

of implementation of the school feeding program are August 2015-2020. As the *Soma Umenye* activity will have nationwide reach, its programming will overlap with the districts targeted by the Food for Education program.

DFID's Learning for All Program: The *Soma Umenye* activity will be implemented alongside and in coordination with other initiatives in the education sector, including those implemented with the support of the UK Department for International Development (DfID). DfID released a solicitation in October 2015 requesting proposals to provide technical assistance for quality improvement and school level accountability, a component of DfID's four-year Learning for All program. DfID expects this program to be implemented June 2016-May 2019. In targeted districts, DfID's program aims to: increase the percentage of children reaching minimum learning standards in numeracy and literacy at primary level; increase the number of children who can read English with sufficient fluency for comprehension by the end of grade 3; increase the average attendance rate of primary boys and girls; and reduce dropout in grade 5 for primary boys and girls.

Government of Rwanda Initiatives: The *Soma Umenye* activity will also be implemented alongside national initiatives of the Government of Rwanda including the roll-out of a new, competency-based curriculum starting in 2016. To support the new curriculum, the GoR is training teachers, as well as English-language School-Based Mentors. More information about Government of Rwanda initiatives can be found on the websites of the Ministry of Education and Rwanda Education Board.

Results Framework

The *Soma Umenye* activity is designed to achieve the following results:

Development Objective: Increased opportunities for Rwandan children and youth to succeed in schooling and the modern workplace

Program Purpose: Improved literacy outcomes for children in early grades	
Required Outcome-Level Indicator, FAF 3.2.1-27: The proportion of students who, by the end of two years of schooling, can read and understand grade-level text	
IR 1: Classroom instruction in early grade reading improved	IR 2: Systemic capacity for early grade reading instruction improved
Sub-IR 1.1: Evidence-based, gender-	Sub-IR 2.1: National advocacy mechanisms

sensitive early grade reading materials available and used	for early grade reading interventions strengthened
Sub-IR 1.2: Teachers' use of evidence-based, gender-sensitive instructional practices in early grade reading increased	Sub-IR 2.2: Student and teacher performance standards and benchmarks for early grade reading applied
Sub-IR 1.3: Capacity of head and mentor teachers to coach and supervise early grade reading instruction strengthened	Sub-IR 2.3: Research-based policies and curricula in support of early grade reading instruction implemented
Sub-IR 1.4: Schools' and teachers' use of student assessment results improved	Sub-IR 2.4: Early grade reading assessment systems strengthened
	Sub-IR 2.5: Capacity of Teacher Training Colleges to prepare effective early grade reading teachers improved
Cross-Cutting: Gender and inclusion of students with special needs, ICT	

Summary of the Activity to be evaluated

The goal of the *Soma Umenye* activity is to improve early grade reading outcomes for at least one million children nationwide. *Soma Umenye* will catalyze evidence-based early grade reading instruction in Kinyarwanda in grades 1-3 and ensure that a majority of students can read in Kinyarwanda with fluency and comprehension by the end of second grade. In keeping with its clear focus on improving early grade reading outcomes in Kinyarwanda, *Soma Umenye* has just two intermediate results: (1) classroom instruction in early grade reading improved, and (2) systemic capacity for early grade reading instruction improved. Activities will include the provision of materials and the training and mentoring of teachers and school leaders on the use of those materials and assessments. In addition, *Soma Umenye* will develop standards and benchmarks, advocate for and support the implementation of improved policy and curriculum, strengthen national assessment systems for early grade reading, and build the capacity of pre-service Teacher Training Colleges to prepare teachers of early grade reading.

Soma Umenye will have nationwide reach, and it will target as direct, primary beneficiaries all students in public and government-aided P1-P3¹² schools in Rwanda.¹³

The *Soma Umenye* Statement of Work (see Attachment 9) serves as the primary technical reference for the impact evaluation. Upon award, the

Soma Umenye implementer and the impact evaluation contractor must collaborate closely and directly with each other, as well as with the Ministry of Education and USAID, to finalize the evaluation design, IE data collection systems, and the selection procedures required for an IE.

Summary of the Activity MEL Plan

The *Soma Umenye* implementer is required to conduct monitoring, evaluation and learning (M&E). This ME will permit the efficient and timely measurement of progress toward outcome and output indicators, aid in the management of *Soma Umenye* activities, and provide other data as agreed with USAID/Rwanda and the implementer of *Soma Umenye*, consistent with USAID's performance monitoring, evaluation and learning policies. In collaboration with the Government of Rwanda, the implementer of *Soma Umenye* will, at a minimum, conduct assessments of teaching practice and early grade reading fluency and comprehension at baseline, midline, and endline of the *Soma Umenye* activity in a representative sample of treatment schools. Additional data collection will be negotiated as part of *Soma Umenye*'s monitoring, evaluation and learning plan. The reading assessment tools used by *Soma Umenye* and the IE Contractor must be complementary, and the early grade reading assessment (EGRA) of *Soma Umenye* and the IE must conform to the guidance in the USAID EGRA Toolkit.¹⁴ The EGRA conducted by the activity will be representative to a sub-national level (to be determined based on the cost-reasonableness analysis of the activity implementer). Data collected by the *Soma Umenye* implementer, including performance monitoring indicators and data and evaluation reports, will be shared with the impact evaluation contractor when they become available.

Illustrative indicators for internal monitoring of the *Soma Umenye* activity include:

- Proportion of students who, by the end of two grades of primary schooling, demonstrate that they can read and understand the meaning of grade level text
- Proportion of teachers implementing evidence-based early grade reading instruction
- Proportion of grade 1-3 teachers of early grade reading successfully trained
- Number of teaching and learning materials for early grade reading provided
- Number of standardized learning assessments supported
- Number of laws, policies, regulations, or guidelines developed or modified to improve primary grade reading programs

Analysis of these results must be disaggregated by gender.

Coordination

The IE Contractor must collaborate effectively with the *Soma Umenye* activity implementer so that neither the evaluation design nor the activity implementations are compromised. The *Soma Umenye* Contractor will have a responsibility and obligation to maintain fidelity to the agreed-upon implementation plan, and the IE Contractor will have a responsibility and obligation to be flexible in the methodology and design, where possible, without compromising the rigor of the impact evaluation design and methodology.

The IE Contractor must also discuss evaluation plans with other key implementers in the sector, including but not limited to DfID's Learning for All program and evaluation and the McGovern- Dole Food for Education Program and evaluation. Discussion is expected to strategize and promote complementarity of evaluations, as well as to coordinate technical details that will facilitate evaluation.

Coordination with the relevant Government of Rwanda agencies is crucial to ensure acceptance by Government of Rwanda officials of the validity of the evaluation's methodology and results. The impact evaluation Contractor must engage with relevant Government of Rwanda staff (such as the Rwanda Education Board) regarding instrument development, enumerator trainings, data collection, dissemination presentations, and other activities or discussions determined by the IE Contractor and USAID to be relevant to promoting Government of Rwanda collaboration, understanding and buy-in with the evaluation.

C.3. GENERAL PROGRAM PARAMETERS

The following guidance is provided with respect to issues such as geographic focus, evaluation questions and design, and personnel requirements.

Target Beneficiaries and Geographic Scope

The *Soma Umenye* activity will be implemented in all public and government-supported primary schools by the end of the activity period. Thus, to measure the impact of *Soma Umenye*, the IE Contractor is expected to be national in its scope in order to capture a representative sample of students, teachers, and schools.

Evaluation Questions

USAID/Rwanda has identified the following evaluation questions:

To what extent are changes in Kinyarwanda reading outcomes for students in Grades 1 attributable to the *Soma Umenye* activity (as a package)?

The purpose of the *Soma Umenye* activity is to improve reading outcomes for students in the early grades. This question will evaluate the extent to which *Soma Umenye* has achieved this goal. For the purposes of this Impact Evaluation, USAID/Rwanda is interested in the impact of *Soma Umenye* on the following measures of student reading outcomes:

- Proportion of students (grade 1) who cannot read a single word of grade level text
- Proportion of students who have mastered sub-skills of reading (including at least letter sound identification and non-word reading)
- Proportion of students (grade 1) who can read and understand grade-level text. This indicator is expected to include assessment of both oral reading fluency (as measured by correct words read per minute) and comprehension.

In addition, *Soma Umenye* seeks to improve the reading skills of all students in lower primary school, including vulnerable student populations. Because of the importance of equity, USAID/Rwanda is keen to identify whether or not the *Soma Umenye* activity differentially affects the reading outcomes of girls and boys, children living in urban and rural areas, and children representing different levels of socio-economic status.

Moreover, USAID/Rwanda needs to understand factors that influence reading outcomes. At a minimum, analysis of student reading outcomes must include the following factors:

Student attendance (as described in the background section)
Influence of the USAID-funded *Mureke Dusome* activity (see background section)

The components of the *Soma Umenye* activity are highly interrelated and the evidence suggests that all are necessary to promote reading achievement. Thus, it is not expected that it will be effective to implement a particular component in isolation from the others, or that the effectiveness of a particular component can be evaluated.

For this reason, the impact evaluator is asked to consider *Soma Umenye* as a package, rather than as component activities.

To what extent has classroom instruction changed as a result of *Soma Umenye*?

For the purposes of the baseline, classroom instruction is defined as the degree to which a teacher of grade 1 is observed implementing key evidence-based pedagogical strategies in Kinyarwanda in the technical approach of the *Soma Umenye* activity. USAID/Rwanda also needs to understand whether the *Soma Umenye* activity has differentially impacted the classroom instruction of male and female teachers. USAID also seeks recommendations to increase teachers' application of the *Soma Umenye* pedagogical approach, including analysis of the challenges teachers still face in changing their teaching practice, and factors that distinguish teachers who successfully apply the instructional principles and practices of *Soma Umenye* from those who do not.

What is the evidence that improved student reading outcomes in Kinyarwanda lead to improved opportunities for students to succeed in schooling?

For the purposes of the baseline, USAID/Rwanda has defined the key indicator of "increased opportunities for Rwandan children to succeed in schooling" as reduced rates of dropout and repetition, as well as improved learning outcomes.

USAID is interested in learning whether or not *Soma Umenye* impacts school repetition and dropout. Through the Education Management Information System (EMIS), the Ministry of Education collects school dropout and repetition data. The data are included in the annual Education Statistical Yearbook and are discussed at bi-annual Joint Education Sector Reviews. However, the quality of the data is unknown, and the impact evaluation contractor must collect independent data about dropout and repetition.

Furthermore, *Soma Umenye* seeks to improve the reading skills of all students in lower primary school in order to increase the opportunity for all Rwandan children to succeed in school.

Because of the importance of equity, USAID/Rwanda is keen to identify whether or not the *Soma Umenye* activity differentially affects the dropout/repetition rates and other learning outcomes for various vulnerable student populations, including girls and boys, students in urban and rural areas, students representing different levels of socio-economic status, and students with disabilities.

Evaluation Design and Methodology

At a minimum, the *Soma Umenye* activity will collect data on student fluency and comprehension at baseline. The *Soma Umenye* activity will also collect data on teaching practice. The impact evaluation contractor will be required to validate and use the performance monitoring data of the *Soma Umenye* implementing partner. The IE Contractor will collect data for the control groups at baseline. The IE's data collection is expected to be separate and independent from the *Soma Umenye* activity but must ensure complementarity between sampling tools, such as EGRA.

The impact evaluation Contractor will also draw on activity quarterly and annual reports, performance reporting, and special purpose publications for a substantial portion of the quantitative data and implementation monitoring data required. At activity start-up, it is expected that the *Soma Umenye* implementer and the impact evaluation contractor will collaborate to develop a mutually supportive and compatible plan for sharing activity (program) monitoring and evaluation output and outcome data.

An impact evaluation design is required to answer the evaluation questions. This design will be supplemented by other methodologies to answer the evaluation questions according to the needs of USAID, as described above (Sections I, "Purpose" and IV "Evaluation Questions"). The design and methodology required for each of the evaluation questions is considered in more detail below.

To what extent are changes in Kinyarwanda reading outcomes for students in Grade 1 attributable to the *Soma Umenye* activity (as a package)?

For this question, a randomized control trial (RCT) is preferred, and it may be possible if *Soma Umenye* plans to phase in implementation. If an RCT isn't possible, the impact evaluation could use a quasi-experimental design. Key outcome data (such as oral reading fluency, reading comprehension, and pre-reading skills that include letter sound recognition and non-word reading at a minimum) will be collected by the implementer of the *Soma Umenye* activity in treatment schools. The impact evaluation contractor's design must include the validation and triangulation of outcome data collected by the activity implementer. If both the activity implementer and the evaluation contractor are assessing oral reading fluency, it is essential that identical or equivalently-leveled passages are used. The design must be able to detect changes in key outcome indicators (such as reading outcomes) at the baseline level determined by the IE

contractor and agreed by USAID based on analysis of the size of expected gains.

The Contractor will conduct primary quantitative and qualitative data collection on subpopulations of interest. Disaggregation and subpopulations of interest related to this evaluation question include: male/female students; urban/peri-urban/rural schools (to capture the influence of relative distance from Kigali¹⁸); socio-economic status groupings; students participating in school-based feeding programs; students exposed to the USAID-funded *Mureke Dusome* activity; students with disabilities; and students entering at different levels (such as non-readers and those at a low benchmark versus more advanced readers).

To what extent has classroom instruction changed as a result of *Soma Umenye*?

In addition to *Soma Umenye*'s data collection on teacher practice, the IE Contractor will collect supplementary data to aid in making recommendations.

What is the evidence that improved student reading outcomes in Kinyarwanda lead to improved opportunities for students to succeed in schooling?

The IE Contractor will collect quantitative data about dropout and repetition.

Evaluation Team Requirements

The Contractor evaluation team must include staff with demonstrated capability and experience in using EGRA methodology, designing impact evaluations, conducting mixed-method and qualitative studies, and conducting advanced statistical analysis. The team should also demonstrate capability for successful team management and logistical coordination required to carry out the above tasks, as well as effective skills to communicate findings and recommendations in both written reports and presentations".

C.4 SERVICES AND TASKS REQUIRED

The Contractor will provide the services and deliverables specified below. Specific services and activities required will be established in the annual work plans approved by the COR and, when necessary, through modifications to work plans approved in writing by the COR. Such technical direction must be consistent with all of the terms and conditions of the Order and the determinations of the Contracting Officer.

Management and Administrative Services

The Contractor will provide all general management and administrative support necessary to perform the Order. The services authorized include, but are not limited to:

- Overall management and administration of the Order, including both expatriate and home office support and administrative services. The Contractor will provide both the key personnel approved by the Contracting Officer and additional personnel, long-term and short-term, necessary to meet recurring general management and administrative support needs under the Order.
- The Order shall procure or lease necessary facilities, supplies and services as necessary to perform the Order.
- The Contractor shall provide the planning necessary for performance of the contract. As discussed in more detail elsewhere, implementation and other plans are required.
- Oversight, quality control and general technical support of all services and deliverables provided pursuant to the Order.
- Provide and assure the proper, efficient and uniform use of modern management and administration, accounting practices, information technology (IT), communications, reporting, human resource management, property control, security, records and other administrative processes and systems required under the Order.
- Manage the overall reporting needs of the program as specified in the Order and developed in work plans in conjunction with the COR.

Technical Services

The Contractor will provide technical services designed to achieve the objectives of the *Soma Umenye* IE. As discussed above, annual work plans and/or modification to the work plans will establish the specific services, activities and deliverables required. In addition, the performance milestones in the work plans will establish the basis on which performance will be evaluated, in addition to the submission of the deliverables stated below.

C.5 PROGRAM REPORTING AND OTHER DELIVERABLES

The required reporting documents and schedule are:

Annual Work Plan: The Contractor will submit for approval annual draft work plans for the evaluation to the Contracting Officer's Representative (COR). The annual work plans will include:

- the anticipated schedule and logistical arrangements for the overall evaluation study plan;
- an anticipated schedule and any logistical arrangements for all data collection activities, analyses, and follow-on study planning (e.g., qualitative studies, smaller scale mixed- methods studies, small-scale cohort studies, and contextual data) based on findings from the baseline;
- a list of the members of the Contractor evaluation team, delineated by roles and responsibilities; and
- an outline of anticipated activities, local travel, and a list of contacts requested from USAID in order to prepare the Evaluation Design and the data collection, analysis, and follow-on supportive study planning. These plans may include qualitative and mixed methods studies that will be designed by the Contractor evaluation team in collaboration with the USAID/Rwanda Education Office and Program Office and will draw on findings from the baseline data analysis. These will be studies designed to address particular questions around sub-population barriers, etc.

Evaluation Design: The Contractor must submit to the COR an evaluation design developed in collaboration with the *Soma Umenye* Contract implementing partner that will describe the methods, sampling, and analysis plans (this will become an annex to the Evaluation report).

- The evaluation design will include: The selection criteria and/or sampling plan. This must include sample size calculations and a justification of sample size, plans as to how the sampling frame will be developed, and a sampling methodology that permits analysis at multiple levels.
- A detailed evaluation design matrix that links the evaluation questions in the SOW to data sources, methods, and the data analysis plan.
- Known limitations to the evaluation design.
- Preliminary data collection plan, with detail on the roll-out and supervision for quality control, and transmission and management of data for all proposed data collection efforts, or a description of how these will

be developed if dependent on findings from the baseline data analysis.

- Plan for data entry and analysis for each data collection effort. The analysis plan could include illustrative versions (i.e., empty shells) of the tables and graphs that will be produced.
- Plan for synthesis (from analysis to findings) and report completion for all data collection efforts.
- Draft survey questionnaires and other data collection instruments, or their main features, along with a detailed plan for how data collection instruments will be developed for all surveys/studies (includes the quantitative surveys of the school children and teacher outcomes of interest), and other key data collection activities, including validation and triangulation studies, as well as proposed anticipated supportive studies that will be developed based on findings from the baseline.
- Methods for selection and training of data collectors (enumerators), data entry personnel, and data analysts.
- Coordination activities (with activity implementer[s], USAID, Government of Rwanda, and others as applicable) related to design, measures, instruments, randomization, etc.
- A description of the process for how informed consent will be obtained, and data collected from study participants and managed in an ethical manner that complies with the Common Rule (Protection of Human Subjects in Research Supported by USAID; A Mandatory Reference for ADS Chapter 200). Also, the design should provide information on how survey review and approval from the National Institute of Statistics of Rwanda will be obtained, if applicable.
- If applicable, a conflict of interest mitigation plan based on the Disclosure of Conflict of Interests submitted with the awardee's proposal.

USAID/Rwanda In-brief: The Contractor will participate in an in-briefing with primary stakeholders in the USAID Program Office, the Education Office, and the Office of Acquisition and Assistance for introductions and to discuss the team's understanding of the assignment, initial assumptions, evaluation questions, methodology, and work plan, and/or to adjust the Statement of Work (SOW), if necessary.

Data Collection Preparation: The preparations for the data collection will include refining the evaluation design, recruiting and training enumerators, pilot testing data collection tools, data entry personnel, and data analysts. It

will also include finalizing logistics associated with data collection. The Contractor can expect that USAID, Government of Rwanda, and/or implementing partner staff may participate in evaluation activities, such as data collection. These participating individuals will complete and sign a conflict of interest disclosure statement.

Interim Update Meetings: During each data collection period, the Contractor will periodically brief USAID/Rwanda stakeholders on the progress of data collection, challenges, and impact on the timeline. If desired or necessary, weekly briefings by phone can be arranged.

Baseline Report: The report will address the findings in relation to each of the questions identified in the SOW and any other issues the team considers to have a bearing on the objectives of the evaluation. Standards for the reports are as follows:

- Report includes background on the local context and the project being evaluated;
- Report includes the main evaluation questions;
- Report clearly explains the methodology and analytic framework for the evaluation approach;
- Report includes an analysis of baseline/midline data;
- Report includes evidence-based recommendations for USAID programming;
- Report proposes follow-on studies that will be useful to inform USAID/Rwanda programming;
- Report includes an executive summary which succinctly lists main findings;
- Report includes at least one 1-page fact sheet with infographics for communications purposes that highlights key data and findings from the report;
- Report's key sections and format follow the evaluation report template;
- Report will be produced in English;
- Report is grammatically correct and contains no spelling or punctuation errors; and

Preliminary Findings Workshops with USAID: The Contractor will hold a preliminary findings workshop for each draft report in-person and/or by virtual conferencing to discuss the summary of findings and any implications for the evaluation study design. An acceptable presentation will include stakeholders from USAID/Rwanda's Program Office and Education Office, the implementing partner staff, and others. A strong presentation

will not exceed 90 minutes. It should include PowerPoint slides, handouts of the slides, other visual media as appropriate, and it should be a collaborative learning and adaptation opportunity.

The annexes to the report will include:

- The Evaluation SOW;
- Any statements of difference regarding significant unresolved differences of opinion by funders, implementers, and/or members of the Contractor's evaluation team;
- All tools used in conducting the evaluation, such as questionnaires, checklists, and discussion guides;
- Sources of information, properly identified and listed; and
- [Disclosure of conflict of interest forms](#) for all evaluation team members, either attesting to a lack of conflicts of interest or describing existing conflicts.

Criteria to Ensure the Quality of the Evaluation Report: Per the USAID Evaluation Policy and USAID ADS 203, draft and final evaluation reports will be evaluated against the following criteria to ensure the quality of the evaluation report:

- The evaluation report should represent a thoughtful, well-researched, and well-organized effort to objectively evaluate what worked in the project, what did not, and why.
- Evaluation reports shall address all evaluation questions included in the SOW.
- The evaluation report should include the SOW as an annex. All modifications to the SOW—whether in technical requirements, evaluation questions, Contractor evaluation team composition, methodology, or timeline—need to be agreed upon in writing by the COR.
- The evaluation methodology shall be explained in detail. All tools used in conducting the evaluation—such as questionnaires, checklists, and discussion guides—will be included in an annex in the report.
- Evaluation findings will assess outcomes and impact on males and females.
- Limitations to the evaluation shall be disclosed in the report, with particular attention to the limitations associated with the evaluation methodology (selection bias, recall bias, unobservable differences between comparator groups, etc.).
- Evaluation findings should be presented as analyzed

facts, evidence, and data and not based on anecdotes, hearsay, or the compilation of people's opinions. Findings should be specific, concise, and supported by strong quantitative or qualitative evidence.

- Sources of information need to be properly identified and listed in an annex.
 - Recommendations need to be supported by a specific set of findings.
 - Recommendations should be action-oriented, practical, and specific, with defined responsibility for the action.
-

ANNEX B. REGULATORY CERTIFICATIONS

REPUBLIC OF RWANDA/REPUBLIQUE DU RWANDA



NATIONAL ETHICS COMMITTEE / COMITE NATIONAL D'ETHIQUE

Telephone: (250) 2 55 10 78 84

E-mail: info@rnecrwanda.org

Web site: www.rnecrwanda.org

Ministry of Health

P.O. Box. 84

Kigali, Rwanda.

FWA Assurance No. 00001973
IRB 00001497 of IORG0001100

August 23, 2017
No.174/RNEC/2017

Edward Jay Allan
Principal Investigator

Your research project: **"Impact Evaluation of the USAID Rwanda Soma Umenye Activity"** has been evaluated by the Rwanda National Ethics committee.

Name	Institute	Involved in the decision		
		Yes	No (Reason)	
			Absent	Withdrawn from the proceeding
Dr.Jean-Baptiste MAZARATI	Biomedical Services (BIOS)	X		
Prof. Eugène RUTEMBESA	University of Rwanda	X		
Dr.Laetitia NYIRAZINYOYE	University of Rwanda		X	
Dr. Egide KAYITARE	University of Rwanda	X		
Sr.Domitilla MUKANTABANA	Kabgayi Nursing and Midwife school	X		
Dr. David K. TUMUSIIME	University of Rwanda	X		
Dr. Lisine TUYISENGE	Kigali Teaching Hospital	X		
Dr. Claude MUVUNYI	Biomedical Services (BIOS)	X		

After reviewing your protocol during the RNEC meeting of July 08, 2017 where quorum was met, and revisions made on the advice of the RNEC submitted on 11 August 2017, we **hereby provide approval for the above mentioned protocol.**

Please note that approval of the protocol and consent form is valid for **12 months**.
You are responsible for fulfilling the following requirements:

1. Changes, amendments, and addenda to the protocol or consent form must be submitted to the committee for review and approval, prior to activation of the changes.
2. Only approved consent forms are to be used in the enrollment of participants
3. All consent forms signed by subjects should be retained on file. The RNEC may conduct audits of all study records, and consent documentation may be part of such audits.
4. A continuing review application must be submitted to the RNEC in a timely fashion and before expiry of this approval.
5. Failure to submit a continuing review application will result in termination of the study.
6. Notify the Rwanda National Ethics committee once the study is finished.

Sincerely,



Date of Approval: August 23, 2017

Expiration date: August 22, 2018

Dr. Jean- Baptiste MAZARATI
Chairperson, Rwanda National Ethics Committee.

C.C.

- Hon. Minister of Health.

- The Permanent Secretary, Ministry of Health

P.O. Vice-chair
Prof. Eugene Rutembeka
Eutembeka

Abarimu

Hifashishijwe impushya zakoreshejwe n' ikigo Save the Children gisuzuma impinduka zabonetse kubera umushinga Literacy Boost!

Izina ry' inyigo: Gahunda y' uRwanda Soma Umenye yatewe inkunga na USAID

INCAMAKE: Utoranyijwe kwitabira ubushakashatsi bujyanye no gufasha guteza imbere kwiga gusoma mu mashuri abanza mu Rwanda. Intego y' ubu bushakashatsi nukurushaho gusobanukirwa uko abarimu mu Rwanda bakorana n' abana mu ishuri babafasha kwiga kumenya gusoma. Uraza gusabwa kuzuza amakuru y' ibanze ajyanye nawe, imyumvire ufite ku myigishirize mu mashuri abanza mugusoma umwe muritwe akurikirana isomo wigisha ikinyarwanda mu ishuri rya P1, P2, cyangwa P3

UMWANYA BITWARA: Kwitabira ubu bushakashatsi biragutwara isomo rimwe n' isaha. Dushobora kuzongera kugusoma umwaka utaha wa 2018 n' ukurikiyeho wa 2019, uramutse udashatse kwitabira ubu bushakashatsi s' itegeko ko ubwitabira, uramutse wemeye ubu, ubutaha ushobora ku mpamvu zawe kureka gukomeza kubwitabira

INGARUKA N' INYUNGU: Nta ngaruka na mba zirimo mu kwitabira ubu bushakashatsi. Nta makuru ajyanye n' ubu bushakashatsi azasanzwe abandi batari mu barimo kubukora, ibyo tunganira bizateranywa n' ibyo bandi barimu bo mu bindi bigo mu karere ku buryo nta Wabasha kumenya ibisubizo byatanze n' umwarimu cyangwa n' ikigo runaka. Inyungu irimo mu kwitabira nuko ubwitabire bwawe muri ubu bushakashatsi buzafasha abanyeshuri bawe na bandi bo mu Rwanda kurushaho kwiga gusoma neza. Nta gihembo cyangwa inyungu twakwizeza kubera ko witabiriyemo ubu bushakashatsi. Uhisemo kwitabira ubu bushakashatsi nta ngaruka na mba wahura nazo

KWISHYURWA: Nta mafaranga ayo ari yose uribubone kuberako witabiriyemo ubu bushakashatsi.

UBURENGANZIRA BW' ABITABIRIYE: Umaze gusoma uru rupapuro, ukemera kwitabira ubu bushakashatsi, menya ko kubwitabira ari amahitamo yawe kandi ko ufite uburenganzira bwo guhitamo kudakomeza kubwitabira igihe icyo aricyo cyose nta ngaruka cyangwa gutakaza amahirwe ayo ari yose wari kuzabona. Andi mahitamo ni kwitabira ubu bushakashatsi. Ufite burenganzira bwo kudasubiza ikibazo icyo aricyo cyose. Ibizava muri ubu bushakashatsi bizashyikirizwa USAID na REB kandi bishobora kugaragazwa mu nama z' inzobere no mu bitabo bijyanye nabyo. Nkuko twavuze haruguru amakuru yose aturutse muri ubu bushakashatsi azateranywa hamwe kuburyo bidashoboka kwitirira cyangwa kumenya umwarimu cyangwa ikigo cya tanze ibisubizo runaka.

NI NDE NAGANIRIZA IBIRENZE BIJYANYE N' UBU BUSHAKASHATSI? Uramutse wifuje gusobanura birenze kuri ubu bushakashatsi cyangwa ufite ikibazo icyo aricyo cyose kuri bwo wakwandikira cyangwa ugahamagara ukuriye iyubahirizwa ry' ibisabwa muri ubu bushakashatsi; Bwana Kenn Ndirangu kuri numero ya telefoni ingendanwa, 0780303164 aguhe ibisobanuro birenze.

Uramutse wifuje gusobanura birenze binjyanye nubu bushakashatsi cyangwa ufite ikibazo icyo aricyo cyose wifuza kuganiriza abandi batakorana muri ubu bushakashatsi wakwandikira cyangwa wahamagara umuyobozi mukuru (Chairman) wa Rwanda National Ethics Commission (RNEC) The chairperson: Dr Mazarati Jean Baptiste kuri telefoni numero 0788309807 cyangwa umunyamabanga Dr. David Tumusiime kuri telefoni numero 0788749398.

Iyi kopi n' iyawe kwibikira.

AMAZINA N' UMUKONO

ITARIKI

Rwanda National Ethics Committee
Approval Date: 23/08/2017
Expiration Date: 22/08/2018

Ku abana bazabazwa

Hifashishijwe impushya zakoreshejwe n' ikigo Save the Children gisuzuma impinduka zabonetse kubera umushinga Literacy Boost!

Izina ry' inyigo: Gahunda y' uRwanda Soma Umenye: Ubushakashatsi busuzuma impinduka kubera umushinga ufasha imyigire mu mashuri abanza

1. Niki kizabaho kubera ubu bushakashatsi?

Intego y' ubu bushakashatsi ni ukurushaho gusobanukirwa uburyo bwafasha abana kwigira gusoma neza kurushaho. Urahabwa isuzuma rireba niba ushobora gusoma inyuguti n' uburyo ubasha gusoma. Iri suzuma riratwara iminota 30. Iri suzuma ryakoreshejwe mu bihugu byinshi by' isi aho ryafashijwe abana benshi mu bihugu byinshi kwigira gusoma neza. Izina ryawe riratorwa ku rutonde rw' abana biga mu ishuli ryanyu, nifuza kuganira nawe hano (mu ishuli rindi ridafite abandi bana biga cyangwa ahandi hatabangamye), ndaza kukubaza ibibazo bike bijyanye no murugo iwanyu, ngusabe gusoma inyuguti, amagambo n'interuro runaka. Ntabwo uri bubone amanota y' ikigikorwa kandi ntukwiye kugira impugenge izo arizo zose uramutse udakoze bimwe muri risuzuma; kuko haribyo tuzi bishobora kuba bigoye. icyo twifuza nukumenya ibyo uzi gusoma. Uramutse udashaka kwitabira iki gikorwa, ntabwo ari itegeko ko ukitabira, wemerewe gusubira mu ishuli ryanyu.

2. Haba hari ikintu kibi cyangeraho?

Nta kintu kibi icyo aricyo cyose cyakugeraho kubera kwitabira ubu bushakashatsi

3. Haba hari ikintu kiza cyangeraho?

Wemeye kudufasha, waba urimo kudufasha kurushaho gusobanukirwa birenzeho uko abana mu Rwanda biga gusoma. Ubu bumenyi bukazakoreshwa mu gufasha abana mu Rwanda kurushaho kwigira gusoma neza..

4. Haba hari abandi bazamenya ko nitabiriye ubu bushakashatsi?

Amakuru yose turibuganireho n'ibanga hagati yawe natwe turimo gukora ubu bushakashatsi. Ntawe tuzaganiriza kubyo twaganiriye, turimo kuganira n'abana benshi baturutse mu bigo byinshi mu Rwanda, ntabwo byashoboka ko undi muntu utari mu barimo gukora ubu bushakashatsi amenya ibyo abana b'ikigo runaka batuganirije.

5. Ni nde naganiriza ibirenze bijyanye n' ubu bushakashatsi?

Uramutse wifuje gusobanuka birenze kuri ubu bushakashatsi cyangwa ufite ikibazo icyo aricyo cyose kuri bwo wakwandikira cyangwa ugahamagara ukuriye iryubahirizwa rw' ibisabwa muri ubu bushakashatsi; Bwana Kenn Ndirangu kuri numero ya telefoni ingendanwa, 0780303164 aguhe ibisobanuro birenze.

Uramutse wifuje gusobanuka birenze binjyanye nubu bushakashatsi cyangwa ufite ikibazo icyo aricyo cyose wifuza kuganiriza abandi batakorwa muri ubu bushakashatsi wowe cyangwa umubyeyi wawe yakwandikira cyangwa yahamagara umuyobozi mukuru (Chairman) wa Rwanda National Ethics Commission (RNEC) The chairperson: Dr Mazarati Jean Baptiste kuri telefoni numero 0788309807 cyangwa umunyamabanga Dr. David Tumusiime kuri telefoni numero 0788749398.

6. Mbye ntashaka kwitabira ubu bushakashatsi?

Uramutse ushatse kwitabira ubu bushakashatsi, wabureka igihe icyo aricyo cyose. Uhisemo kwitabira ubu bushakashatsi nta ngaruka izo arizo zose zakugeraho kubera kutabwitabira

Iyi kopi n' iyawe n'abandi bo murugo iwanyu. Niba wemeye kwitabira ubu bushakashatsi, shyira umukono wawe hano, cyangwa umbwire ko wemeye:

Umukono w' umwana (niba umwana ashoboye gushyira umukonowe: niba atabishoboye andika "yemeye mu magambo")

Andika amazina y' umwana mu cyapa

Itariki y' ikiganiro

Umukono w' ubaza cyangwa ikindi kimuranga

 Rwanda National Ethics Committee
Approval Date: 23.08.2017
Expiration Date: 22.08.2018



Kigali,

N° ~~0145~~.../2017/10/NISR

**NATIONAL INSTITUTE OF
STATISTICS OF RWANDA**

P.O. Box 6139 Kigali

Tel: +250-571035 Fax: +250-570705

E-mail : info@statistics.gov.rw

Edward Jay Allan, PhD
Director,
Monitoring and Evaluation
USAID/Rwanda
KIGALI

Dear Sir,

RE: Visa Approval

Reference is made to your letter dated July 3rd, 2017 requesting for authorization to conduct the study on "*Impact Evaluation of USAID/Rwanda/ Soma Umenye ("Read and Know) Activity"*".

After examining your request and according to the law N° 45/2013 of 16/06/2013, stating on statistical activities organization in Rwanda, we have the pleasure to inform you that the visa has been granted, with the conditions that the data and final report will be submitted to NISR before the publication. In addition, the customised international standard classification manuals (ISIC Rev 4, ISCO 08, and ISCED 97) that NISR has published must be used during the data processing stage to improve the quality of data. Given that the study is targeting the program intervention area, the findings should not be generalized beyond the scope of the program. The NISR would also like to advise you to work closely with the Ministry of Education and Rwanda Education Board for contextual support. However, field work should start after 10th August 2017.

Thank you for your collaboration.

A handwritten signature in blue ink, appearing to read "Yusuf MURANGWA".

Yusuf MURANGWA
Director General



CC: - Ministry of Education
- Rwanda Education Board

MINEDUC Research Clearance Certificate

Application submitted; pending naming of REB Focal Person

Kigali: 8 /08/ 2017
N°: 1866 /REB/2017

Hon. Minister of Education
Ministry of Education
Kigali, Rwanda

Re: Letter of Recommendation for the Impact Evaluation of the USAID/Rwanda Soma Umenye Activity.

Dear Honorable Minister:

This letter is to confirm that REB is collaborating with the Impact Evaluation of the USAID/Rwanda Soma Umenye Activity. USAID/Rwanda has contracted International Business and Technical Consultants, Inc. (IBTCI) to conduct the evaluation.

The purpose of the Soma Umenye Impact Evaluation is to evaluate the impact of the Soma Umenye activity, specifically the extent to which Soma Umenye has improved classroom instruction, improved early-grade reading outcomes and led to success in schooling, as well as to evaluate and provide recommendations to improve the success and sustainability of programming. As the Soma Umenye activity is supporting REB to meet its goals for improved student learning, REB finds this research to be important.

The contact person for the Soma Umenye Impact Evaluation is:

Name and Academic Rank: Edward Jay Allan, Ph.D.

Position: Director, Monitoring & Evaluation, and Project Director, Soma Umenye Impact Evaluation
International Business and Technical Consultants, Inc. (IBTCI)

Phone: +1-617-877-1619

Email: eallan@ibtci.com

Yours sincerely,



GASANA I. Janvier
Director General



CC:

- Hon. Minister of State in Charge of Primary and Secondary Education
- Director General of Science, Technology, and Research, MINEDUC

ANNEX C. EVALUATION DESIGN AND YEAR I WORKPLAN



USAID | **RWANDA**
FROM THE AMERICAN PEOPLE

SOMA UMENYE IMPACT EVALUATION

AID-696-TO-16-00002

YEAR I WORK PLAN AND EVALUATION DESIGN

SEPTEMBER 30, 2016 – SEPTEMBER 30, 2017

Originally Submitted: December 14, 2016

Approved: August 3, 2017

This Work Plan and Evaluation Design was prepared for USAID/Rwanda by International Business & Technical Consultants, Inc. (IBTCI) under Task Order AID-696-TO-16-00002. The views expressed in this document do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

SOMA UMENYE IMPACT EVALUATION

AID-696-TO-16-00002

YEAR I WORK PLAN AND EVALUATION DESIGN (revised)

SEPTEMBER 30, 2016 – SEPTEMBER 30, 2017

APPROVED: AUGUST 3, 2017

International Business & Technical Consultants, Inc. (IBTCI)

8618 Westwood Center Drive
Suite 400
Vienna, VA 22182
USA

DISCLAIMER

This Annual Work Plan and Evaluation Design is made possible by the support of the American people through the United States Agency for International Development (USAID). The contents are the sole responsibility of IBTCI and do not necessarily reflect the views of USAID or the United States Government.

TABLE OF CONTENTS

Acronyms	3
Section 1. Introduction	4
1.1 USAID Soma Umenye Impact Evaluation Overview and Evaluation Questions	4
1.2 Description of Deliverables	5
1.3 Consortium and Impact Evaluation Team Members	8
Section 2. Technical Activities	10
2.1 Evaluation Design of the USAID Soma Umenye Activity	10
2.2 Random Assignment of Sectors with Longitudinal Sample	13
2.3 Statistical Power and Sample Size Requirements	16
2.5 Managing and Implementing the USAID Soma Umenye Impact Evaluation	21
2.6 Activities Related to Soma Umenye Research Questions	22
IE Question # 1. To what extent are changes in Kinyarwanda reading outcomes for students in Grades 1-3 attributable to the Soma Umenye activity (as a package)?	22
IE Question # 2. To what extent has classroom instruction changed as a result of Soma Umenye?	24
IE Question # 3. What is the evidence that improved student reading outcomes in Kinyarwanda lead to improved opportunities for students to succeed in schooling?	26
IE Question # 4. To what extent are the improvements in the systemic capacity for early grade reading instruction a result of the Soma Umenye activity?	27
Protection of Human Subjects	28
2.7 Instruments	28
2.8 Data Collection and Quality Assurance	30
2.9 Data Analysis	33
2.11 Cost and Cost Effectiveness of the Intervention	40
2.12 Coordination with Other Implementers	40
Section 3. Dissemination of Findings	41
Annex A. Soma Umenye Impact Evaluation Statement of Work	42
Annex B. Year I Work Plan	60
Annex C. List of Sectors in Phases 1, 2, and 3	63
Annex D. Selected Documents	72
Annex E. Conflict of Interest Statements	75

TABLE OF FIGURES

Table 1. Soma Umenye Impact Evaluation Deliverable Schedule	7
Table 2. Illustration of the Randomized Rollout	15
Table 3. Minimum Detectable Impacts (MDIs) for Correct Words per Minute under Different Students Level Attrition Scenarios	17
Table 4. Sample Size and Scheduling	18
Table 5. Analytical Approach to Evaluation Questions	19
Table 6. List of Instruments to Collect the Data Needed by the Impact Evaluation	29
Table 7. Analytical Approach to Evaluation Questions	32
Table 8. Students to be Assessed as part of the Impact Evaluation	33

ACRONYMS

CBO	Community-based organization
COR	Contracting Officer's Representative
CPD	Continuing professional development
CPMD	Curriculum and Pedagogical Materials Development Department
CWPM	Correct words per minute
DOS	Director of Studies
EDC	Education Development Center, Inc.
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
FBO	Faith-based organization
GoR	Government of Rwanda
IBTCI	International Business & Technical Consultants, Inc.
ICC	Intra-cluster correlation coefficient
IE	Impact evaluation
IP	Implementing partner
L3	EDC's Literacy, Language, and Learning project
P1	Grade Primary 1
P2	Grade Primary 2
P3	Grade Primary 3
M&E	Monitoring and evaluation
MDI	Minimum detectable impact
MINEDUC	Ministry of Education
NGO	Non-governmental organization
NISR	National Institute of Statistics of Rwanda
PDA	Personal digital assistant
REB	Rwanda Education Board
RNEC	Rwanda National Ethics Committee
SD	Standard deviation
SEO	Sector education officer
SSL	School subject leader
SU	Soma Umenye
TBD	To Be Determined
TLM	Teaching and Learning Materials
TTC	Teacher Training College
URCE	University of Rwanda College of Education
USAID	United States Agency for International Development
VPN	Virtual private network

SECTION I. INTRODUCTION

On September 30, 2016, USAID/Rwanda awarded a five-year task order, AID-696-TO-16-00002, to International Business & Technical Consultants, Inc. (IBTCI) under the PPL-LER Monitoring and Evaluation (M&E) IDIQ, AID-OAA-I-15-00022, to conduct an Impact Evaluation (IE) of its Soma Umenye (SU) (“Read and Know”) activity implemented by Chemonics International, the implementing partner (IP).

This deliverable constitutes IBTCI’s submission of the Soma Umenye *Impact Evaluation Annual Work Plan and Evaluation Design*, as required by the task order (AID-696-15-00025, section C.4.3). In this document we detail IBTCI’s planned activities for Year 1 of its task order, covering the period between the awarding of the contract to IBTCI to conduct the Impact Evaluation (September 29, 2016) to the end of the federal fiscal year (September 30, 2017). IBTCI’s intended approach to the Impact Evaluation of Soma Umenye (SU) is discussed in Section 2 Evaluation Design. While the Evaluation Design covers the whole period of the contract, activities to be implemented by the IE in Year 2 (October 2017 to September 2018) and in subsequent years are referred to but not formally included in the Year 1 work plan.

The Soma Umenye activity is expected to contribute to the goal of improving reading outcomes in Kinyarwanda for at least 1 million children (unique direct beneficiaries) in public and government-aided schools in Rwanda by the end of Grade 3. The IBTCI IE team will work closely with the IP to ensure we conduct the best possible evaluation of the impact of Soma Umenye activity. To ensure that at least 70 percent of P1-P3 students are able to read grade-level text with fluency and comprehension (ultimate project target), Soma Umenye will implement activities to achieve two intermediate results:

- Intermediate Result 1 (IR 1) is related to the classroom and school-level interventions necessary to improve evidence-based reading instruction, including provision of materials, training and coaching, supportive leadership, and analysis and use of student assessment results.
- Intermediate Result 2 (IR 2) focuses on strengthening the capacity of the education system in Rwanda to implement and support high-quality, evidence-based reading instruction throughout the country during and beyond the life of the Soma Umenye activity.

I.1 USAID Soma Umenye Impact Evaluation Overview and Evaluation Questions

“An impact evaluation helps demonstrate attribution to the specific intervention by showing what would have occurred in its absence.”¹ The various links in the chain of Soma Umenye results will be analyzed using a variety of quantitative and qualitative methods, building up an argument as to whether the theory on which the project is based has been realized in practice. For example, we will examine student reading scores on the EGRA subtests with and without the project as well as practices exhibited by teachers who have received Soma Umenye training and by those who have not. We will seek evidence that improved student reading outcomes in Kinyarwanda lead to improved opportunities for students to succeed in schooling. This will be done by comparing results on a Math test to be developed in 2018 by REB and IBTCI.² We will also assess the extent to which SU initiatives designed to strengthen the capacity of the education system in Rwanda have been implemented and the degree to which they are successful.

¹ USAID Technical Note on Impact Evaluations

² Math was selected due to the fact that most international tests administer Reading and Math tests and use the scores to assess the state of education in the country.

Examining different links in the results chain avoids ‘black box’ impact evaluations, or IEs that provide findings on impact without any indication as to why impact happened. Answering the “why” question requires looking into the intermediate results—starting in 2017—and the final results that the project expects to achieve. In addition, we will assess the cost effectiveness of Soma Umenye, i.e., what costs are involved in taking children from where they are without the project to a higher level of competence, which we hypothesize will be observed on the average scores of the intervention group in 2018. USAID/Rwanda has given the initial push to the process by identifying four evaluation questions:

- To what extent are changes in Kinyarwanda reading outcomes for students in Grades 1-3 attributable to the Soma Umenye activity (as a package)?
- To what extent has classroom instruction changed as a result of Soma Umenye?
- What is the evidence that improved student reading outcomes in Kinyarwanda lead to improved opportunities for students to succeed in schooling?
- To what extent are the improvements in the systemic capacity for early grade reading instruction a result of the Soma Umenye activity?

Each of the questions above has embedded a number of sub-questions with their specific variables that requires data to be collected and analyzed in order to provide a satisfactory answer. In Sections 2.5 and 2.6 we detail how we plan to respond to the set of evaluation questions and the sub-questions that enlarge their reach. While the role of the Soma Umenye M&E system is to collect the data to demonstrate that the project has been implemented according to plan (process evaluation) and has achieved its targets (performance evaluation), the role of the Impact Evaluation is to document the average effects or impact that the project had on the population of interest and the size of these effects. The combination of these three types of evaluation will give USAID/ Rwanda information about how well the project was implemented and managed as well as the size (and therefore, relevance) of the impact it caused.

We assume that Soma Umenye is a well-designed project that addresses the perceived needs of the population of interest as well as the development objectives of USAID. We also assume that the project will be competently managed and will reach its targets partially or fully. To date we have seen nothing that would call these assumptions into doubt. But even when a project reaches its targets, impact is not guaranteed or impact may be too small to justify the investment. The role of the Impact Evaluation is to select an evaluation design, collect the necessary data, and conduct the analyses that will clarify to USAID and to the Government of Rwanda (GOR) the extent to which the implementation of Soma Umenye caused an impact on the reading competency of Rwandan children in grades P1 to P3 and the extent to which changes in reading competency affect success in schooling.

1.2 Description of Deliverables

In this section, we list the required deliverables and the submission schedule for the first year as per IBTCI’s Soma Umenye Impact Evaluation task order.

Annual Work Plan Year I (C.5.1): This document constitutes the Work Plan deliverable for Soma Umenye IE Year I and includes (i) a discussion of the various elements of the IE such as design, instruments, data collection analyses and dissemination of results; (ii) an outline of anticipated activities to be conducted by the IE team; and (iii) an anticipated schedule for the implementation of the IE activities. This deliverable was submitted to USAID on December 14, 2016 and a revised version addressing USAID comments was

resubmitted on January 30, 2017. This fourth version addresses a third round of USAID comments and seeks to clarify remaining concerns expressed when the third version was submitted. As requested by the IE COR, this Annual Work Plan reflects the 2017 Federal fiscal year ending in September 30, not the 2017 academic year, which ends in November 2017.

Evaluation Design (C.5.2): The development of the Evaluation Design, which is incorporated in this document, took into consideration the phased-in implementation plan shared with us by the IP and discussions with the IE COR and other staff of USAID/Rwanda's Education Office. All members of the IE team attended the retreat organized by the IP for Soma Umenye stakeholders on November 3-4 as well as the Education Partners' meeting organized by USAID. These meetings and discussions and the fact-finding discussions with the various implementing partners and with the GOR represented by the Director General of the Rwanda Education Board (REB) provided the feedback we needed to refine the IE design and adapt it to the reality of the country and to the plans of the implementer. The next step will be to formally present the Evaluation Design and the Work Plan to REB. We anticipate that following USAID acceptance of the IBTCI Work Plan, our Methodologist will be travelling to Kigali to conduct an **Interim Update Meeting** with USAID/Rwanda, with the IP, and with REB and other stakeholders to present and discuss the various technical aspects of the design and methodology of the IE.

USAID/Rwanda In-Brief (C.5.3): This deliverable was completed in November 2016 with discussions by IBTCI's Project Director Ed Allan, Team Leader IBTCI, Methodologist IBTCI, and Field Manager Kenn Ndirangu with primary stakeholders in the USAID Program Office and the Education Office.

Data Collection Preparation (C.5.4): The preparations for the first data collection (to be conducted in September/October 2017) will include (i) working with the IP in order to finalize a package of instruments that will allow both Soma Umenye and the IE to collect the data needed to answer the evaluation questions assigned to each; (ii) refining the logistics associated with school visits and data collection in the sectors/schools which are part of the IE sample; (iii) recruiting and training enumerators; (iv) pilot testing data collection tools, data entry processes, and the production of datasets. The Team Leader, IBTCI, will have primary responsibility for the work related to IBTCI's use of the instruments, described below, while the Field Manager, Kenn Ndirangu, managing director of Incisive Africa, will have primary responsibility for the data collection activity.

Instrumentation: For the EGRA, the IE will use the same instruments adapted or developed by the implementer of Soma Umenye and the same data collection procedures used by the IP in order to ensure comparability of data. IBTCI's Field Manager participated in Soma Umenye's EGRA planning workshops and IBTCI's international personnel are also providing feedback into the refinement of the instruments, as already done with the Teacher Observation Protocol. This is necessary because the plan is to share data sets between the two organizations to allow additional analyses of the data collected. In some cases, the IE may need additional questions added to the instruments or even additional instruments—for example, instrument(s) needed to collect data to answer evaluation Question # 3 or the interview protocol to collect data from key informants in order to address Question # 4. To answer Question # 3, whether improved reading skills leads to improved performance in other areas, IBTCI will take advantage of the opportunity offered by Dr. Michael Tusiime, DDG of REB, to work with REB experts in the development of a suitable test.³ Following the rejection of the L3 exam and at USAID's request, IBTCI prepared a memo

³ IBTCI had received technical direction from USAID to use the L3 Mathematics exam developed by EDC. However, when we sought REB's approval to use the test (July 2017) Dr. Michael Tusiime, DDG of REB, informed us that the L3 exam was developed

describing options and points for and against each.

- *Option 1: Postpone developing math assessment until 2018 (preferred option)*
- *Option 2: Develop math assessment immediately and use the August/September data collection period to gather data for reliability*
- *Option 3: Develop math assessment immediately and use it for the August/September data collection period without establishing reliability (not recommended)*

After review, USAID agreed on our preferred option of postponing development of the mathematics assessment until 2018, when we are to prepare mathematics assessment tests for P1, P2, and P3. We anticipate doing this work in collaboration with REB, which is in the process of preparing P2 and P3 assessments. If feasible, we anticipate using the REB-developed instruments. The IE Team Leader will coordinate the work to develop the assessment tool.

Preliminary Findings Workshop with USAID (C.5.7) and Dissemination Presentations (C.5.8): When the preliminary findings from the data collected in September/October 2017 and have been analyzed (January 2018), IBTCI will hold a meeting with USAID to discuss these findings and their implications. The second step will be to make a presentation of the key findings included in the IE report to stakeholders from USAID/Rwanda's Program Office and Education Office, REB, the implementing partner staff, and other stakeholders. The scheduling and format as well as the audience for these deliverables will be based on further discussion with USAID. Table I summarizes the deliverables schedule.

Table I. Soma Umenye Impact Evaluation Deliverable Schedule

Deliverable	Contract Section	Initial Submission Date
USAID/Rwanda In-Brief	C.5.3	October 30, 2016
Annual Work Plan	C.5.1	December 14, 2016 (resubmission 01/30/2017, 02/27/2017, 03/17/2017, 4/30/17, and 06/15/2017)
Evaluation Design	C.5.2	December 14, 2016 (resubmission 01/30/2017, 02/27/2017, 03/20/2017, 4/30/17, and 06/15/2017)
Data Collection Preparation (logistics of data collection, training of enumerators, etc.)	C.5.4	June-August 2017
Interim Meeting Updates	C.5.5	IBTCI anticipated to be in Rwanda after the approval of this document to present and discuss Final Evaluation Design and Work Plan with USAID and REB. IBTCI anticipated to be in Rwanda after the approval of this document to work with the IP regarding activities that will be implemented by SU to improve the systemic capacity for early grade reading instruction and with Kenn Ndirangu on the aspects of the data collection effort and conduct discussions with USAID and the IP.
First Impact Evaluation Report	C.5.6	First Draft January 2018; Final report due after the Preliminary Findings Workshop and 15 days after receiving USAID feedback. ⁴

during the previous curriculum and is obsolete now that there is a new curriculum.

⁴ IBTCI is aware of the current requirements in the Technical Notes of the Education Strategy to submit the report and the data

Preliminary Findings Workshop with USAID	C.5.7	TBD
Dissemination Presentations (REB, stakeholders)	C.5.8	TBD

I.3 Consortium and Impact Evaluation Team Members

The Soma Umenye Impact Evaluation is a consortium of three organizations: International Business & Technical Consultants, Inc. (IBTCI), the prime contractor; Mathematica Policy Research; and Incisive Africa, a Rwanda-based firm.

As the representative of the prime contractor, the **Project Director, Edward Allan**, will provide technical and administrative/contractual direction for the Task Order, oversee the management of budgets and expenses, manage staff and resources to ensure tasks and deliverables are completed properly, maintain control of quality, promote effective collaboration with Chemonics, the IP, and serve as the primary point of contact with USAID.

The **Team Leader, IBTCI**, augmented by support from the home office, will be the primary technical point of contact for this evaluation with USAID, particularly the COR, as well as with our partners, Mathematica and Incisive Africa, and with Chemonics, the IP. The Team Leader will provide overall technical and managerial oversight to ensure that all project deliverables and processes are timely and meet the highest performance and data quality standards. These include, but are not limited to: design and implement the IE in collaboration with the Methodologist and Field Manager, including analysis and interpretation and providing supervision and guidance to the project staff and sub-contractors; ensuring timely and quality submission of all contract deliverables; manage positive and productive relationships with USAID/Rwanda personnel as well as with the staff of USAID implementing partners, GoR, project beneficiaries, and other stakeholders; manage the evaluation team, workload, and relationships, including sub-contractors and short-term technical assistance (STTA); ensure that the evaluation process, data collection, and deliverables are systematic, grounded in verified evidence, and fully answer each evaluation question; provide support and technical assistance during critical times of implementation of the IE; lead the analysis and writing up of results; and maintain regular communication with the IBTCI Project Director.

The **Methodologist, Ali Protik**, from our sub-contractor Mathematica will provide technical input into the design and implementation of the IE and will ensure that the techniques and methods applied in the impact evaluation meet the established rigor, analytical approaches, and methodological principles and quality standards for IE. The Methodologist will ensure that all deliverables and related materials and tasks are in compliance with applicable USAID M&E policies and guidance evaluations and provide guidance to the evaluation team and enumerators; on request, he will provide technical input to the IP in matters related to the IE methods and technical activities.

The **Field Manager, Kenn Ndirangu**, who directs sub-contractor Incisive Africa, will lead and oversee the field data collection, data integrity, aggregation, recording, and reporting of impact evaluation data and ensure that collected and reported data is disaggregated as appropriate. He will liaise with the IP to help

sets to be transferred to the SART within 90 days of the closing of data collection. Please refer to Task Order p. F-2.

assure the quality and compatibility of the data and datasets produced. Mr. Ndirangu will provide the logistics for the data collections effort and the technical guidance, recommendations, and necessary quantitative and qualitative analysis. He will also serve as the in-country point of contact for USAID on a day-to-day basis.

As Quality Control at our Home Office, IBTCI's Vice President for Technical Programs, Annette Bongiovanni, will be reviewing the reporting, deliverables, evaluation design and ongoing progress of the evaluation.

SECTION 2. TECHNICAL ACTIVITIES

Impact evaluation (IE) is an assessment of how the intervention being evaluated affects outcomes. Specifically, in the case of Soma Umenye, the IE will compare results obtained by the intervention and the control groups to assess the extent to which improvements observed in student reading ability, teacher instructional behaviors, and rates of student, teacher, and head teacher tardiness and absenteeism can be attributed to SU. The IE will also identify the apparent contribution of the Soma Umenye activities to improvements noted in the institutional capacity of the Rwanda education system. A well-designed impact evaluation can also assist the IP to understand the impact on the main variables of interest, thus providing information for the redesign of certain activities. A key function of an IE is to produce policy-relevant results that can influence the design of future programs.

An important distinction needs to be made between performance evaluations and impact evaluations. Performance evaluations provide a description of the *factual*. Impact evaluations utilize a *counterfactual* to attribute observed outcomes to the intervention. Counterfactual analysis is also called **with** versus **without**. This is not the same as **before** versus **after**, as the situation “before” may differ in respects other than the intervention.

The proper analysis of impact requires a counterfactual of what outcomes would have been in the absence of the intervention. The most rigorous evaluation design is a randomized controlled trial (RCT), where a randomly assigned group that does not receive the intervention—the control group—is used as the counterfactual. The counterfactual compares actual outputs and outcomes to what they would have been in the absence of the intervention. The difference in outcomes between the beneficiaries of the intervention (the treatment group) and the comparison or control group is a measure of impact. By constructing a counterfactual—a control group—an IE is able to assign observed changes in outputs and outcomes to the intervention. We plan to analyze data to show the impact or effects of the project—differences in the outcome for treatment and comparison groups—and the size of that effect.

2.1 Evaluation Design of the USAID Soma Umenye Activity

The Evaluation Design described in this section covers the period between the awarding of the contract to IBTCI to conduct the Impact Evaluation (IE) on September 29, 2016 and the end of the IE in September 2021. In this section of the Work Plan we describe our intended approach to the activities to be conducted as part of the USAID Soma Umenye Impact Evaluation. Due to the length of the contract (five years), the design is subject to revision as the Soma Umenye project evolves. The first steps for the conduct of this activity were taken by USAID Rwanda when it decided to conduct an Impact Evaluation and identified key evaluation questions.

The IE design for Soma Umenye reflects the current plan by REB and the IP to roll out the implementation of the project in three phases, starting with grade P1 and adding a grade with each subsequent year. Each phase is one year long and aligned with Rwanda’s school year, which begins in January/February and continues until October/November. The phased implementation allows time for the IP team to prepare materials and to adjust activities as needed. Summing up, the following activities are likely to be implemented as part of Soma Umenye:

I. Teacher training

2. Print-rich classroom environment
3. Formative assessment
4. Headmaster training
5. Coaching/mentoring support for teachers
6. Reading materials

The IP plans to deliver a *package* that include all or a subset of the above six activities. The IE design allows for the evaluation of the impact of this *intervention package*. Based on our discussion with the IP, the first five activities are likely to be implemented as a package in phases, but the reading materials (activity 6) are likely to be delivered to all schools nationwide during the 2017 academic year as leveled texts have been requested by the Rwanda Education Board (REB). For the IE, the intervention package comprises the activities that SU will implement in the intervention schools. For example, if four of the activities listed above are implemented in phases (with Phase 1 schools getting them in the 1st year, Phase 2 schools in the 2nd year, and Phase 3 schools in the 3rd year) then these four activities will constitute the package of activities. It is possible that the implementation plan for these activities will evolve over time. We also acknowledge that the implementation of the package can be uneven across schools. The IE estimates the average impact of the package “as it is implemented.” In other words, we estimate the average impact of SU in terms of the difference between the treatment group that received the intervention package as it was implemented and the control group that did not receive any of the intervention. As is the standard practice, we will carefully describe the implementation to inform the reader so that the impacts can be interpreted in the right context.

Phase 1 (school year 2017). As described in Soma Umenye’s Year I Work Plan, the Soma Umenye activity will be implemented in first grade Kinyarwanda classes in all government-supported and public primary schools in about 30 percent of sectors in each district. Activities involved in Soma Umenye include:

- a set of supplementary Kinyarwanda readers for P1 (numbers will be determined through a situation analysis and consultation with the Curriculum and Pedagogical Materials Development Department (CPMD)) as well as a locking classroom bookshelf to hold them;
- teacher continuing professional development (CPD) (10 days of face-to-face training plus structured and guided mentoring support as described below); only P1 students will receive these in Phase 1;
- support (10 days of face-to-face training) for Kinyarwanda mentors—who might be directors of studies (DOS), head teachers, or Kinyarwanda school subject leaders (SSL) — to conduct structured classroom observations and mentoring;
- support (four days of face-to-face training) for head teachers as instructional leaders in their schools (aligned with REB’s head teacher standards);
- print-rich classrooms (Soma Umenye will provide three posters per classroom, one display board per class to display word cards and children’s work, and paper/pen supplies for teachers to use to make their own posters) and teachers will be trained to make their own teaching aids using locally available materials.⁵

Phase 2 (school year 2018). In 2018, Soma Umenye will cover all government-supported and public primary schools in 60 percent of sectors and will work with P1 and P2 teachers in these schools (with the same intervention package).

⁵ Soma Umenye Work Plan, Year I. October 17, 2016.

Phase 3 (school year 2019). In 2019, Soma Umenye will expand to all public and government-aided primary schools in the country (2,476 schools in 2015) and will work to improve literacy instruction in P1, P2, and P3 Kinyarwanda classes.

Soma Umenye will be implemented in three phases as described below. Ignoring for now provincial and district administrative divisions, the ~2,500 public and government-aided schools are assigned to 416 sectors.

- Phase 1 (school year 2017): Implementation will occur at 30 percent of the 416 sectors and will involve only P1 students.⁶ For the purpose of the IE design and work plan, we will call these schools the Phase 1 schools.
- Phase 2 (school year 2018): Implementation will continue at Phase 1 schools and will start in an additional 30 percent of sectors meaning that both P1 and P2 students will benefit from the intervention. For the purpose of the IE design and work plan, we will call the schools in the additional 30 percent sectors the Phase 2 schools. (Note: the majority of P1 students from Phase 1 schools will be P2 students in 2018).
- Phase 3 (school year 2019): Implementation will continue at Phase 1 and Phase 2 schools and include the remaining 40 percent of sectors, which form the comparison or control group. In 2019 the three grades—P1, P2, and P3 students—will be receiving the intervention. For the purpose of this work plan, we will call the additional schools in the final 40 percent sectors the Phase 3 schools. (Note that the P1 students from Phase 1 schools in 2017 will be P2 students in 2018 and will be P3 students in 2019; P1 students from Phase 2 schools in 2018 will be P2 students in 2019; however, Phase 2 schools are not included in the IE.)

The proposed breakdown of number of sectors per phase follows.

Phase 1, 2017 academic year	30 percent of 416 sectors = 125 sectors
Phase 2, 2018 academic year	30 percent of 416 sectors = 125 sectors
Phase 3, 2019 academic year	40 percent of 416 sectors = 166 sectors

Activities for 2020-2021. To a significant extent, what IBTCI does in 2020-2021 depends on what Soma Umenye, other donors, and the Government of Rwanda plan to do and this is not known in 2017. From the end of data collection in September/October 2019 to about May 2020, we will be involved with data analyses, preparation and revisions to the report on the IE itself, including conducting the cost effectiveness analysis, and dissemination of information on the IE and on findings to date on classroom instruction, pre-service and in-service teacher training, and relevant aspects of institutionalization. During FY 2020 and FY 2021, we will likely be conducting special studies, e.g., on various sub-populations, as will be agreed upon between USAID and IBTCI. For example, as a subset of findings from the IE, IBTCI will have identified high-performing and low-performing intervention and control schools which could be a

⁶ We discussed the distribution of sectors into phases with the IP. Differences in distribution have implications for statistical power for the IE. We discuss the statistical power of sampling sectors vs. schools in Section 2.3.

useful starting point for various studies, “TBD,” and we will endeavor to interact with URCE to develop joint research activities.

In addition, the overall focus of the remainder of the Task Order will be on the further study of Question #2, “To what extent has classroom instruction changed as a result of Soma Umenye?” and Question #4, “To what extent are the improvements in the systemic capacity for early grade reading instruction a result of the Soma Umenye activity?” We will begin this research in 2017 by identifying the current state of teacher training and classroom instruction

2.2 Random Assignment of Sectors with Longitudinal Sample

A major concern in selecting the evaluation approach is the way in which the problem of selection bias will be addressed. How this will be done depends on an understanding of how such biases may be generated, which requires a good understanding of how the beneficiaries are identified by the program.

Considering that the Soma Umenye IE is being designed ex-ante, randomization was possible. When the treatment group is chosen at random, then a random sample drawn from the sample population is a valid comparison group, and will remain so provided contamination can be avoided. The comparison group sample must be of adequate size and serve as the basis for a credible counterfactual, addressing issues of selection bias (for example, if the comparison group is drawn from a different population than the treatment group) and contagion (if the comparison group is affected by the intervention or a by a similar intervention implemented by another agency).

To rigorously evaluate the impacts of the Soma Umenye package of interventions, we propose an RCT, where the 416 sectors in Rwanda are randomly assigned into the three phases of the planned implementation, with longitudinal tracking of students. In our original proposal, we had proposed a random assignment of schools to intervention or control. However, in discussion with the IP we proposed the random assignment of the 416 sectors in Rwanda into the three phases of the planned implementation.⁷ Because schools in sectors selected for implementation in Phase 2 and Phase 3 will not receive the package of interventions for one and two years, respectively, they can represent the counterfactual condition—what would have happened without the intervention. With random assignment of sectors into phases, schools in sectors selected for Phase 1, Phase 2, and Phase 3 are likely to be the same, on average, in terms of their characteristics. Thus, comparing schools in Phase 1 sectors schools in Phase 3 sectors will allow us to measure the impacts of the Soma Umenye package of interventions on student reading ability and on other variables suggested by the IE evaluation questions.⁸

⁷ There are advantages and disadvantages to randomly assigning sectors as opposed to schools. First, random assignment of sectors makes it more feasible to implement the intervention. Random assignment of all 2,493 schools (list of schools provided by the IP) to the intervention or control groups could result in schools in each of the phases being scattered around the country, which would pose significant implementation challenges to the IP. Second, because sector education officers (SEOs) are in charge of all schools in their sectors, selecting some schools in a sector for Phase 1 and others for Phase 2 and/or Phase 3 could result in contamination. The disadvantage is the smaller sample size of the unit of assignment (416 sectors vs. 2,493 schools), which results in lower statistical power to detect impacts compared to random assignment of schools. We discuss the issue of statistical power in sub-section 2.3.

⁸ Sectors in either Phase 2 or Phase 3 can represent the counterfactual condition. We explain why phase 2 sectors are not included in the IE in the next paragraph.

For the IE, we will consider the schools in Phase 1 sectors as the intervention group and the schools in Phase 3 sectors as the control group. We will not include schools in Phase 2 sectors as leaving them out will allow us to measure impacts of a two-year implementation of Soma Umenye package of interventions and will avoid any risk of contamination between schools in Phase 1 and Phase 2. We will select random samples of students from both the intervention and the control group schools and will longitudinally track them across years for the IE. This design will permit analysis at multiple levels, allowing us to compare the intervention versus the control group at the end of 2018 in order to observe the two-year impacts for P2 students and the one-year impact for P1 students in schools selected for the SU in Phase 1 sectors compared to students in the control group schools selected for SU in Phase 3 sectors.

Random assignment process and stratification. The IE team guided the IP in the random assignment of sectors into phases conducted in mid-December 2016 so that the implementation of Soma Umenye could be planned accordingly before the beginning of the intervention in January of the 2017 school year. It was important that random assignment of all three phases took place prior to the implementation of the project so that it became clear from the start which sectors would receive the intervention and which sectors would be part of the comparison group.

The random assignment was carried out according to the following stratification:

- *Stratification by district.* We stratified the random assignment of sectors by district for two reasons. First, there are geographic variations in education outcomes in Rwanda, which could confound the effectiveness of *Soma Umenye* if it is implemented in one region of the country and not in the others. By assigning sectors within each district, we ensure that the intervention and the control group sectors are evenly spread across the country. Second, because all districts and thus provinces are represented in both the intervention and the control groups, we will be able to explore any variation in impacts by province. Third, this strategy is helpful for the IP as they will not have to exclude any district from their implementation plan in any of the phases.
- *Stratification by sector size.* Within each district, we further stratified the random assignment by the size of the sectors in terms of the number of schools. Because sectors with a large number of schools are likely to operate differently than sectors with a small number of schools, the IP recommended that we follow this stratification process. For example, larger sectors may be better able to attract better quality teachers or they may have more resources to monitor student progress. We stratified each sector into two pots – one containing the larger sectors and one the smaller sectors.

The IP and IE team concurred that the random assignment should be carried out publicly in each district where sector representatives are present. In preparation for this public random assignment, we compiled a list of sectors by district and, within each district, grouped them into two pots according to their sizes in terms of the number of public and government-aided schools. The first pot in each district contained sectors that have more public and government-aided schools than the average number of these schools in the district. The second pot contained sectors that have less than or equal to the average number of these schools in the district. In addition, we calculated the number of sectors in each pot that should be assigned to Phase 1, Phase 2, and Phase 3 so that the final distribution would be roughly 30 percent, 30 percent, and 40 percent across the three phases. In case where the percent distribution resulted in a fraction, we always rounded up in favor of Phase 1 and rounded down for Phase 2. This resulted in a slightly higher percentage of Phase 1 sectors (33%) and lower percentage of Phase 2 sectors (27%) than

was originally proposed. The IP used the list of sectors for each district and publicly conducted lotteries in the presence of sector officials to assign sectors to Phases 1, 2, and 3 for each pot.⁹ The random assignment by pot ensured a more even distribution of large and small sectors, in terms of number of public and government-aided schools, into the three phases. The list of sectors by pot and districts selected for the roll out implementation phases is included in Annex C.

In summary, the randomization of sectors into phases was carried out in the following manner:

- Random assignment took place in each of the 30 districts to ensure equity across districts in receiving the intervention. Sectors within each district were randomly assigned to Phase 1, phase 2, and phase 3. Across all districts, 33 percent of the sectors were randomly assigned to phase 1; 27 percent to Phase 2; and 40 percent to Phase 3.
- The assignment of sectors into Phases 1, 2 and 3 was strictly random, per guidance provided by the IE to the IP.
- Once a sector was assigned to a phase, all schools in that sector are to receive the intervention in that phase.

The following table shows the plan for carrying out the above randomly assigned rollout plan that will allow us to estimate the impacts of the Soma Umenye package of interventions, labeled “T”.

Table 2. Illustration of the Randomized Rollout

Groups	Year 1 (2017)	Year 2 (2018)	Year 3 (2019)
Group 1 (Intervention)= Phase 1 schools (all schools in 33% sectors)	T in P1	T in P1, P2	T in P1, P2, P3
Group 2 (Non-study)= Phase 2 schools (all schools in 27% sectors)	-	T in P1, P2	T in P1, P2, P3
Group 3 (Control)= Phase 3 schools (all schools in 40% sectors)	-	-	T in P1, P2, P3

Data collected by the IE will be disaggregated by sub-populations of interest related to the impact evaluation questions, including student sex; scores obtained by students in urban and rural schools, (We will use GoR definition of urban/rural as well as relative distance from Kigali,¹⁰) and other variables of interest, including students with disabilities and proxy measures for socio-economic status. Note that when we disaggregate data according to the classification of the school (urban and rural) we will utilize the GoR definition of urban and rural, the same that is used in LARS II. In our analyses of the data we will pay attention to students exposed to the USAID-funded *Mureke Dusome* activity, and students with disabilities. These sub-populations have the potential to become the object of special or focused studies of interest to USAID/Rwanda and to the Government of Rwanda (GoR). Starting in Year 2, as data are analyzed after the first assessment in September 2017, the IE team will prepare a list of possible special studies topics and, in consultation with USAID/Rwanda and REB, determine which topics are more

⁹ IBTCI provided the IP with guidelines on how to conduct random selection of sectors and this was done in December 2016, before the start of the 2017 school year in January.

¹⁰ The 2012 census differentiates between urban and rural. Previously collected baseline data found that the lowest performing 20 percent of schools average a distance of 116 kilometers from Kigali and the highest performing 20 percent average 59.2 kilometers from Kigali.

relevant, show more promise, and could become the focus of special quantitative or qualitative studies.

2.3 Statistical Power and Sample Size Requirements

An IE can detect true intervention impacts only if it has sufficient statistical power. The statistical power depends on factors such as the scale of the true difference or impact that occurs, the sample size, the evaluation design, the number of assigned units (sectors), the correlation of outcomes within clusters, the explanatory power of covariates, and the variance of the outcome measure. Ideally, impact evaluations are powered to detect, measure and validate smaller, rather than larger minimum detectable impacts (MDIs), so that at least small yet meaningful impacts—which are common in educational interventions—can be documented and lessons regarding what works can be learned.

For this impact study, the major uncertainty to statistical power is related to the attrition rate because we will track students longitudinally and dropout rates can be as high as 15% as estimated by Statistical Yearbook. However, the estimates from the Statistical Yearbook are conservative and dropout rates can be higher in reality. In consultation with the IP and USAID, we decided to start with a relatively large sample size of 18 students per cohort per school. To determine the statistical power of the IE with this sample size and given the uncertainty regarding student-level attrition, we estimate MDIs under three attrition rate scenarios of 15%, 18%, and 20%. We estimate the MDIs for one of the key outcomes of the study—students’ scores on the oral reading fluency assessment of the EGRA—based on the impact evaluation’s design, proposed sample size, and parameter estimates from previous studies. The MDIs show the smallest effects that the evaluation can reliably detect, measured in the units of the outcome variable. We also report MDIs in terms of the percent change in standard deviation units.

Specifically, Table 3 shows MDIs for a random assignment design with a longitudinal sample of students, in which 33 percent of Rwandan sectors are randomly assigned to the treatment (Phase 1) and 40 percent of sectors to the control (Phase 3) group and for one selected outcome, Oral Reading Fluency. We assume that the standard deviation (SD) of this outcome will be 24.7 correct words per minute (CWPM) based on findings from a previous study in similar settings.¹¹ We used numbers for the P3 students, as they will be the subject of the end-line impact evaluations. Our calculated SD is larger than the SD of 17 wpm for P2 students that was provided in a comment made by USAID on the previous version of this document. Our power calculation is more conservative and the IE will be able to detect slightly smaller impacts “if” SD for P3 students is actually smaller. For the purpose of this design report, this slight change in MDI has little implication, and so we decided to keep our conservative estimates. The MDIs are based on our data collection plan of one school per sector. Because the impact analysis will be conducted by grade, we show the power calculation for one illustrative grade.

¹¹ RTI (2011) measured comparable mean and standard deviations of EGRA oral reading fluency results for Grade 3 students in Senegal. We computed the coefficient of variation from the Senegal results, and used it to estimate a standard deviation value for Grade 3 students in Rwanda based on the mean value for Rwandan P3 students reported in USAID (2014). We were not able to locate both means and SDs for P3 students, which we need for the power calculations, in the relevant Rwanda documents. The available SD for Rwandan P2 students in oral reading fluency is 17 words per minute, which is smaller than the calculated SD we used in the power calculation. As such, our power calculation is more conservative and at the design stage, we decided to continue using these estimates rather than more optimistic ones.

For this impact study, our assumptions are consistent with established statistical procedures (see Duflo et al., 2008, Bloom, 2005). As a rule of thumb we assume (i) 80 percent power, (ii) a 95 percent Confidence Level (5 percent level of statistical significance), (iii) a two-tailed test, (iv) a 45 percent probability of assignment to the intervention group (i.e., to sectors in Phase I), and (v) an Intra-Cluster Correlation Coefficient (ICC), a measure of the similarity of students in the same school) of 0.25. Further, we assume that there will be no cluster-level attrition—all sectors assigned to treatment and control will be part of the study. We finally assume that covariates in our analysis such as baseline test scores (if available), school size, and student age and sex will account for 30 percent of the variation in students' scores at the individual level and 30 percent at the school level.

Table 3. Minimum Detectable Impacts (MDIs) for Correct Words per Minute under Different Student-Level Attrition Scenarios

	Attrition Rate at the Student Level		
	15 Percent	18 Percent	20 Percent
Sample size at the sector and school level			
Number of Phase I sectors	136	136	136
Number of Phase 3 sectors	168	168	168
Number of total sectors for IE	304	304	304
Number of total schools for IE	304	304	304
Student sample size and MDIs for full sample			
Number of students per school per cohort	18	18	18
Number of total students for IE	5,472	5,472	5,472
Correct words per minute (% SD)	3.67 (.148)	3.68 (.149)	3.69 (.149)
Student sample size and MDIs for subgroup with 50% sample			
Number of students per school per cohort	9	9	9
Number of total students for IE	2,736	2,736	2,736
Correct words per minute (% SD)	3.96 (.160)	3.98 (.161)	3.99 (.161)

Note: MDIs are for a two-tailed test with 80 percent power and a 95 percent level of significance, 45% probability of being assigned to the treatment group, an intra-cluster correlation coefficient of 0.25, regression R^2 values of 0.3 both at the student and the school level that indicate the amount of variation in outcome explained by baseline and other control variables, and no cluster-level attrition.

With 15% attrition rate, the MDI is 3.67 correct words per minute from a baseline mean of 22.1 correct words per minute as per the National Oral Reading Fluency and mathematics Assessment Baseline Report and when we sample 18 students per cohort per school. So the evaluation will be able to detect an impact of 3.67 correct words per minute or greater under this scenario. The MDIs with 18% and 20% attrition rates are slightly higher at 3.68 and 3.69 correct words per minute, respectively. The total number of schools and students (for one cohort) that would be included in the impact analysis is 304 and 5,472, respectively. Table 3 also shows MDIs for a subgroup analysis involving half of the sample, based on gender for example. The IE will be able to detect impacts of 3.96, 3.98, and 3.99 correct words per minute under the three attrition rate scenarios, respectively, for a 50% sub-sample. The total number of schools and students (for one cohort) that would be included in the impact analysis of a sub-sample involving half of the sample is 304 and 2,736, respectively.

The roll-out plan described in Table 2, combined with the calculations of MDI, in this and subsequent sections, yields the maximum annual and cumulative sample size indicated in Table 4. The longitudinal

nature of the sample is indicated in the table through color codes – same cohorts of students are marked by the same color across years. For example, the cohort that starts P1 in 2017, marked by blue, will be in P2 in 2018 and in P3 in 2019. Similarly, the cohort that starts P1 in 2018, marked by green, will be in P2 in 2019. The sample sizes in each year indicate the maximum available sample size for analysis in that year. However, the actual available sample size will be smaller depending on the observed level of attrition across years. Table 8, in Section 2.8, indicates the breakdown of the sample size, disaggregated by Treatment vs. Control.

Table 4. Sample Size (students) and Scheduling

	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>Total</i>
2017	5,472	X	X	5,472
2018	5,472	5,472 ^a	X	10,944
2019	5,472	5,472 ^a	5,472 ^a	16,416
Total	16,416	10,944	5,472	32,832

Note: Students in the same cohort are indicated by the same color code across different years.

^a These numbers indicate the maximum sample size of 18 students per school per cohort given that all students sampled for the cohort for the first time are tracked in subsequent years. The actual sample size in these years will depend on the observed level of attrition.

As this table shows, data are collected only for grades where there can be comparisons between students in intervention schools and those in control schools. Because teachers teach multiple grades, teacher training for P1 can result in desirable spillover for P2 and P3 students in intervention schools in 2017 and in P3 in 2018. While valuable information on this possible spillover can be collected directly from P2 and P3 students who might be beneficiaries of the spillover, the cost of doing so would be excessive considering the budget to conduct the IE.

Note that in 2019, comparing P1 cohort students in the treatment schools, a cohort that received the intervention in 2017 and 2018 and 2019, with P1 students in the control schools, a cohort that received the intervention only for one year, will give us an impact estimate as the difference between receiving three years of SU versus receiving one year. The comparison between P1 in control schools in 2018 and P1 in control schools in 2019 will provide an impact estimate between no intervention and one year of the Soma Umenye intervention.

2.4 Procedures for Longitudinal Sampling

Sampling Procedures for Schools. As the first step for selecting the sample for analysis, IBTCI will select one school within each of the treatment and control sectors to be visited randomly. The decision to select one school in each sector is guided by the calculation of MDI presented above. Because the random assignment is conducted at the sector level, the calculated MDI is extremely sensitive to the number of sectors. This is a common feature in cluster random assignment studies where relatively large sample size of the unit of assignment (sector in this case) is required to gain a targeted level of precision compared to non-cluster random assignment studies (e.g. at the student level) where the unit of assignment and analysis are the same (Schochet, 2008). By selecting one school per sector, we ensure that each intervention and control sector is represented in the sample and thus maintain the number of sectors

required for the MDI calculation presented above. The resulting sample size of intervention and control schools are presented in Table 5.

Table 5. Sample Size of Schools

	Sample size
Number of intervention (phase I) schools	136
Number of control (phase 3) schools	168
Total number of schools for IE	304

One consequence of the above selection process is that the sample will contain an equal number of schools—one to be exact—from each of the sectors irrespective of the size of the sectors in terms of the number of schools they have. In other words, the probability of a school to be selected for the sample will not be the same—schools in a large sector will have a lower probability of selection to our sample and vice versa. We will use sampling weights in our analysis to correct for this issue.

Sampling Procedures for Students. To sample students to be included in the IE and for longitudinal tracking, from the randomly selected schools, we will first randomly select one classroom from all classrooms in a grade and then randomly select students from the selected grade. This selection process is common in many education studies and is practical given that in many Rwandan schools PI is taught in two shifts – morning and afternoon. In addition, average enrollment of PI students in Rwandan public and government-aided schools is over 500 students. Thus, selecting students and collecting data from them across both shifts would have posed a significant challenge.

Instead, we will administer the sampling and data collection within the morning shift. This will give enumerators ample time to follow the necessary data collection steps and will also have minimum instructional disruption for students.¹² Specifically, IBTCI will proceed in the following steps to sample students from each randomly selected control and intervention school:

1. *Random selection of a classroom:* First, we will randomly select one classroom from the total number of PI classes in the morning shift.
2. *Collecting student roster.* Next, we will collect the student roster as maintained by the classroom teacher for the randomly selected classroom.
3. *Random selection of students stratified by gender.* After collecting the roster of students for the randomly-selected classroom, we will use the roster to establish the original sample of 5,472 PI students. Sampling of students at each school will be stratified by student gender. At each school in our sample, we will randomly select 9 boys and 9 girls from the randomly-selected classroom. This random selection of students will be carried out according to a pre-set protocol detailing the rules of selection in a written statement that the data collection teams will adhere to. In addition, the data collection team will execute the random selection in the presence of the headmaster and/or a representative from the school staff and a representative from the community. This process will represent fairness while assuring the random selection of students according to the necessary research protocol. We will follow the same procedure to randomly select an additional

¹² Since parents are not given a choice to select the shift for their child(ren) and since in-school cohorts alternate periodically between morning and afternoon, there is no a priori reason to believe that there would be any significant difference between students from morning shifts and those from afternoon shifts.

5,472 students from the entering P1 cohort in each of 2018 and 2019 (Table 4).

4. *Longitudinal tracking of students.* Upon random selection of students for inclusion in the IE sample, we will longitudinally track these students for participation in the assessment for the one-year impact evaluation in September 2017, for the two-year impact evaluation in September 2018, and for three-year impact evaluation in September 2019.
 - Treatment of repeaters: As the 2017 cohort progresses from P1 to P2 and then to P3, we will attempt to test and interview students who were selected in P1 for the study irrespective of their enrollment status. For example, students who repeat P1, (repetition rates are as high as 27%) will be tested and interviewed in 2018 as part of the cohort that began in 2017, even though the rest of the cohort has moved to P2 while s/he remains behind. Repeating students are not considered to be attrited from the sample in longitudinal studies as long as the program continues to serve them.
 - Treatment of absent students: If a child is absent on the day of data collection,¹³ we will make one attempt to collect data for the child by locating the child in the community based on the home address of the child in the school record.
 - Treatment of students no longer enrolled: Similarly, we will make one attempt to collect data for children who are no longer enrolled at the school (e.g., they have dropped out or transferred to a different school) in 2018 or 2019 by locating them in the community. If we cannot locate them in the community, they will be considered attrited from the sample. The remaining sample will still serve as the longitudinal sample for the study. In calculating the MDI for the IE, we have accounted for a 15-20% attrition of student-level sample (Section 2.3). While MINEDUC statistics show repetition rates as high as 27% in P1, drop-out rates are much lower in the early grades—about 15%. This suggests that there would be little likelihood that attrition will be a problem. Please note that repeaters continue to be part of the longitudinal sample.

For the initial assessment, we will obtain assent from each student in accordance with REB and NISR procedures and will assign each student participating in the IE a discrete identification number which will be maintained securely by IBTCI. For subsequent assessments with a student, we will also obtain parental consent. No individually identifiable information will be used in preparation or dissemination of findings from this IE.

Sampling Procedures for Teachers. To examine impacts of SU on teacher instructional practices, we will compare a sample of teachers in the intervention group with a sample of teachers in the control group. In each intervention and control school, one teacher will be automatically and randomly selected when his/her P1 classroom is randomly selected during the process of student-level sampling.¹⁴ This will result in a teacher sample of 304—136 in the treatment group (one from each of the 136 treatment schools) and 168 in the control group (one from each of the 168 control schools). We will longitudinally track these teachers in 2018 and 2019 as long as they are teaching in the same school. Note that teachers selected in the first year may not teach all of his/her students in the sample in subsequent years. However, we will follow the selected teachers and students irrespective of their respective assignments.

¹³ Because we will be selecting students for the first time in 2017, we will not encounter this problem in 2017 but in subsequent years.

¹⁴ Teachers frequently teach multiple grades in Rwanda. So a teacher assigned to P1 may also be assigned to teach P2 and/or P3 classrooms.

2.5 Managing and Implementing the USAID Soma Umenye Impact Evaluation

There are challenges to the validity of the results obtained by any research study and challenges to be considered in order to maintain the integrity of the Impact Evaluation.

Clear understanding of what an Impact Evaluation entails. Many types of evaluation use the difference in the outputs or outcomes before versus after the intervention. Before versus after is a good measure of the achievement of objectives or targets but it is not a good impact measure as it fails to control for other factors external to the program. Using a *counterfactual* (control or comparison group), the Soma Umenye IE will measure differences in outcomes that can be observed between intervention and comparison groups after implementation cycles—usually one or two school years but it could be longer. The IE will also examine the size of the effects that can be attributed to the project.

Fidelity of Implementation. It is important that different components of a project are implemented in congruence with the evaluation design and in substantially the same way across all intervention schools. For example, if teachers are non-randomly selected to receive training, it will likely bias the impact estimates, upward if more effective teachers are selected or downward if less effective teachers are selected. Also, the level of possible impacts may be affected if the planned implementation package is not fully deployed during the evaluation period.

Contagion or contamination. The comparison group is contaminated if it is subject to a similar intervention, either by spillover effects from the intervention or another donor starting a similar project. This becomes especially challenging when there are changes in government priorities and decisions are made that affect the integrity of the control group.

Attrition. Because of the longitudinal design, attrition may become a problem. Random attrition will only reduce the statistical power of our study, but attrition that is correlated with the intervention may bias our impact estimates. For example, if students who are benefiting the least from *Soma Umenye* tend to drop out of the sample, ignoring this fact will lead us to overestimate the impact. While randomization ensures independence of potential outcomes in the initial intervention and control groups, it does not hold after non-random attrition. For a cluster randomized design like ours, attrition of the unit of assignment (sector) is of greatest threat, but attrition at the unit of analysis (students) can also create meaningful bias. While we do not expect attrition at the sector level, we do expect attrition at the student level. Because attrition is a difficult problem to solve ex-post, we have included our detailed plan above to manage attrition during the data collection process.

Despite our best attempts to collect data from children who are no longer enrolled at the school (e.g., they have dropped out or transferred to a different school) in 2018 or 2019 by locating them in the community, we may still lose some of these students if they move out of the community. In that case, we will follow a two-step process to determine attrition bias:

1. *Assess the extent of attrition bias.* We will assess the extent of the attrition problem based on the standards developed by the What Works Clearinghouse (U.S. Department of Education 2008). If the sample meets an attrition standard based on a combination of overall attrition and differential attrition between research groups (intervention and control), the risk of serious attrition bias is low. However, if the sample used for the impact analysis fails to meet the attrition standard, we will proceed to:

2. *Equivalence of baseline characteristics.* In this step we will test intervention and control groups in the analysis sample for equivalence of observable baseline characteristics (Duflo et al., 2008). We are limited in our ability to check for baseline equivalence of student performance as no baseline testing was conducted, but we will check all available demographic characteristics.

Analyses that fail to meet both the attrition and equivalence standards will be determined to have risk of bias. We will clearly discuss these results in the context of the attrition bias.

Generalization possible from impact evaluations. Impact evaluations are usually of specific interventions in a specific context. It is not necessarily the case that the findings of the USAID Soma Umenye program in Rwanda can be generalized to the same intervention in different contexts. Our approach is to collect data on variables that will help the understanding of the context in which the intervention did or did not work, and so that findings can be generalized to other similar contexts in which the same findings may be expected.

2.6 Activities Related to Soma Umenye Research Questions

The overall question to be answered by the Impact Evaluation refers to whether the program has made a difference and, if so, what the size of the program's effects are and whether they fully justify the cost of the intervention. An observation of the control group will tell us what would have happened in the absence of the program. Specifically, the Soma Umenye Impact Evaluation has been designed to provide answers to the questions of interest as stated in the RFTOP. In this section we provide an overview of the IE activities and discuss important issues and requirements associated with providing answers to each of the four IE evaluation questions posed by USAID.

IE Question # 1. To what extent are changes in Kinyarwanda reading outcomes for students in Grades 1-3 attributable to the Soma Umenye activity (as a package)?

When random assignment is used to place students in intervention and control groups, an Impact Evaluation does not require a Baseline given that the impact comparisons are made between those who receive the intervention and those who do not. However, having students or schools completely randomly assigned to treatment and control groups all over the country would represent a challenge for the IP in terms of implementation strategies and financial cost. In conversation with the IP it was decided to place sectors, and by extension, schools, in the control group, the excluded group (schools slated to begin Soma Umenye in 2018), or the intervention group.

We intend to collect data from PI students in September/October 2017 as the first measure of student reading skills taken by the IE. The findings of this first round of data collection will assist us in the identification of context variables and in our thinking regarding special, targeted studies to be conducted in 2018 and 2019 to examine specific issues related to small subsets of the population. These targeted studies will seek answers to the two sub-questions that are of interest to USAID Rwanda and to the Government of Rwanda: *To what extent has USAID Soma Umenye had an impact on various categories of especially vulnerable children?* and *To what extent have students in different parts of Rwanda benefitted from Soma Umenye?* Section 2.10 provides more detail related to these studies. The data collected in September 2017 will allow us to make the first comparison between the outcomes of students in intervention and the control group and at the same time examine the size of the effects observed, which we expect will increase as Soma Umenye matures.

As explained in the Early Grade Reading Assessment Toolkit, Second Edition (March, 2016) EGRA subtasks are aligned to the components of reading. Because skills are acquired in phases, at a given point in time, some subtasks are likely to have a floor effect—that is, most children would not be able to perform at a sufficient skill level to allow for analysis—depending on where the children are in their development. However, USAID and the GoR are interested in the indicator stated as “...proportion of students (grades 1 – 3) who can read and understand grade-level text.” To answer questions related to this indicator we will use the same EGRA sub-tests adapted or developed by Soma Umenye that include Listening Comprehension, Letter Identification, Familiar Word Reading, Syllable Pronunciation, Fluency, and Reading Comprehension to assess students in our sample in 2017, 2018, and 2019. Data will be analyzed in 2017, 2018, and 2019 to show the level of attainment of this indicator at each grade and compare with scores of the control group. The assessment will allow the **with** and **without** comparisons (outcomes of intervention vs. control group students) needed by the IE.

Even though Soma Umenye will have started activities during 2017, implementation on the first year of a project is often slowed down by external and internal factors, thus reducing the size of the expected impact. It is possible that we find only minimal differences between children *with* the program (intervention) and those *without* the benefits of Soma Umenye (control) in the September/October 2017 2017 assessment. In case significant—even if minimal—differences are found between the groups, we will also examine the size of the effects using Cohen’s d statistics. This same comparison will be made at key points—at the end of 2018, and 2019 school years. Data will be analyzed each time to determine the impact of Soma Umenye on students’ EGRA outcomes and the size of effects observed between the two groups in the IE sample as well as the impact on other variables of interest such as repetition and dropout.

In September 2018 the first cohort (P1 in 2017) will now be P2 and the second assessment of reading competence will be conducted using the same EGRA subtasks tools used in September 2017 with the exception of the Fluency and Reading Comprehension subtasks that will require the use of a P2 grade-level passage. We assume that having now received the intervention for two years, the differences between the control and treatment groups will have increased and that significant differences and larger effects will be observed in the number of letters and in the number of words or syllables recognized/read correctly per minute. Most important, it is possible that significant differences will be observed between control and treatment groups in Fluency and Reading Comprehension.

In September 2019 we will conduct the third assessment using the same instruments as used in 2018 with the exception of the reading passage (a P3 level reading passage will be needed). We expect that after three years of implementation, comparisons between average scores of the intervention and comparison groups will show larger and significant differences, representing the impact of the intervention on student reading outcomes. We will also examine data collected by SU MEL system to understand the impact the SU has had on other variables such as repetition and dropout rates. Given that the IE proposed to use an RCT model, we will be able to say that any differences observed could be attributed to Soma Umenye while showing what the situation would be in the absence of the program.

In order to fully answer Question # 1, we need to provide information on the subgroups that are of interest to USAID and to the Government of Rwanda. We will disaggregate the data collected at each measurement occasion by sex, urban and rural location of the school (as explained in the sampling section we will use GoR definition of urban/rural), and province. This will allow us to report more clearly on the

extent to which boys and girls in different parts of Rwanda benefitted from Soma Umenye. Given the high rate of repetition in PI, we will also disaggregate performance in PI by repeaters vs first-time students.

The Student Context Form, developed by Soma Umenye, has several sections designed to elicit information about a child and her/his background. This questionnaire will be administered to children after they will have been assessed for their reading skills, so they will remain fresh for the EGRA. Section 2 is intended to develop information on socio-economic status. Section 3 addresses repetition and prior participation in pre-school and similar activities, such as Mureke Dusome. To allow an estimate of the number of vulnerable children (orphans and children with disability) included in the IE sample,¹⁵ USAID has requested that we ask each child assessed the six questions of the Washington Group Short Set of Questions on Disability, as adapted and translated into Kinyarwanda, and this information is requested in Section 7 of the Student Context Form. Because power diminishes with every disaggregation level, our ability to detect meaningful impacts decreases as the sample size decreases with each disaggregation. Smaller groups such as composed by students with disabilities will be better served by focused targeted studies that aim to uncover impacts of Soma Umenye on subsets of the population and identify the challenges faced by its members (please refer to Section 2.9 Special Studies).

IE Question # 2. To what extent has classroom instruction changed as a result of Soma Umenye?

The literature is clear that there are strategies that teachers can utilize to improve early grader students' reading ability. We assume that Soma Umenye in-service training program is based on latest research and that teachers will be made familiar and eventually will integrate these strategies in their day-to-day teaching of reading. For example, observations of teachers by Taylor, Pearson, Peterson, and Rodriguez (2003) suggest that teaching variables such as (a) small-group instruction, (b) skill instruction in comprehension, (c) teacher modeling, and (d) coaching for teachers explained substantial variation in student achievement.¹⁶

Due to their proximity to field processes, Soma Umenye is in a privileged position to track individual teachers as they go through training and to observe their performance in the classroom. By conducting these routine teacher observations the IP will be able to document how well the intervention is taking hold and address specific issues teachers might be experiencing. For example, teachers may be unwilling or unable to adapt to new pedagogy, there might be an overall negative reaction to the changes promoted by the project, teachers may feel little incentive to adopt new instructional behaviors, or they may experience difficulty implementing the changes due to school environment and regulations.

The purpose of the IE is not to detect exact changes that occurred within specific schools or teachers. The IE seeks to understand the impact of the intervention on the whole population that received the intervention, rather than the impact on specific teachers, the latter being the domain of the IP. We will assess whether and to what extent a population—in this case PI teachers observed in 2017 —has on average changed (or not) their instructional behavior. Note that we have removed systematic differences

¹⁵ The World Health Organization estimates that 15% of any population has some form of disability, with a higher incidence of disability in countries that are post-conflict or disaster-stricken areas. This figure translates to an estimated 1.7 million Rwandans with a diversity of disabilities and needs.

¹⁶ Taylor, B., Pearson, P., Peterson, D., & Rodriguez, M. (2003, September 1). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *Elementary School Journal*, 104.

by randomly selecting sectors/schools to be assigned to the intervention and comparison groups and that in 2017 teachers will be randomly selected for observation. We will track the sample (one teacher per school included in the IE sample) of these PI teachers in both treatment and control groups through 2019. Please note that as teachers exit the intervention by transferring to a school that is not receiving the benefits of SU, for example, start teaching an upper grade such as P4 or higher, resign, or retire they are no longer part of the sample being tracked.

In September 2017 we will start to draw a profile of the group of teachers observed in the IE sample along some constructs included in the SU Lesson Observation Protocol such as Classroom Management, Time-on-Task, Delivery of Instruction, and others to be decided. At the first measure to be taken in September/October 2017 we will analyze the data collected using the Lesson Observation Protocol using descriptive statistics to profile the manner in which teachers currently teach reading to PI students. We have been working with SU to develop a protocol for class observation¹⁷ that addresses the needs of the IE and SU's own monitoring. Note that this is different from the routine observations to be conducted by the IP's MEL system, which focuses on individual teachers and tracks changes over time in order to make before and after the training comparisons. The IE compares teachers who have received the intervention and those who have not.

In 2018 and 2019 we will be able to make a comparison between teachers who have received the intervention from the beginning (treatment group) and those who have received no training or different “doses” of training. For example, in 2018 some teachers will have received the intervention for one year while others will have participated in the training program for two years. In 2019 some teachers will have been part of the program for three years while others will have received only one year of training. These comparisons will allow us to answer evaluation question # 2.

We understand that in Rwanda, as in many school systems, the reality is that teachers are reassigned, switch schools or grades, take leaves of absence, or retire from one year to the next. However, because we are conducting an Impact Evaluation with a randomized selection of sectors to form the intervention and the control groups, conditions that affect individuals will occur both in the Intervention and in the Control groups. The IE is only concerned with *a priori* conditions that *systematically* create differences between the groups. As mentioned previously, the RCT model takes into account external conditions. The random selection ensures that on the average we will find these conditions in both groups at approximately the same levels. Because teachers in treatment schools will be exposed to the intervention and teachers in the control group will not be, the impacts will still be internally valid estimate of SU's impact, irrespective of the grades they teach. However, we will carefully describe this aspect in the report.

In addition to classroom instruction as a lever for reading improvement, school environment, is highly dependent on the performance of the school principal or head teacher and on the processes that he or she has been put into place in the school in order to encourage and improve reading.

For this reason, we plan to examine and integrate data collected by SU on processes that promote reading and have been implemented in the school under the direct influence of the head teacher. For example, is there a school-wide system to record teacher and student tardiness and absenteeism? Are there incentives

¹⁷ The Observation protocol will require that observers record what they actually see happening in the classroom rather than express their opinion regarding the competence of the teacher.

for students to come to school on time? Are students who arrive on time and do not miss classes recognized? Are teachers required to record student absenteeism? How is reading encouraged in the school? What support do teachers receive to teach reading? Are there incentives for students to become better readers? We are convinced that failing to examine school processes that encourage and support improvement would be a limitation in the description of the context in which the teaching and learning to read occurs. And as we know, school principals or head teachers put school processes into place.

Evaluation findings are strengthened when several pieces of evidence point in the same direction. Often a single data set will allow a variety of impact assessments to be made, although it is better if different data sets and approaches come to broadly the same conclusion. We will coordinate with the IP's MEL and triangulate our findings and their findings—obtained through routine observations of teacher performance in the reading class—aided by the use of qualitative information (teacher and Head Teacher interviews) to reinforce findings and add depth to them. The interviews with the teacher prior to the class observation will help the assessor obtain demographic information and information related to other variables of interest related to teachers while the interview with the Head Teacher will help us to obtain information related to the encouragement and support provided to teachers in order to improve reading teaching and learning.

IE Question # 3. What is the evidence that improved student reading outcomes in Kinyarwanda lead to improved opportunities for students to succeed in schooling?

There are two major aspects to this question: Does improved reading help to improve performance in other subjects? And does improved reading help to keep a student in school and productive? To examine whether improved reading skills in Kinyarwanda increase the rate of success in other areas of schooling such as solving word problems in Mathematics, together with the EGRA. Research shows that the ability to solve mathematics problems is often limited by low levels of reading ability. Question # 3 tests the hypothesis that students' opportunities to succeed in school are enhanced as their reading competency improves. As described on page 7 of this Work Plan, IBTCI will take advantage of the opportunity offered by Dr. Michael Tusiime, DDG of REB and work with REB experts in the development of a Mathematical test.¹⁸ Comparing Math scores obtained by the treatment and the control groups we will determine whether there is a correlation between children's improved reading skills as measured by the EGRA and their ability to solve Math calculations or problems. The exam will be administered in both treatment and control schools by an IE enumerator since SU is not responsible for the administration of the numeracy measure. What we want to examine is whether after two or three years there has been sufficient impact of the program on student reading skills to spill onto other areas of schooling—specifically on the numeracy skills.

The second aspect of Question # 3 to be considered is whether schools that receive the intervention show improved attendance punctuality, retention, and grade promotion rates. Because we will be tracking PI students through 2018 and 2019, we can estimate repetition and drop-out rates of students included in the IE sample. These data will be triangulated with information provided by GoR and also gathered by the Soma Umenye MEL system to fully respond to question #3.

¹⁸ The Mathematical Assessment in Rwandan Schools tool, which was developed under USAID's Literacy, Language and Learning (L3) program was not considered suitable by REB.

IE Question # 4. To what extent are the improvements in the systemic capacity for early grade reading instruction a result of the Soma Umenye activity?

The Soma Umenye Work Plan states that the IP intends to implement activities that address perceived needs in three main areas:

- National advocacy mechanisms for early-grade reading interventions and research-based policies and curricula in support of early-grade reading instruction;
- Student and teacher performance standards and benchmarks for early-grade reading and early-grade reading assessment systems; and,
- The capacity of Teacher Training Colleges (TTCs) to effectively prepare teachers of early-grade reading areas.

Providing answers to Question # 4 starts with a desk review of relevant documents and of existing policies related to early grade reading currently in place. We are aware that the IP is planning to conduct a Gap Study that will provide insights into the difference between what is needed and what is currently in place. To avoid duplication of efforts, we plan to coordinate our activities with the IP so that we complement and expand the information that the IP collects.

At the national level, and with guidance from USAID and the IP we propose to conduct a series of interviews at the various government institutions and agencies whose work touches, in one way or another, on early literacy issues. These interviews will focus on gaining an understanding of the functioning of each institution and agency in relation to specific issues in early literacy, gathering updates on policies and programs, and discussing working relationships among institutions in order to compare actual performance with potential or desired performance. We anticipate that the sample of government institutions and agencies will include representatives from the Directorate General of Education Planning at the Ministry of Education (MINEDUC); the Education Quality and Standards Department, the Curricula, Materials Production and Distribution Department, and Teacher Development and Management Department at the Rwanda Education Board (REB); the University of Rwanda College of Education (URCE); District Education Officers; and Sector Education Officers. We will rely on USAID and the IP guidance and assistance in order to identify key informants and schedule the interviews to better understand how the national policies are operationalized. The interviews will be recorded (when we receive permission of the interviewees) and transcribed to facilitate analysis. This takes time and requires planning. The data the IE collects will be triangulated with the data obtained by the IP's Gap study.

Our Field Manager, who is Managing Director of our sub-contractor (Incisive Africa), is a resident of Rwanda and very cognizant of the educational scenario in the country; he will be a great asset as we conduct this piece of the work. This activity will start during Year I (2017) and continue throughout the life of the contract. The Team Leader and the Field Director will work together on the assessment of what is currently in place and use the information as a basis for comparison in subsequent years.

Another area to be examined is the manner in which early grade reading is assessed nationally and whether benchmarks have been established at the end of each grade, for example, number of words read correctly by grade. In 2012, EDC submitted to USAID and REB proposed benchmarks for P3 and P5 Kinyarwanda and English; however, these were never fully institutionalized. Sub-Intermediate Result 2.2 for the Soma Umenye activity is “Student and teacher performance standards and benchmarks for early-grade reading

applied.”

First, before reaching out to REB, we will request information from USAID and the IP on the existing assessment system in order to highlight successes, identify weaknesses, and make recommendations for improvement. At the same time, we will examine and document the existing initiatives and policies meant to increase time-on-task for reading and draw a profile of the current situation.

Finally, we propose to visit teacher training college programs, review current teacher-training curricula and courses offered, and conduct qualitative interviews with future teachers being trained and their instructors. From this information, we will draw a profile of the pre-service training as currently conducted, make recommendations where applicable, and use this information for later comparisons that could show changes resulting from Soma Umenye. Annex B includes the time line for the Work Plan discussed in this section.

Protection of Human Subjects

Longitudinal designs inherently require the collection of individually identifiable data in order to make tracking of respondents possible. In collecting and analyzing data from both children and adults, IBTCI staff will follow the requirements of the “Common Rule” (22 CFR 225) or of a Rwandan equivalent, whichever is more stringent, and its own “IBTCI Ethical Standards and Protocols for Field Research” (copy available on request) and will seek prior approval and annual renewals from the Rwanda National Ethics Committee (RNEC) and the National Institute of Statistics of Rwanda (NISR). We will inform respondents in appropriate ways that their participation is entirely voluntary, that records of personally identifiable information will be linked to the relevant responses only by code numbers kept separately and securely and available only to project personnel such as enumerators with a strict “need to know” basis, and that data will not be presented in a way that would allow them to be readily identified, e.g., per the 2014 Update to the Education Strategy Reporting Guidance, which states:

“De-identified = Steps have been taken to protect the privacy and anonymity of individuals and schools associated with an assessment. The implementer should work with their Institutional Review Board (IRB) to ensure that assessment participants are properly protected. Here are some examples of what should be done:

- Remove names of students, teachers, headmasters, schools. ...”

In keeping with the requirement and principles of retaining anonymity per the Reporting Guidance and the EGRA Toolkit, given that only one school is selected for each sector IBTCI will not be able to provide disaggregation on findings for levels lower than the district level.

2.7 Instruments

The key instrument to measure student outcomes is the EGRA that will be administered in September of 2017 to P1 students, September 2018 to P2 students, and September 2019 to P3 students. The IE will use the same instrument that the IP adapted to Rwanda in order to allow the IE and the IP to share data sets, conduct further analyses of the data, and compare findings.

The same is true with the Lesson Observation Protocol that the IP is in the process of developing. The IE team has already provided feedback on the draft observation protocol to the team developing the instrument and will continue to provide feedback until the instrument is finalized. The primary purpose of

this feedback is to assure that the instruments provide IBTCI with the information needed to conduct the IE. The same is true with any other instrument develop by the IP that will also needed for the impact evaluation. Table 6 below provides a list of the instruments that the IE will use.

Table 6. List of Instruments to Collect the Data Needed by the Impact Evaluation

Question # 1. Student reading skills	
Instrument	Details
1. Early Grade Reading Assessment (EGRA) sub tests: Listening Comprehension, Letter Identification, Familiar Word Reading, Syllable Pronunciation, Fluency, and Reading Comprehension 2. Student Interview Questionnaire	The IE will use the same EGRA subtask assessment tools that Soma Umenye will be administering to all students in the intervention group. <i>Please note that the IE will always be comparing student performance in the intervention group (with the program) to student performance in the control group (without the program).</i>
Question # 2. Class Observation Protocol	
Instrument	Details
3. Class Observation Protocol	The IE will use the same Observation Protocol that Soma Umenye developed to be administered to teachers in the intervention group. The IE will use the data collected to establish comparisons between the instructional performance of teachers (when teaching reading to early graders) who have received SU training to the performance of those who have not received training from SU (control group).
Question # 3. Improved opportunities of success in schooling	
Instrument	Details
4. Math test	The Impact Evaluation will use a Mathematical Assessment tool—to be developed jointly with REB in 2018—to examine whether improved reading ability results in improved ability to solve numerical Math Problems. Math competency data will be collected from P1 and P2 students in September/October 2018 and P1, P2 and P3 students in September/October 2019 for intervention and control groups.
Question # 4. Improvements in the systemic capacity to support early grade reading	
Instrument	Details
5. Desk review	Desk Review. The Impact Evaluation will conduct a desk review of

(documents, reports, policies, etc.) 6. Interviews with key informants (Interview Protocol) 7. Interviews with educators at the Ministry, REB members, and with faculty of Teachers Colleges and Universities 8. Assessment of the Teachers Colleges Curriculum (reading for early graders only) and its delivery	documents and reports available on the processes in place that help or hinder teaching of reading to early graders. We will focus on this activity in 2017, 2018 and 2019. We understand that SU plans to conduct a Gap analysis and will work with SU to avoid duplication. Key Informants With the assistance of USAID Rwanda, REB and the IP we will identify and interview key informants familiar with processes at the Ministry that are essential to reading improvement. It is possible that some of these processes may not be in place or need to be upgraded. Interviewees will include Ministry and REB key staff as well as staff at Teacher Training Colleges and at the University of Rwanda, as well as district and sector officials. Curriculum alignment. We plan to examine the alignment between the curriculum at Teachers Colleges and the skills that future teachers require (as defined by SU) in order to teach reading to early graders. To the extent possible we will also attempt to collect data to assess the competences of TTC tutors and TTC graduates/Newly Qualified Teachers.
---	--

2.8 Data Collection and Quality Assurance

Please note that the IP will collect the data in all the treatment schools while the IE is responsible for data collection at the control schools. The data collection plan is informed by the need to cover all the schools in the sample within the shortest time possible to ensure all students are at similar levels of learning. We therefore consider two weeks as the optimum period to start and finish the collection of data, with a total of 10 workdays for data collection. To ensure that the assessors have sufficient time to collect quality data, we estimate that each team will cover one single school per day. This will give the team sufficient time to adhere to the rigorous data collection procedures as envisaged in the proposal and contract. Given the probability that data collection will not go entirely as hoped, we will also allow an additional week to take care of loose ends, such as schools that may need to be revisited. Ideally, data collection for the control schools will take place simultaneously with data collection for the implementation schools and in keeping with standard EGRA practice this would be near the end of the Rwandan academic year, in September/October.

Each of the 30 IE data collection teams will comprise of seven team members, including one team leader, and a non-interviewing driver traveling in an all-wheel-drive vehicle to ensure that every team reaches its allocated school in time and without major mishaps. Each team will be allocated a zone of ten (10) sample schools, with one team being allocated eleven (11) sample schools, to ensure that the sample is covered within the ten-day data collection timeframe.

Assessors will have a minimum of two years of university education and will participate in one week of specialized EGRA assessment, observation, and interviewing training. When recruiting assessors, we will consider recruiting people that have conducted EGRA/EGMA before and strive to maintain a balanced man/woman mix while selecting the best qualified among them. Assessors must be fluent Kinyarwanda speakers, and they should have experience working with children. To promote consistency among

assessors recruited by Soma Umenye to collect data from intervention schools and assessors recruited by the IE to collect data from control schools, we plan to have training for both sets of assessors conducted jointly; IBTCI and Chemonics have already held discussions on how this can be carried out, including preliminary discussions as to sharing costs. We envisage that each team will conduct at least two full-length pilot tests in real scenario schools to ensure dexterity and fluency during data collection in the IE sample schools.

All data will be collected via tablets to ensure rapid turnaround; we will confer with Soma Umenye to ensure that tablets and software are compatible. Assessors will send data to a central server every day immediately after data collection to avoid data loss. The assessors will be instructed on back-up procedures whenever immediate transmission of the data is not possible due to field conditions. Data will be transmitted to a secure encrypted cloud server via a virtual private network (VPN) to ensure data integrity, and will only be available for download to authorized individuals via use of tamper-proof passwords and authentication. This ensures that data are available to authorized staff as closely to real-time as possible. Each assessments and observation collected will bear start and end time stamps as well as GPS coordinates of the location where it is conducted. Procedures for ensuring data are properly uploaded or backed up will be the same during both field-testing and full data collection. Once uploaded to the server, the data will be aggregated and downloaded by the data management team (Incisive Africa) and checked for inconsistencies and outliers. The data will then be cleaned and the data sets shared with all members of the IE team for analyses. The Team Leader and the Methodologist will lead the data analyses effort, the reporting of the findings, and the preparation of reports.

A detailed description of how to conduct the data collection activity will be part of the assessors' training manual that we will prepare in advance of training. In the manual, the data collection procedures will be developed in detail and expanded. The assessors will be provided with a script to guide their initial interchanges with the Head Teacher, the teacher to be observed and the student who will be assessed. The same data collection procedures that are utilized at baseline will be used at midline and end line to ensure continuity and comparability. Annex E provides an example of the level of detail for procedures that assessors will need to follow as they visit schools to collect data. This necessary level of detail will promote consistency and comparability. The same detailed procedures will be used any time an instrument is administered.

The Team Leader and the Field Director will share the responsibility of interviewing key informants from MINEDUC, REB, Province and District education officials, and other informants from Teachers Colleges whose perspectives are essential to address evaluation Question # 4. Both are experienced researchers and familiar with interviewing as a data collection method. Their experience will ensure the quality of the information collected.

The chart that follows summarizes our analytical approach and the proposed timeline for the data collection.

Table 7. Analytical Approach to Evaluation Questions

Impact Evaluation Approach		
Research Question	Analytical Approach	Timeline for data collection
1. To what extent are changes in the Kinyarwanda reading outcomes for students in Grades 1-3 attributable to the Soma Umenye activity (as a package)? <i>a. To what extent has USAID Soma Umenye had an impact on various categories of especially vulnerable children?</i> <i>b. To what extent have students in different parts of Rwanda benefitted from Soma Umenye in comparison to each other? *</i>	Compare outcomes of students in the intervention group to outcomes of those in the comparison group (counterfactual). Outcomes will be disaggregated by sex, urban/rural location, province, and other variables. When verified that the size of the population is too limited to allow statistical differences to be of significance—for example, children with disabilities—these populations could become the object of special studies.¹⁹	September 2017 (P1) September 2018 (P1 and P2) September 2019 (P1, P2, and P3)
2. To what extent has teaching of reading changed as a result of the intervention?	Compare the instructional performance of teachers in the intervention group to that observed in comparison group. The comparison of EGRA scores obtained by students in the intervention and in the control groups will enable us to make an assumption regarding the changes in the teaching of reading.	September 2017 September 2018 September 2019
3. What is the evidence that improved student Kinyarwanda reading outcomes lead to improved opportunities for students to succeed in schooling? <i>a. To what extent do the reading results of Soma Umenye students correlate with student repetition and dropout vis-à-vis control students?</i> <i>b. To what extent do the reading results of Soma Umenye students correlate with their apparent results in other subjects vis-à-vis the control students?</i>	Compare Math test scores obtained by students in intervention and in control groups Compare the treatment and control groups on variables such as: - Punctuality, absenteeism, repetition, and drop-out rates - Time available for the teaching of reading (time-on-task).	September 2017 (P1) September 2018 (P1 and P2) September 2019 (P1, P2, and P3) September 2017, 2018, and 2019
4. To what extent are the improvements in the systemic capacity for early grade reading instruction a result of the Soma Umenye activity? <i>a. Strengthening a system for delivery of</i>	Interview key informants and examine Soma Umenye documents (such as the Gap Study) as well as other related documents to assess of the situation prior to project implementation and to record progress in the areas of early	Starting in 2017 (after approval of the Work Plan) and at various points throughout the life of the project as Soma Umenye

¹⁹ Please see section 2.9 for an initial description of the special studies.

<i>teacher continuing professional development (CPD) for early grade reading?</i> <i>b. A national assessment system for early grade reading?</i> <i>c. A teacher training college program that prepares relevant graduates to be effective teachers of early grade reading?</i> <i>d. Policies that increase time-on-task for reading?</i>	grade reading that could be attributed to Soma Umenye activities. - Teacher continuing professional development - National assessment - TTC and URCE programs focused on early grade reading - Policies that increase time-on-task for reading	conducts activities described in Question # 4 and the IE contract is active.
--	--	--

We will be able to longitudinally track the cohort of students entering PI in 2017 for three years and the cohort of students entering PI in 2018 for 2 years. This will allow us to make several comparisons to examine impacts of the Soma Umenye intervention. For example, at the end of 2019 (end line), students in the treatment group entering in PI in 2017 (shown in blue in Table 7) will have received Soma Umenye benefits for three years. The control group students of the same cohort, on the other hand, will have received no benefits in 2017 and 2018 but will have been in the program for one year at the end of 2019. Thus the evaluation design will allow for multiple comparisons such as PI with the program and PI without the program; two years in the program versus no program; three years with the program versus only one year of benefits. These comparisons will clarify the various facets of the impact of Soma Umenye. The number of students assessed will remain the same each year for the cohorts longitudinally tracked to the extent that attrition from the longitudinal sample is not an issue. But because sectors have been randomly selected, this condition would not cause a systematic difference and would affect equally treatment and control groups. Table 8 clarifies the number of students, disaggregated by membership in Treatment or Control Schools.

Table 8. Students to be Assessed as part of the Impact Evaluation

Year	P1		P2		P3		Total		Grand Total
	T	C	T	C	T	C	T	C	
2017	2,448	3,024					2,448	3,024	5,472
2018	2,448	3,024	2,448	3,024			4,896	6,048	10,944
2019	2,448	3,024	2,448	3,024	2,448	3,024	7,344	9,072	16,416
Total	7,344	9,072	4,896	6,048	2,448	3,024	8,160	10,080	18,240

Note: Students in the same cohort are indicated by the same color code across different years. T=Treatment and C=Control.

2.9 Data Analysis

Quantitative information

The IE will manage and analyze quantitative data from the EGRA and the Math assessment to analyze

impacts at the student level. The main goal of the impact analysis is to assess the effects of the interventions on student outcomes such as reading ability. Random assignment ensures that there are no systematic differences between the research groups other than access to the SU intervention. Consequently, the difference between the mean of the outcome of interest of the intervention group and control group following implementation of the intervention provides an unbiased estimate of the impact of SU. In addition, because we are using a longitudinal sample, we will be able to estimate the impact of SU at different time points as the implementation progresses and students continue to be exposed to the intervention. In other words, we will be able to examine how the impact of the SU intervention (difference in levels between the intervention and the control group) changes over time (changes in the level difference).

Although the difference in mean of the outcome for the intervention and the treatment groups provides an unbiased impact estimate, we will use a multivariate statistical model to estimate program impacts for two reasons:

1. **Correctly specify the precision of impact estimates.** Although random assignment is conducted at the sector level and implementation is carried out at the school level, outcomes are measured at the student level. This situation can cause us to overstate the precision of impact estimates if it is not addressed statistically. We will use a regression specification that will correctly estimate standard errors of impact estimates.
2. **Improve the precision of impact estimates.** Despite random assignment, it is possible that small differences in observable baseline characteristics across research groups may arise by chance or from missing data. The statistical model controls for any such differences in observable characteristics. Estimating program impacts in this way will improve the precision of the impact estimates (Raudenbush 1997).

The impact analysis will rely on a linear regression model. This model can be used for both continuous and binary outcomes. It can be expressed as follows:²⁰

$$(1) \quad y_{iscft} = \alpha + \beta x_{iscf0} + \gamma z_{s0} + \delta T_s + B_s + \mu_f + \tau_s + \varepsilon_{ist}$$

where y_{iscft} is the outcome of interest for student i in cohort c in school s , teacher f , and at time t (such as correct words per minute). The vector x_{iscf0} represents the baseline (time 0) characteristics of student i in cohort c and school s , which may include age, gender, or other factors²¹. The vector z_{s0} represents the baseline characteristics of school s , such as enrollment, rural/urban status of school, or school-level baseline measures from the same domain as outcome y (if available).²² The variable T_s is an indicator equal to one for students in treatment group schools and zero for those in control group schools. The term B_s represents a set of dummy variables indicating the randomization block²³, μ_f is a set of dummy variables indicating teacher specific fixed effects, τ_s is a school-specific error term (a group or cluster effect), while ε_{ist} is a random error term for student i in school s observed at time t .

The parameter δ is the regression-adjusted difference in mean outcomes between the treatment and the control groups, which gives the impact of SU on the outcome of interest. The program effect estimates provide what is known as the “intent-to-treat” (ITT) effect. It estimates the average difference in outcomes between all students in target grades at time t in treatment and control schools, regardless of whether the students in the treatment schools participated in program activities (by attending class everyday) and regardless of whether the school implemented the intervention. Such an estimate can be interpreted as the average impact on the population of students in schools being given the option to implement the intervention. This is the key parameter of interest, since from a policy perspective this shows the impact of introducing the intervention into schools. This will not necessarily involve take-up by all students or full implementation by the school.

How significant are the effects of the intervention? For the impact estimate of each outcome, we will use a two-tailed t-statistic to test the null hypothesis that there is no difference between the regression-adjusted means for the treatment and control groups. We will use the associated p-value, which reflects the probability of obtaining the observed impact estimate when the null hypothesis of no effect is true, to judge the likelihood that SU had a statistically significant impact. Also, it is important to understand that not all impacts are the same and that for different reasons some interventions impact the beneficiaries more than others. In order to assess the size of the effects observed, we will use Cohen’s d statistics.²⁴

²⁰ For binary outcomes, this specification is referred to as a linear probability model. For binary variables, we will also estimate a logistic regression model as a robustness check. These two approaches tend to produce similar findings if the mean of the outcome is not close to 0 or 1. The linear probability model is the preferred specification because of its ease of interpretation.

²¹ Ideally, we will include baseline of outcome y for each student t . However, we will not have baseline data for any of the outcomes on PI students related to academic performance as beginning-of-the year academic performance test through EGRA or a math assessment was not administered.

²² The model can also be run without the student or the school-level baseline variables. However, we gain precision when these variables are included. If these data are not available or not included, the model would not include the x or z terms.

²³ Recall that randomization was carried out within two blocks in each district, one that contained sectors with more than the district average number of schools (larger sectors) and one that contained less than or equal to the average number of these schools in the district. Because the blocks were constructed using pre-existing sector characteristics to match them as closely as possible, controlling for block fixed effects are appropriate for this stratified random assignment design (Duflo et al., 2008).

²⁴ For Cohen’s d an effect size of 0.2 to 0.3 might be a “small” effect, around 0.5 a “medium” effect and 0.8 to infinity, a “large”

Multiple Comparison Issue.

As the number of estimated impacts increases, it becomes more likely that we will find a false positive, meaning that we will conclude there is an effect when there is none. This fact is often referred to as the multiple comparisons problem. When only one hypothesis test is conducted, it is common to conclude that an impact is statistically significant when the probability that it is due to chance (and not to the program) is 5 percent. However, when using a 5 percent significance threshold and more than one test is conducted at the same time, the probability of a false positive—that is the probability that a significant impact is due to chance and not to the program—is larger than 5 percent. For example, if we estimate impacts on “correct word per minute” separately for each of three grades in 2019 and do not adjust for multiple comparisons, the probability of a false positive is 14 percent.²⁵ In addition to analysis of multiple outcomes, multiple comparison concerns apply to analysis of multiple subgroups as well.

The central element of our recommended strategy for addressing multiple comparison concerns is to focus the primary analysis of program effectiveness in each domain, such as reading ability or math aptitude, on a single key outcome identified prior to beginning the analysis.²⁶ Using a small set of outcomes within each domain makes it less likely that statistically significant findings will emerge by chance. Limiting the primary analysis to a single outcome means that it will not be necessary to do statistical corrections for multiple comparisons. Selecting the primary measures before beginning analysis prevents focusing the assessment of program effectiveness on outcomes that happen to emerge as statistically significant (or the perception that this may have been the case). We will work with USAID, the IP, and REB to identify whether there are key primary outcomes in each domain that we can focus on before we begin the data analysis.

However, if we do not identify a key primary outcome and thus there are more than one primary outcome per domain, we would use the Benjamini-Hochberg corrections. Other common corrections include the Bonferroni and the Tukey-Kramer adjustment. For a detailed discussion of these methods and its advantages and disadvantages see Schochet (2009). Note that, any correction for multiple comparisons reduces statistical power to identify significant effects. Thus, we will only apply the multiple comparison corrections to the full sample as *Soma Umenye*'s primary target is the entire primary grade student population in Rwanda and not any specific subgroup. This approach is in line with the recommendations provided in Schochet (2009).

Teacher observation data. As mentioned before, the IE will use the same instrument developed by the IP and administered to the treatment group. Observation data collected in September 2017, 2018, and 2019 will be analyzed using descriptive statistics (percentages, standard deviations, etc.) to provide a description of how teachers conduct the reading class. Our hypothesis is that with adequate training provided by Soma Umenye there will be significant statistical differences between teachers who received SU training and those that did not.

Our experience in the evaluation of projects that conduct teacher in-service training to improve student performance tells us that it takes time for innovations to be fully accepted. Some teachers have been teaching reading in a certain manner for many years and may not be sure of the benefits to be accrued by

effect. (Cohen's d might be larger than 1.)

²⁵ The probability of finding a false positive when conducting n independent two-tailed t tests, given the null hypothesis is true for each test, is found using the formula $1 - (1 - \alpha)^n$, where α is the significance threshold used in the test.

²⁶ This multiple comparison framework is consistent with that suggested by Schochet (2009).

changing their instructional behavior. This, combined with the obstacles encountered during the first year of project implementation such as limited time available for training during Year I of Soma Umenye, leads us to hypothesize that in 2017 we will not be able to observe much difference between the instructional behaviors of teachers in the intervention and in the comparison groups. As teachers who receive the intervention become more comfortable with using new strategies and see the benefits to the students, the tendency is for the gap between the two groups to increase.

Qualitative Information

Good evaluations are almost invariably mixed method evaluations. Qualitative information informs both the design and the interpretation of quantitative data. Many evaluations under-exploit qualitative methods, both in the techniques they use and the way in which analysis is undertaken.

Interview data. When the interviewee grants us permission, we will record the semi-structured qualitative interviews to ensure that no details are missing. If the interviewee objects to the recording, we will take notes. In both cases the interviews will be transcribed in a format that facilitates analysis. To analyze the data obtained through semi-structured qualitative interviews, we will take the following steps:

1. Read the transcribed interviews or the interview notes and identify recurrent themes or the idea categories that emerge from the data;
2. Note patterns in the data by examining the content of each response in order to categorize verbal data for the purpose of classification, summarization and tabulation;
3. Code the data by attaching labels to the lines of text in order to group and compare similar or related pieces of information and then compile similar blocks of text from different sources into a single file; and,
4. Search for answers to the evaluation questions.

2.10 Special Studies

There are many cases where the size of the population does not allow significant comparisons to be drawn. For example, on a large sample such as the IE sample (2,910 students) disaggregating data by treatment and control or by sex allows the testing for significance. In this and similar cases if significant differences exist between intervention and comparison groups they will easily documented.

In other cases such as students with disabilities, the numbers of students may not be sufficient to give us significant differences because the power diminishes with every disaggregation level. For example, according to Tables 3.7 and 3.9 of the 2015 Statistical Yearbook, 4,871 of the 639,656 PI students (0.76%) enrolled in 2015 had a disability of some sort.²⁷ With small subsets of a population, we will not be able to detect impact or the effect sizes. Assuming even distribution, with a sample size of 2,910, we would assume a total of approximately 22 individual students in the total sample of children in intervention and control schools, hardly adequate to draw useful conclusions. These smaller groups will better served by focused special studies with specific methodology designed to give us insights related to the performance to expect and the challenges faced by its members.

²⁷ Students from private schools are included in these figures.

Although at this point we do not have data to make decisions regarding additional topics well suited for special studies, our experience allows us to point to some important areas that may be viable for further study. Examples of possible areas for special studies are presented below. We anticipate that these studies will be conducted after the 2017 and the 2018 data collection and analyses, so special studies will be more definitively elaborated in future workplans.

The performance of children with disabilities: What is the impact of Soma Umenye on students with disabilities?

Based on MINEDUC's figures, we can expect very few children with disabilities as part of our IE sample. We will, therefore, go beyond the IE sample to obtain information on children with disabilities, using at least three sources of data to identify children with disabilities in the relevant grade level of the schools which we are visiting. Please note that extra sensitivity is needed to identify and track children with disabilities, and we will need to confer with Disabled People's Organizations and relevant social service agencies with respect to potential referrals should we identify a child with a disability.

1. Before conducting the student interview, the assessor will check the statement *Describe in the space below if the student has any observable physical disability*. In the course of conducting interviews with students selected for assessment, the assessor will ask each child the six-question Washington Group Short Set of Questions on Disability. Regardless of response, the child will remain part of the IE sample and will be tracked.
2. As part of the teacher interview, s/he will be asked if there are other children with apparent disabilities in the class. Any child identified will be tracked from one year to the next as to persistence in school.
3. Teachers of other PI classes will be asked if there are children with apparent disabilities in their classes. Any child identified will be tracked from one year to the next as to persistence in school.

Analysis of dropouts.

According to Table 4.2 of the 2014 Education Statistical Yearbook, on a national basis at primary level, most relevantly for Soma Umenye and specifically at early grade level, 2013 drop-out rates were significantly worse for boys than for girls, overall drop-out is 10.2% for P1, 13.8 for P2, 12.7% for P3, and 12.9% for P4,²⁸ and there are significant variations across districts. There would be merit in attempting to ascertain differences in dropout rates, if any, between treatment and control schools, both during the school year and between school years, and possible causes, particularly in the most severely impacted sectors.²⁹

Do schools that receive food support perform better than those that do not?

One of the subgroups of interest for the Soma Umenye (SU) impact evaluation is students in schools participating in school-based feeding program. The World Food Programme (WFP) and its sub-grantee, World Vision International, are implementing the "Home-Grown School Feeding Program" (HGSF) funded by the U.S. Department of Agriculture (USDA) in 104 schools in 16 sectors in four districts

²⁸ There do not seem to be comparable statistics in the 2015 Statistical Yearbook.

²⁹ While the data on repetition between the L3 report and the Statistical Yearbook are consistent, the data on drop-outs are not. Perhaps the substantially lower L3 drop-out rate refers to children who drop out during the school year and does not reflect children who complete one grade but who do not return to school to begin the next grade.

(Karongi, Nyamagabe, Nyaruguru, and Rutsiro) in the Western and Southern Provinces of Rwanda. However, there are only eight sectors that overlap with the Soma Umenye impact evaluation study sample — four in the treatment group or Phase I sectors (Ruhango, Musange, Kamageri, and Kivu) and four in the control group or Phase 3 sectors (Murundi, Kivumu, Gasaka, and Ngoma).³⁰ All schools in these eight sectors, 48 in total, will receive the HGSF program.

The sampling plan for the SU impact evaluation, as discussed in Section 2.3, calls for the random selection of one school per sector in each of the Phase I sectors (intervention) and Phase 3 sectors (control) for inclusion in the SU impact study. Thus, our sampling plan only provides a maximum sub-sample of four schools (from the intervention group) that will have both SU and HGSF program starting in 2017 and a sub-sample of four schools (from the control group) that will have “HGSF only” in 2017 and 2018 before both programs are implemented in these sectors in 2019. An analysis to estimate the impacts of SU for this subgroup to examine whether school-based feeding has different (and possibly larger) impacts, compared to the overall impact of SU, will lack the required statistical power given this small sample size. In fact, this small sample size, divided across intervention and comparison, is of limited value even in providing substantive descriptive results.

Therefore, we propose examining the additional impacts of the HGSF program with a separate sample, in addition to the sample of schools that will be selected for inclusion in the SU impact evaluation. While data will be collected at the same time as data are collected for the impact evaluation, a separate sample will keep the SU impact evaluation plan unaffected, while providing a larger sample for examining the potential additional impacts of the HGSF program. Because we want to examine the additional impacts of the HGSF program on top of the SU intervention, the sample for this analysis will be drawn from the three districts where HGSF is being implemented and will involve the sectors that are part of the SU intervention group (Phase I sectors).³¹ Specifically, we will compare outcomes between students in schools in the following two groups of sectors:

- 1) Group 1: All schools in the four sectors where both SU and the HGSF programs are being implemented in 2017. There are 27 schools in these four sectors.
- 2) Group 2: All schools in an additional four sectors, where only SU will be implemented in 2017.³² These four sectors will be purposively selected so that they are comparable to the four sectors in Group 1. We will explore administrative data on sector characteristics (such as the number of schools, distance from the capital, etc.) to select these four comparison sectors.

Comparing outcomes between Group 1 (SU+HGSF) and Group 2 (SU only) will give us an estimate of

³⁰ There are eight Phase 2 sectors where the HGSF program is implemented as well. However, Phase 2 sectors are not part of the SU impact evaluation study sample to avoid possible contamination and, therefore, are not considered for inclusion in this special study.

³¹ HGSF is being implemented in four districts. However, none of the sectors in Rutsiro where HGSF is implemented is part of a SU intervention group.

³² There are 19 sectors in the three HGSF districts (Karongi, Nyanagebe, and Nyaruguru) where SU is being implemented in 2017. Excluding the four sectors where HGSF is being implemented, there will be 14 sectors remaining to select these additional four sectors from. It is the use of these four sectors where HGSF is not being implemented, making them appropriate comparison sectors, plus the fact that we are conducting a census rather than a sample of schools in sectors that receive both HGSF and SU interventions in 2017 that helps to differentiate this study from the overlapping performance evaluations that WFP began conducting of its program, starting in 2016.

additional impacts, if any, of the HGSF program. We will analyze outcomes for P1 students in 2017 (who will receive the SU intervention in 2017), outcomes of P1 and P2 students in 2018 (who will receive the SU intervention in 2018, and outcomes of P1-P3 students in 2019 (who will receive the SU intervention in 2019). However, because of the non-random placement of the HGSF program, these impacts are non-rigorous and suggestive. Also, while this additional sampling provides a larger sample compared to the sample we will have under the SU impact evaluation sampling plan, it is still small as it is constrained by the number of sectors where the HGSF program is implemented within the SU intervention group.

2.11 Cost and Cost Effectiveness of the Intervention

In addition to providing answers to the four Impact Evaluation questions and sub-questions, IBTCI is called on “to provide recommendations to improve the potential for the sustainability of Soma Umenye activities and outcomes.” A key aspect of this is the cost and budgetary aspects of sustaining activities similar to Soma Umenye once USAID funding comes to an end. Therefore IBTCI will be taking into consideration any cost capture and reporting requirements in Chemonics’ contract.³³

IBTCI has taken part in training by USAID’s Education Office in the new cost reporting guidance, and conferred with Chemonics and reviewed their intended cost recording structure for its suitability and adequacy. As Soma Umenye progresses, IBTCI will work with Chemonics to assemble from their records the data needed to conduct the cost-effectiveness analyses of the Soma Umenye activity. We need to take cost data for the intervention group, and divide it by the number of beneficiaries—students in the intervention group (IE sample) to arrive at unit costs. This will provide the cost for the absolute and relative gain in EGRA scores in 2017. We will repeat and extend this analysis to include the 2018 and 2019 school years. At end line, effectiveness will be determined by grade for each of the EGRA subtests in terms of absolute gains relative to the comparison group.

2.12 Coordination with Other Implementers

While our primary interactions will be with Soma Umenye, we are well aware that there are other activities supported by USAID, USDA, other donors, and the GoR itself that are likely to interact with Soma Umenye now and in the future. The IBTCI team had the opportunity to be introduced to several of these other education partners during the October-November 2016 mobilization trip and intends to interact more substantively with the education partners after approval of the Evaluation Design subject to the guidance of our COR and preferably in coordination with the Soma Umenye implementer. While details would be handled on a case-by-case basis, in general we will advocate that interventions provided by other donors be provided to both intervention and control sectors comparably and that we be kept informed as to of what the interventions consist.

³³ IBTCI participated in the USAID/Office of Education’s workshop on *Cost Reporting Guidance in USAID-funded Education Activities* held in early February 2017.

SECTION 3. DISSEMINATION OF FINDINGS

Because we anticipate that analyses of the data collected in September 2017 will disclose primarily only what the status quo is of PI children with limited if any exposure to Soma Umenye, we anticipate that dissemination will take place in early 2018 and will be done primarily through reports to stakeholders that will include USAID Rwanda, the IP, REB, the education community at the central level, colleges and universities, and researchers.

In 2018 and 2019, as more relevant findings are available, we will add more diversified strategies such as links to the evaluation results on relevant websites; press releases; media dossiers; and, highlighting of key findings in the media. Under the leadership of USAID Rwanda and REB, we will identify people and organizations in Rwanda that could potentially be idea champions of early grade reading. As the IE progresses to Midline and End line, donor agencies, national and international NGOs/FBOs, CBOs, local media, community leaders and other members of civil society will be included.

The information obtained during the IE will be disseminated to the different audiences in accordance with their specific interest. Information is power and the needs of each of the target audiences will be considered in order to provide information that will empower them to make decisions and advocate for more quality and more relevant reading strategies. For example, as we share findings of the 2017 data collection, the REB and the Ministry of Education may be interested in understanding children's level of mastery of pre-reading skills such as Letter Identification, and Familiar Word Reading since early assessment of the pre-reading and foundational skills required for fluency allows the implementation of measures to correct deficiencies where they exist ³⁴ while Teacher Training Colleges may be interested in the results of class observation and on how teachers are actually teaching skills essential to reading to PI students. This becomes important when we learn that the average repetition rate is 25.7%, in PI.³⁵

Reaching the target audiences involves a three-step process: presenting fresh and relevant information in a timely manner; packaging information in user-friendly formats; and disseminating information to the target audiences at the level of details that makes it actionable. We propose to engage key stakeholders in being message multipliers, so that an increasingly broad audience learns how the results Impact Evaluation can be used to guide educational policies and actions. The purpose is to encourage debate, problem solving and joint ownership of plans to address the opportunities and challenges identified by the Impact Evaluation.

³⁴ Abadzi, Helen (2009) "Instructional Time Loss in Developing Countries: Concepts, Measurement, and Implications." World Bank Research Observer. 24 (2):267-290.

³⁵ GoR Yearbook Report 2013

ANNEX A. SOMA UMENYE IMPACT EVALUATION STATEMENT OF WORK

ANNEX B. YEAR I WORK PLAN

The following work plan reflects activities anticipated from inception of the Task Order through September 30, 2017.

Year I - USAID/Soma Umenye Impact Evaluation Activities

IMPACT EVALUATION ACTIVITY	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep
Gather information needed to contextualize the evaluation—desk review of documents	■	■										
Travel to Rwanda to collect information, meet with USAID, REB, Chemonics, subcontractor, and other relevant stakeholders.	■	■										
Meetings and conference calls with USAID, Chemonics, EdIntersect, REB and other relevant stakeholders.	■	■	■									
Finalize IE design in collaboration with the implementer and USAID. Make recommendations regarding phasing of implementation and sample selection.		■	■	■	■	■	■					
Finalize IE Work Plan and submit together with IE design to COR		■	■	■	■	■	■	■	■			
Design the strategy and assist the IP in the randomizing of sectors to Phases 1, 2 and 3 of implementation			■									
Present Work Plan and Design to REB, if requested										■	■	
Participate in Tools Workshop, conduct initial research on teacher training and attitudes, and dialogue with REB and other stakeholders.							■	■	■			
Collaborate with the IP to develop/adapt the package of instruments for the September/October 2017 assessment:						■	■	■	■	■	■	■
Collect school rosters								■	■			
Participate in piloting of instruments									■		■	
Collaborate with the IP, if requested, in getting necessary approvals.									■			
Select assessors to conduct data collection in the IE sample									■			

Soma Umenye Impact Evaluation – Draft Year I Workplan and Evaluation Design

Train assessors to collect the data on intervention and control schools on the IE sample (anticipated dates Sept. 13-23)												■
Finalize logistics arrangements and preparations for the collection of baseline data for the IE											■	■
Collect EGRA and math data on PI control students on the IE sample. (anticipated dates Sept. 25-Oct. 13)												■
Meet with REB to plan the development of the Math assessment tool												■
Assess current approaches to pre-service training in early grade reading in order to provide answers to Evaluation Question # 4										■	■	
Prepare Year 2 Work Plan										■	■	
Submit Year 2 Work plan												■

ANNEX C.**LIST OF SECTORS IN PHASES 1, 2, AND 3**

Northern Province			
District	Phase 1	Phase 2	Phase 3
1. Rulindo	Pot 1 Sectors		
	1. Buyoga	1. Bushoki	1. Cyinzuzi
	2. Rukozo	2. Kisaro	2. Mbogo
	3. Shyorongi	3. Tumba	3. Murambi
	Pot 2 Sectors		
	4. Base	4. Cyungo	4. Burega
	5. Kinihira	5. Ngoma	5. Masoro
	6. Ntarabana		6. Rusiga
2. Gakenke	Pot 1 Sectors		
	1. Gakenke	1. Busengo	1. Muzo
	2. Ruli		2. Rushashi
	Pot 2 Sectors		
	3. Coko	2. Gashenyi	3. Karambo
	4. Cyabingo	3. Kamubuga	4. Mataba
	5. Janja	4. Mugunga	5. Minazi
	6. Kivuruga	5. Rusasa	6. Muhondo
			7. Muyongwe
			8. Nemba
3. Burera	Pot 1 Sectors		
	1. Kinyababa	1. Kivuye	1. Butaro
	2. Nemba	2. Rugarama	2. Cyanika
	Pot 2 Sectors		
	3. Bungwe	3. Gahunga	3. Cyeru
	4. Ruhunde	4. Gatebe	4. Gitovu
	5. Rwerere	5. Rugengabari	5. Kagogo
			6. Kinoni
			7. Rusarabuye
4. Musanze	Pot 1 Sectors		
	1. Cyuve	1. Kinigi	1. Muhoza
		2. Musanze	2. Nkotsi
	Pot 2 Sectors		
	2. Gataraga	3. Busogo	3. Gacaca
	3. Nyange	4. Gashaki	4. Kimonyi
	4. Remera		5. Muko
	5. Shingiro		6. Rwaza

5. Gicumbi	Pot 1 Sectors		
	1. Bukure	1. Nyankenke	1. Bwisige
	2. Cyumba	2. Rushaki	2. Byumba
	3. Mutete	3. Rutare	3. Kaniga
	4. Rukomo	4. Ruvune	4. Miyove
			5. Muko
	Pot 2 Sectors		
	5. Giti	5. Kageyo	6. Manyagiro
	6. Nyamiyaga	6. Mukarange	7. Rubaya
	7. Rwamiko		8. Shangasha

Eastern Province			
District	Phase 1	Phase 2	Phase 3
1. Bugesera	Pot 1 Sectors		
	1. Nyamata	1. Ngeruka	1. Juru
	2. Rweru		2. Mayange
	Pot 2 Sectors		
	3. Mwogo	2. Gashora	3. Kamabuye
	4. Ntarama	3. Musenyi	4. Mareba
	5. Rilima	4. Ruhuha	5. Nyarugenge
			6. Shyara
2. Gatsibo	Pot 1 Sectors		
	1. Gatsibo	1. Gitoki	1. Muhura
	2. Kageyo	2. Kabarore	2. Nyagihanga
	3. Kiziguro	3. Kiramuruzi	3. Rugarama
			4. Rwimbogo
	Pot 2 Sectors		
	4. Remera	4. Murambi	5. Gasange
			6. Ngarama
3. Kayonza	Pot 1 Sectors		
	1. Gahini	1. Rukara	1. Murundi
	2. Kabare		2. Mwiri
	Pot 2 Sectors		
	3. Murama	2. Ndego	3. Kabarondo
	4. Rwinkwavu	3. Nyamirama	4. Mukarange
			5. Ruramira
4. Kirehe	Pot 1 Sectors		
	1. Musaza	1. Gahara	1. Kirehe
	2. Nasho	2. Mushikiri	2. Mpanga
	Pot 2 Sectors		
	3. Mahama	3. Kigarama	3. Gatore
	4. Nyamugali		4. Kigina
			5. Nyarubuye
5. Ngoma	Pot 1 Sectors		
	1. Kibungo	1. Murama	1. Kazo
	2. Mugesera		2. Rukira
	Pot 2 Sectors		
	3. Mutenderi	2. Jarama	3. Gashanda
	4. Sake	3. Karemba	4. Rukumberi
		4. Remera	5. Rurenge
			6. Zaza

6. Nyagatare	Pot 1 Sectors		
	1. Karangazi	1. Rwempasha	1. Mukama
	2. Tabagwe	2. Rwimiyaga	2. Nyagatare
	Pot 2 Sectors		
	3. Gatunda	3. Karama	3. Kiyombe
	4. Katabagemu	4. Matimba	4. Mimuli
			5. Musheru
			6. Rukomo
7. Rwamagana	Pot 1 Sectors		
	1. Muhazi	1. Nyakariro	1. Kigabiro
	2. Mwulire		2. Munyaga
	Pot 2 Sectors		
	3. Muyumbu	2. Gahengeru	3. Fumbwe
	4. Nzige	3. Karenga	4. Gishali
		4. Musha	5. Munyiginya
			6. Rubona

Southern Province			
District	Phase I	Phase 2	Phase 3
1. Gisagara	Pot 1 Sectors		
	1. Muganza	1. Mamba	1. Gikonko
	2. Mugombwa	2. Ndora	2. Gishubi
	3. Save		3. Kansi
	Pot 2 Sectors		
	4. Kigembe	3. Mukindo	4. Kibirizi
	5. Nyanza		5. Musha
2. Huye	Pot 1 Sectors		
	1. Ruhashya	1. Kigoma	1. Maraba
	2. Rusatira		2. Mbazi
	Pot 2 Sectors		
	3. Karama	2. Ngoma	3. Gishamvu
	4. Kinazi	3. Rwaniro	4. Huye
	5. Mukura		5. Simbi
3. Muhanga	Pot 1 Sectors		
	1. Kibangu	1. Nyamabuye	1. Nyarusange
			2. Rongi
	Pot 2 Sectors		
	2. Cyeza	2. Kiyumba	3. Mushishiro
	3. Kabacuzi	3. Shyogwe	4. Nyabinoni
	4. Muhanga		5. Rugendabari
4. Kamonyi	Pot 1 Sectors		
	1. Kayenzi	1. Gacurabwenge	1. Nyarubaka
	2. Mugina		2. Rukoma
	3. Ngamba		
	Pot 2 Sectors		
	4. Rugarika	2. Kayumbu	3. Karama
		3. Runda	4. Musambira
5. Ruhango	Pot 1 Sectors		
	1. Byimana	1. Ruhango	1. Bweramana
	2. Kinazi		2. Ntongwe
	Pot 2 Sectors		
	3. Mbuye	2. Kinihira	3. Kabagali
6. Nyanza	Pot 1 Sectors		
	1. Busasamana	1. Mukingo	1. Busoro

	2. Cyabakamyi	2. Nyagisozi	2. Muyira
	Pot 2 Sectors		
	3. Rwabicuma	3. Ntyazo	3. Kibirizi
			4. Kigoma
7. Nyaruguru	Pot 1 Sectors		
	1. Busanze	1. Ruramba	1. Ngera
	2. Kibeho		2. Ngoma
	Pot 2 Sectors		
	3. Cyahinda	2. Nyabimata	3. Muganza
	4. Kivu	3. Ruheru	4. Munini
	5. Mata		5. Nyagisozi
			6. Rusenge
8. Nyamagabe	Pot 1 Sectors		
	1. Mugano	1. Tare	1. Gasaka
	2. Musange	2. Uwinkingi	2. Kaduha
			3. Kitabi
	Pot 2 Sectors		
	3. Buruhukiro	3. Cyanika	4. Gatare
	4. Kamegeri	4. Kibirizi	5. Kibumbwe
	5. Musebeya	5. Mbazi	6. Mushubi
			7. Nkomane

Western Province			
District	Phase I	Phase 2	Phase 3
1. Nyamasheke	Pot 1 Sectors		
	1. Macuba	1. Gihombo	1. Kagano
	2. Mahembe	2. Karengera	2. Kanjongo
	3. Nyabitekera		3. Kirimbi
	Pot 2 Sectors		
	4. Bushenge	3. Rangiro	4. Bushekeri
	5. Cyato	4. Ruharambuga	5. Karambi
			6. Shangi
2. Rusizi	Pot 1 Sectors		
	1. Gashonga	1. Butare	1. Bweyeye
	2. Nyakabuye	2. Mururu	2. Gihundwe
			3. Nzahaha
	Pot 2 Sectors		
	3. Bugarama	3. Muganza	4. Gikundamvura
	4. Giheke	4. Nkombo	5. Kamembe
	5. Gitambi	5. Nyakarenzo	6. Nkanka
3. Rubavu	Pot 1 Sectors		
	1. Gisenyi	1. Cyanzarwe	1. Busasamana
			2. Kanama
	Pot 2 Sectors		
	2. Kanzenze	2. Bugeshi	3. Mudende
	3. Nyakiriba	3. Nyamyumba	4. Nyundo
	4. Rugerero		5. Rubavu
4. Ngorprero	Pot 1 Sectors		
	1. Kabaya	1. Nyange	1. Kavumu
	2. Sovu		2. Ngororero
	Pot 2 Sectors		
	3. Bwira	2. Hindiro	3. Kageyo
	4. Gatumba	3. Ndaro	4. Matyazo
	5. Muhanda		5. Muhororo
5. Rutsiro	Pot 1 Sectors		
	1. Mukura		1. Mushonyi
	2. Ruhango		2. Nyabirasi
	Pot 2 Sectors		
	3. Gihango	1. Kigeyo	
	4. Murunda	2. Manihira	3. Boneza
	5. Rusebeya	3. Mushubati	4. Kivumu

			5. Musasa
6. Karongi	Pot 1 Sectors		
	1. Murambi	1. Bwishyura	1. Mutuntu
	2. Rugabano	2. Gitesi	2. Rubengera
	3. Twumba	3. Ruganda	3. Rwankuba
	Pot 2 Sectors		
	4. Mubuga	4. Gashari	4. Gishyita
			5. Murundi
7. Nyabihu	Pot 1 Sectors		
	1. Rugera	1. Shyira	1. Bigogwe
	2. Rurembo		2. Rambura
	Pot 2 Sectors		
	3. Karago	2. Kabatwa	3. Jenda
	4. Kintobo	3. Mukamira	4. Jomba
			5. Mulinga

Kigali City			
District	Phase 1	Phase 2	Phase 3
1. Nyarugenge	Pot 1 Sectors		
	1. Nyakabanda	1. Kimisagara	1. Kigali
	2. Nyamirambo	2. Mageragere	2. Nyarugenge
	Pot 2 Sectors		
	3. Kanyinya	3. Muhima	3. Gitega
			4. Rwezamenyo
2. Kicukiro	Pot 1 Sectors		
	1. Nyarugunga	1. Kigarama	1. Gatenga
		2. Masaka	2. Kanombe
	Pot 2 Sectors		
	2. Gahanga	3. Niboye	3. Gikondo
	3. Kagarama		4. Kicukiro
3. Gasabo	Pot 1 Sectors		
	1. Bumbogo	1. Jali	1. Jabana
	2. Rusororo	2. Rutunga	2. Ndera
			3. Nduba
	Pot 2 Sectors		
	3. Gisozi	3. Kinyinya	4. Gatsata
	4. Kimihurura	4. Remera	5. Gikomero
	5. Kimironko		6. Kacyiru

Note: Pot 1 indicates that the number of schools in the sector is higher than the average number of schools in the district. Pot 2 indicates that the number of schools in the sector is lower than the average number of schools in the district.

ANNEX D. SELECTED DOCUMENTS

USAID Documents

Education Strategy, 2011- 2015

-- Implementation Guidance, revised 2012

-- Technical Notes

-- Reference Materials

-- 2015 Update to Reporting Guidance

Early Grade Reading Assessment (EGRA) Toolkit, Second Edition, March 2016

Evaluation Policy, updated October 2016

ADS Chapter 200. Development Policy

ADS Chapter 200. Mandatory Reference. Protection of Human Subjects in Research Supported by USAID

ADS Chapter 201. Program Cycle Operational Policy

ADS Chapter 201. Mandatory Reference. Criteria to Ensure the Quality of the Evaluation Report

ADS Chapter 201. Mandatory Reference. USAID Evaluation Report Requirements

Strengthening Evidence-Based Development: Five Years of Better Evaluation Practice at USAID, 2011–2016

Technical Note on Impact Evaluations, September 2013

JBS International - Aguirre Division. The Power of Coaching: Improving Early Grade Reading Instruction in Developing Countries. USAID, February 2014

Cost Reporting Guidance for USAID-Funded Education Projects, Pilot Version, February 2017

Other USG Documents

Centers for Disease Control and Prevention. National Center for Health Statistics. Washington Group Short Set of Questions on Disability, 2010

Government Accountability Office. Designing Evaluations, 2012 Revision

Government Accountability Office. USAID Has Implemented Primary Grade Reading Programs but Has Not Yet Measured Progress toward Its Strategic Goal (GAO-15-479), May 2015

Documents from Predecessor and Related Activities

Education Development Center, Literacy, Language and Learning Initiative (L3). Proposed National Reading Standards, Kinyarwanda and English, P3 & P5, 2012

Education Development Center, Literacy, Language and Learning Initiative (L3). National Fluency and Mathematics Assessment of Rwandan Schools: Midline Report, January 2016

Education Development Center, Literacy, Language and Learning Initiative (L3). National Fluency and Mathematics Assessment of Rwandan Schools: Endline Report, January 2017

RTI International. Early Grade Reading and Mathematics in Rwanda, Final Report. EdData II Task Number 7, February 2012

Save the Children. Early Literacy & Maths Initiative (ELMI) Rwanda: Endline Report, June 2015
Save the Children. Endline Evaluation of Rwandan Children's Book Initiative (RCBI), May 2015
Save the Children. Friedlander, E. & Goldenberg, C. (eds.). (2016). Literacy Boost in Rwanda: Impact Evaluation of a 2-year Randomized Control Trial. Stanford, CA: Stanford University
Save the Children. Friedlander, E. & Goldenberg, C. (eds.). (2016). Literacy Boost in Rwanda: Impact Evaluation of a 2-year Randomized Control Trial. Report Annex, Data Collection Stanford, CA: Stanford University
Save the Children. Literacy Boost in Rwanda: Impact Evaluation of a Two-Year Randomized Control Trial (Endline Summary 2016)
Save the Children. Early Literacy Promotion in Rwanda: Opportunities and Obstacles, April 2014.
Save the Children. Rwanda: Baseline Survey Tracking Literacy Knowledge, Attitudes and Practices at the School and Community Level, September 2016. Available at <https://dec.usaid.gov/dec/content/Detail.aspx?ctID=ODVhZjk4NWQtM2YyMi00YjRmLTkxNjktZTcxMjM2NDhmY2Uy&rID=MzIxNDEz>

Soma Umenye Documents

Soma Umenye Scope of Work (from IE RFTOP)
Soma Umenye Year I Workplan as of October 17
Soma Umenye Rwanda Education Sector Capacity Analysis Report

Other Governmental Documents

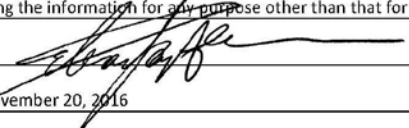
National Institute of Statistics of Rwanda (NISR), Demographic and Health Survey [DHS] – Rwanda, 2014/2015, Key Findings, June 2015
National Institute of Statistics of Rwanda (NISR), Statistical Yearbook, 2015 Edition, November 2015
Rwanda Education Board, Guide to Inclusive Education in Pre-primary, Primary, and Secondary Education, 2016
Rwanda Ministry of Education, 2014 Education Statistical Yearbook. March 2015
Rwanda Ministry of Education, 2015 Education Statistical Yearbook. June 2016
UNICEF, System Dynamics in Primary and Secondary Education in Rwanda – Inception Report, August 2016
World Food Program, Baseline Study: Home Grown School Feeding Program 2016-2020, July 22, 2016

Other Relevant Documents

Bloom, H.S. (2005): *Randomizing groups to evaluate place-based programs*. NY: Russell Sage Foundation, chap. Learning more from social experiments, pp. 115–172.
Bryk, A., Sebring, P.B., Kerbow, D., Rollow, S., Easton, J. (2000). *School Leadership and the Bottom Line*. Consortium on Chicago School Research. University of Chicago.
Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2008. “[Using Randomization in Development Economics Research: A Toolkit](#).” T. Schultz and John Strauss, eds., *Handbook of Development Economics*. Vol. 4. Amsterdam and New York: North Holland, 4.).
Fullam, M. (2000) *Principals as Leaders in a Culture of Change*. San Francisco. Jossey-Bass.
Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. 2016. *Impact Evaluation in Practice*, second edition. Washington, DC: Inter-American

- Development Bank and World Bank. doi:10.1596/978-1-4648-0779-4.
- IBTCl. Impact Evaluation of the USAID/ Aprender A Ler Project in Mozambique: Year 2 (Midline 2) IE/RCT, February 2015
- IBTCl. Impact Evaluation of the USAID/ Aprender A Ler Project in Mozambique: Year 3 IE/RCT, Final Report, May 2016
- Popova, Anna and David Evans. "Training teachers on the job: What we know, and why we know less than we should." World Bank Development Impact, September 28, 2016. <http://blogs.worldbank.org/impactevaluations/training-teachers-job-what-we-know-and-why-we-know-less-we-should>
- Schochet, Peter. "Statistical Power for Random Assignment Evaluations of Education Programs." *Journal of Educational and Behavioral Statistics*, vol. 33, no. 1, 2008, pp. 62–87.
- Schochet, Peter. "An Approach for Addressing the Multiple-Testing Problem in Social Policy Impact Evaluations." *Evaluation Review*, vol. 33, no. 6, 2009, pp. 539–567.
- Taylor, B., Pearson, P. & Pressley, C. (2002) Supporting Schools as they Implement Reading Reform. Research on Effective Schools.
- Taylor, B., Pearson, P., Peterson, D., & Rodriguez, M. (2003, September 1). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *Elementary School Journal*, 104.
- Vilenius-Tuohimaa, P, Aunolab, K., Nurmiba, J. The association between mathematical word problems and reading comprehension Department of Educational Sciences/Special Education, University of Jyväskylä, Finland; Department of Psychology, University of Jyväskylä, Finland and "Improving Problem Solving by Improving Reading Skills" (2009). Paper 9. <http://digitalcommons.unl.edu/> University of Nebraska, Lincoln.
- Visser, I; Kusters, C.S.L; Guijt, I; Roefs, M; Buizer, N. 2014. Improving the Use of Monitoring & Evaluation Processes and Findings; Conference Report. Centre for Development Innovation, Wageningen UR (University & Research centre). Report CDI-14-017. Wageningen.

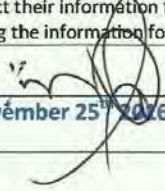
ANNEX E. CONFLICT OF INTEREST STATEMENTS

Name	Edward Jay Allan
Title	Project Director
Organization	IBTCI
Evaluation Position?	☐ Team Leader ☐ Team member ☒ Project Director
Evaluation Award Number (contract or other instrument)	AID-696-TO-16-00002
USAID Project(s) Evaluated (Include project name(s), implementer name(s) and award number(s), if applicable)	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes ☒ No X
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	
I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.	
Signature	
Date	November 20, 2016

Soma Umenye Impact Evaluation – Draft Year I Workplan and Evaluation Design

Name	Magdala Raupp
Title	
Organization	IBTCI
Evaluation Position?	☒ Team Leader ☐ Team member
Evaluation Award Number (contract or other instrument)	AID-696-TO-16-00002
USAID Project(s) Evaluated (include project name(s), implementer name(s) and award number(s), if applicable)	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes ☒ No
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	
I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.	
Signature	
Date	November 25, 2016

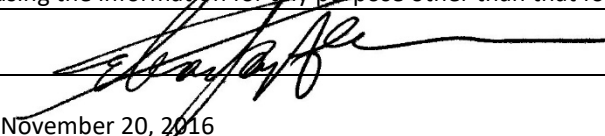
Name	ALI PROTIK
Title	SENIOR RESEARCHER
Organization	MATHEMATICA POLICY RESEARCH
Evaluation Position?	☐ Team Leader ☒ Team member
Evaluation Award Number (contract or other instrument)	AID-696-TO-16-00002
USAID Project(s) Evaluated (Include project name(s), implementer name(s) and award number(s), if applicable)	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes ☒ No
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	
I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.	
Signature	
Date	12/2/2016

Name	Kenn Ndirangu
Title	Field Manager
Organization	Incisive Africa
Evaluation Position?	☐ Team Leader ☒ Team member
Evaluation Award Number (contract or other instrument)	AID-696-TO-16-00002
USAID Project(s) Evaluated (Include project name(s), implementer name(s) and award number(s), if applicable)	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes ☒ No
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	
I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.	
Signature	
Date	November 25th 2016

ANNEX D. CONFLICT OF INTEREST CERTIFICATIONS

Name	Edward Jay Allan
Title	Project Director
Organization	IBTCI
Evaluation Position?	☐ Team Leader ☐ Team member ☒ Project Director
Evaluation Award Number (<i>contract or other instrument</i>)	AID-696-TO-16-00002
USAID Project(s) Evaluated (<i>Include project name(s), implementer name(s) and award number(s), if applicable</i>)	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes ☐ No ☒ X
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	

I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.

Signature	
Date	November 20, 2016

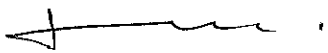
Name	Magdala Raupp
Title	
Organization	IBTCI
Evaluation Position?	☐ X Team Leader ☐ Team member
Evaluation Award Number <i>(contract or other instrument)</i>	AID-696-TO-16-00002
USAID Project(s) Evaluated <i>(Include project name(s), implementer name(s) and award number(s), if applicable)</i>	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes ☒ No
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	

I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.

Signature	
Date	November 25, 2016

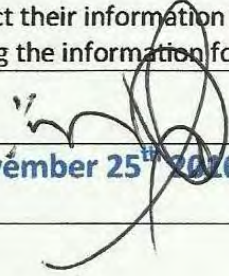
Name	ALI PROTIK
Title	SENIOR RESEARCHER
Organization	MATHEMATICA POLICY RESEARCH
Evaluation Position?	☐ Team Leader ☒ Team member
Evaluation Award Number (contract or other instrument)	AID-696-TO-16-00002
USAID Project(s) Evaluated (Include project name(s), implementer name(s) and award number(s), if applicable)	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes ☒ No
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	

I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.

Signature	
Date	12/2/2016

Name	Kenn Ndirangu
Title	Field Manager
Organization	Incisive Africa
Evaluation Position?	☐ Team Leader ☒ Team member
Evaluation Award Number (contract or other instrument)	AID-696-TO-16-00002
USAID Project(s) Evaluated (Include project name(s), implementer name(s) and award number(s), if applicable)	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes ☒ No
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	

I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.

Signature	
Date	November 25th 2026

Name	Nikita Ramchandani
Title	International Research Analyst
Organization	Mathematica Policy Research
Evaluation Position?	☐ Team Leader X Team member
Evaluation Award Number (contract or other instrument)	AID-696-TO-16-00002
USAID Project(s) Evaluated (Include project name(s), implementer name(s) and award number(s), if applicable)	USAID/Rwanda Soma Umenye
I have real or potential conflicts of interest to disclose.	☐ Yes X No
If yes answered above, I disclose the following facts: <i>Real or potential conflicts of interest may include, but are not limited to:</i> 1. Close family member who is an employee of the USAID operating unit managing the project(s) being evaluated or the implementing organization(s) whose project(s) are being evaluated. 2. Financial interest that is direct, or is significant though indirect, in the implementing organization(s) whose projects are being evaluated or in the outcome of the evaluation. 3. Current or previous direct or significant though indirect experience with the project(s) being evaluated, including involvement in the project design or previous iterations of the project. 4. Current or previous work experience or seeking employment with the USAID operating unit managing the evaluation or the implementing organization(s) whose project(s) are being evaluated. 5. Current or previous work experience with an organization that may be seen as an industry competitor with the implementing organization(s) whose project(s) are being evaluated. 6. Preconceived ideas toward individuals, groups, organizations, or objectives of the particular projects and organizations being evaluated that could bias the evaluation.	

I certify (1) that I have completed this disclosure form fully and to the best of my ability and (2) that I will update this disclosure form promptly if relevant circumstances change. If I gain access to proprietary information of other companies, then I agree to protect their information from unauthorized use or disclosure for as long as it remains proprietary and refrain from using the information for any purpose other than that for which it was furnished.

Signature	<i>N. Ramchandani</i>
Date	10/22/2017

ANNEX E. DISTRICTS AND SECTORS BY PHASE

ANNEX C.**LIST OF SECTORS IN PHASES 1, 2, AND 3**

Northern Province			
District	Phase 1	Phase 2	Phase 3
1. Rulindo	Pot 1 Sectors		
	1. Buyoga	1. Bushoki	1. Cyinzuzi
	2. Rukozo	2. Kisaro	2. Mbogo
	3. Shyorongi	3. Tumba	3. Murambi
	Pot 2 Sectors		
	4. Base	4. Cyungo	4. Burega
	5. Kinihira	5. Ngoma	5. Masoro
	6. Ntarabana		6. Rusiga
2. Gakenke	Pot 1 Sectors		
	1. Gakenke	1. Busengo	1. Muzo
	2. Ruli		2. Rushashi
	Pot 2 Sectors		
	3. Coko	2. Gashenyi	3. Karambo
	4. Cyabingo	3. Kamubuga	4. Mataba
	5. Janja	4. Mugunga	5. Minazi
	6. Kivuruga	5. Rusasa	6. Muhondo
			7. Muyongwe
			8. Nemba
3. Burera	Pot 1 Sectors		
	1. Kinyababa	1. Kivuye	1. Butaro
	2. Nemba	2. Rugarama	2. Cyanika
	Pot 2 Sectors		
	3. Bungwe	3. Gahunga	3. Cyeru
	4. Ruhunde	4. Gatebe	4. Gitovu
	5. Rwerere	5. Rugengabari	5. Kagogo
			6. Kinoni
			7. Rusarabuye
4. Musanze	Pot 1 Sectors		
	1. Cyuve	1. Kinigi	1. Muhoza
		2. Musanze	2. Nkotsi
	Pot 2 Sectors		
	2. Gataraga	3. Busogo	3. Gacaca
	3. Nyange	4. Gashaki	4. Kimonyi
	4. Remera		5. Muko
	5. Shingiro		6. Rwaza

Soma Umenye Impact Evaluation – Draft Year I Workplan and Evaluation Design

5. Gicumbi	Pot 1 Sectors		
	1. Bukure	1. Nyankenke	1. Bwisige
	2. Cyumba	2. Rushaki	2. Byumba
	3. Mutete	3. Rutare	3. Kaniga
	4. Rukomo	4. Ruvune	4. Miyove
			5. Muko
	Pot 2 Sectors		
	5. Giti	5. Kageyo	6. Manyagiro
	6. Nyamiyaga	6. Mukarange	7. Rubaya
	7. Rwamiko		8. Shangasha

Eastern Province			
District	Phase 1	Phase 2	Phase 3
1. Bugesera	Pot 1 Sectors		
	1. Nyamata	1. Ngeruka	1. Juru
	2. Rweru		2. Mayange
	Pot 2 Sectors		
	3. Mwogo	2. Gashora	3. Kamabuye
	4. Ntarama	3. Musenyi	4. Mareba
	5. Rilima	4. Ruhuha	5. Nyarugenge
			6. Shyara
2. Gatsibo	Pot 1 Sectors		
	1. Gatsibo	1. Gitoki	1. Muhura
	2. Kageyo	2. Kabarore	2. Nyagihanga
	3. Kiziguro	3. Kiramuruzi	3. Rugarama
			4. Rwimbogo
	Pot 2 Sectors		
	4. Remera	4. Murambi	5. Gasange
			6. Ngarama
3. Kayonza	Pot 1 Sectors		
	1. Gahini	1. Rukara	1. Murundi
	2. Kabare		2. Mwiri
	Pot 2 Sectors		
	3. Murama	2. Ndego	3. Kabarondo
	4. Rwinkwavu	3. Nyamirama	4. Mukarange
			5. Ruramira
4. Kirehe	Pot 1 Sectors		
	1. Musaza	1. Gahara	1. Kirehe
	2. Nasho	2. Mushikiri	2. Mpanga
	Pot 2 Sectors		
	3. Mahama	3. Kigarama	3. Gatore
	4. Nyamugali		4. Kigina
			5. Nyarubuye
5. Ngoma	Pot 1 Sectors		
	1. Kibungo	1. Murama	1. Kazo
	2. Mugesera		2. Rukira
	Pot 2 Sectors		
	3. Mutenderi	2. Jarama	3. Gashanda
	4. Sake	3. Karemba	4. Rukumberi
		4. Remera	5. Rurenge
			6. Zaza

6. Nyagatare	Pot 1 Sectors		
	1. Karangazi	1. Rwempasha	1. Mukama
	2. Tabagwe	2. Rwimiyaga	2. Nyagatare
	Pot 2 Sectors		
	3. Gatunda	3. Karama	3. Kiyombe
	4. Katabagemu	4. Matimba	4. Mimuli
			5. Musheru
			6. Rukomo
7. Rwamagana	Pot 1 Sectors		
	1. Muhazi	1. Nyakariro	1. Kigabiro
	2. Mwulire		2. Munyaga
	Pot 2 Sectors		
	3. Muyumbu	2. Gahengeru	3. Fumbwe
	4. Nzige	3. Karenga	4. Gishali
		4. Musha	5. Munyiginya
			6. Rubona

Southern Province			
District	Phase I	Phase 2	Phase 3
1. Gisagara	Pot 1 Sectors		
	1. Muganza	1. Mamba	1. Gikonko
	2. Mugombwa	2. Ndora	2. Gishubi
	3. Save		3. Kansi
	Pot 2 Sectors		
	4. Kigembe	3. Mukindo	4. Kibirizi
	5. Nyanza		5. Musha
2. Huye	Pot 1 Sectors		
	1. Ruhashya	1. Kigoma	1. Maraba
	2. Rusatira		2. Mbazi
	Pot 2 Sectors		
	3. Karama	2. Ngoma	3. Gishamvu
	4. Kinazi	3. Rwaniro	4. Huye
	5. Mukura		5. Simbi
3. Muhanga	Pot 1 Sectors		
	1. Kibangu	1. Nyamabuye	1. Nyarusange
			2. Rongi
	Pot 2 Sectors		
	2. Cyeza	2. Kiyumba	3. Mushishiro
	3. Kabacuzi	3. Shyogwe	4. Nyabinoni
	4. Muhanga		5. Rugendabari
4. Kamonyi	Pot 1 Sectors		
	1. Kayenzi	1. Gacurabwenge	1. Nyarubaka
	2. Mugina		2. Rukoma
	3. Ngamba		
	Pot 2 Sectors		
	4. Rugarika	2. Kayumbu	3. Karama
		3. Runda	4. Musambira
5. Ruhango	Pot 1 Sectors		
	1. Byimana	1. Ruhango	1. Bweramana
	2. Kinazi		2. Ntongwe
	Pot 2 Sectors		
	3. Mbuye	2. Kinihira	3. Kabagali
6. Nyanza	Pot 1 Sectors		
	1. Busasamana	1. Mukingo	1. Busoro

	2. Cyabakamyi	2. Nyagisozi	2. Muyira
	Pot 2 Sectors		
	3. Rwabicuma	3. Ntyazo	3. Kibirizi
			4. Kigoma
7. Nyaruguru	Pot 1 Sectors		
	1. Busanze	1. Ruramba	1. Ngera
	2. Kibeho		2. Ngoma
	Pot 2 Sectors		
	3. Cyahinda	2. Nyabimata	3. Muganza
	4. Kivu	3. Ruheru	4. Munini
	5. Mata		5. Nyagisozi
			6. Rusenge
8. Nyamagabe	Pot 1 Sectors		
	1. Mugano	1. Tare	1. Gasaka
	2. Musange	2. Uwinkingi	2. Kaduha
			3. Kitabi
	Pot 2 Sectors		
	3. Buruhukiro	3. Cyanika	4. Gatare
	4. Kamegeri	4. Kibirizi	5. Kibumbwe
	5. Musebeya	5. Mbazi	6. Mushubi
			7. Nkomane

Western Province			
District	Phase I	Phase 2	Phase 3
1. Nyamasheke	Pot 1 Sectors		
	1. Macuba	1. Gihombo	1. Kagano
	2. Mahembe	2. Karengera	2. Kanjongo
	3. Nyabitekera		3. Kirimbi
	Pot 2 Sectors		
	4. Bushenge	3. Rangiro	4. Bushekeri
	5. Cyato	4. Ruharambuga	5. Karambi
			6. Shangi
2. Rusizi	Pot 1 Sectors		
	1. Gashonga	1. Butare	1. Bweyeye
	2. Nyakabuye	2. Mururu	2. Gihundwe
			3. Nzahaha
	Pot 2 Sectors		
	3. Bugarama	3. Muganza	4. Gikundamvura
	4. Giheke	4. Nkombo	5. Kamembe
	5. Gitambi	5. Nyakarenzo	6. Nkanka
3. Rubavu	Pot 1 Sectors		
	1. Gisenyi	1. Cyanzarwe	1. Busasamana
			2. Kanama
	Pot 2 Sectors		
	2. Kanzenze	2. Bugeshi	3. Mudende
	3. Nyakiriba	3. Nyamyumba	4. Nyundo
	4. Rugerero		5. Rubavu
4. Ngorprero	Pot 1 Sectors		
	1. Kabaya	1. Nyange	1. Kavumu
	2. Sovu		2. Ngororero
	Pot 2 Sectors		
	3. Bwira	2. Hindiro	3. Kageyo
	4. Gatumba	3. Ndaro	4. Matyazo
	5. Muhanda		5. Muhororo
5. Rutsiro	Pot 1 Sectors		
	1. Mukura		1. Mushonyi
	2. Ruhango		2. Nyabirasi
	Pot 2 Sectors		
	3. Gihango	1. Kigeyo	
	4. Murunda	2. Manihira	3. Boneza
	5. Rusebeya	3. Mushubati	4. Kivumu

			5. Musasa
6. Karongi	Pot 1 Sectors		
	1. Murambi	1. Bwishyura	1. Mutuntu
	2. Rugabano	2. Gitesi	2. Rubengera
	3. Twumba	3. Ruganda	3. Rwankuba
	Pot 2 Sectors		
	4. Mubuga	4. Gashari	4. Gishyita
			5. Murundi
7. Nyabihu	Pot 1 Sectors		
	1. Rugera	1. Shyira	1. Bigogwe
	2. Rurembo		2. Rambura
	Pot 2 Sectors		
	3. Karago	2. Kabatwa	3. Jenda
	4. Kintobo	3. Mukamira	4. Jomba
			5. Mulinga

Kigali City			
District	Phase 1	Phase 2	Phase 3
1. Nyarugenge	Pot 1 Sectors		
	1. Nyakabanda	1. Kimisagara	1. Kigali
	2. Nyamirambo	2. Mageragere	2. Nyarugenge
	Pot 2 Sectors		
	3. Kanyinya	3. Muhima	3. Gitega
			4. Rwezamenyo
2. Kicukiro	Pot 1 Sectors		
	1. Nyarugunga	1. Kigarama	1. Gatenga
		2. Masaka	2. Kanombe
	Pot 2 Sectors		
	2. Gahanga	3. Niboye	3. Gikondo
	3. Kagarama		4. Kicukiro
3. Gasabo	Pot 1 Sectors		
	1. Bumbogo	1. Jali	1. Jabana
	2. Rusororo	2. Rutunga	2. Ndera
			3. Nduba
	Pot 2 Sectors		
	3. Gisozi	3. Kinyinya	4. Gatsata
	4. Kimihurura	4. Remera	5. Gikomero
	5. Kimironko		6. Kacyiru

Note: Pot 1 indicates that the number of schools in the sector is higher than the average number of schools in the district. Pot 2 indicates that the number of schools in the sector is lower than the average number of schools in the district.

ANNEX F. SUPERVISOR MANUAL



USAID
FROM THE AMERICAN PEOPLE

USAID RWANDA SOMA UMENYE

SEPTEMBER 2017

Team Supervisor Manual

A guide to ensure the quality of the data collection process

This guide is made possible by the support of the American People through the United States Agency for International Development (USAID.) The contents of this guide are the sole responsibility of the Soma Umenye partners and do not necessarily reflect the views of USAID or the United States Government.

Table of Contents

I.	Overview.....	3
2.	Overall Responsibilities of EGRA Team Leaders.....	3
2.1	Pre-Visit Checklist.....	5
2.2	Key Contact Numbers.....	6
2.3	Briefing School Staff.....	9
2.4	Location to Administer EGRA.....	10
2.5	Random Selection of the Class for EGRA Assessment.....	11
2.6	Random Selection of Students for Assessment.....	12
2.7	During the EGRA Assessment.....	15
2.8	Assessor Workload.....	15
2.9	Other Instruments Being Used for Data Collection.....	16
2.10	Typical Problem Scenarios during Data Collection.....	12
2.11	End of Visit Checklist.....	20
	Annex A.....	15
	Annex B.....	16
	Annex C.....	17
	Annex D.....	19
	Annex E.....	21
	Annex F.....	23

I. OVERVIEW

This information booklet is based on the guidance provided by USAID's Early Grade Reading Assessment (EGRA) Toolkit, Second Edition (2016) and builds on the experience developed by Chemonics, the Soma Umenye Implementer, EdIntersect, Chemonics' subcontractor in charge of instrument development and training for data collection, IBTCI, the implementer of the Impact Evaluation of Soma Umenye in Rwanda and in other countries, and Africa Incisive, IBTCI's subcontractor managing logistics and field operations. The overall purpose and background for the data collection are described below.

The Rwanda Education Board (REB) and the US Agency for International Development (USAID) are working together to develop ways to improve the reading skills of Rwanda's children, especially those in grades P1-P3. To do this, it is important to know the average level of reading at P1. The information will be of great interest for Soma Umenye in their efforts to implement activities that will improve the reading ability of children in the early grades.

To collect this information, in September/October 2017 we are visiting a randomly selected sample of 329 schools throughout Rwanda. At each school, one P1 class will be randomly selected for participation. From the class selected we will randomly select nine boys and nine girls for the assessment. The results of these assessments will be combined with the results obtained by students in other classes and other schools throughout the district and the country. Nobody will know the scores obtained by individual students and the scores will have no impact on a child's grade at the end of the year or at any time. Participation is completely voluntary. If a child does not want to take part in the EGRA, he or she does not have to. We will ask another child instead. There will be no penalty or reward for lower or higher scores. We will also administer a questionnaire, the Student Context questionnaire, to the 18 participating students. The purpose of the assessment is to find out how well, on average, P1 children are reading.

In addition to conducting the EGRA, we will be collecting school-related data. In each of the 329 schools we will interview one P1 Kinyarwanda teacher, observe one P1 Kinyarwanda class, and interview the head teacher. As with children, the participation of teachers and head teachers is voluntary. Prior to starting any assessments or interviews you will ask the interviewee if she/he wants to participate. You can learn more about the EGRA Toolkit at <https://globalreadingnetwork.net/resources/early-grade-reading-assessment-egra-toolkit-second-edition>. If you have any questions or issues prior to or during the data collection, please call:

Organization	Contact	Mobile phone/email
Chemonics/Soma Umenye	Laura Harrington, MEL Director	+ 250 785 471 885 lharrington@soma-umenye.org
Incisive Africa	Kenn Ndirangu, Field Manager	+250 788305422 kenn@incisiveafrica.com

2. OVERALL RESPONSIBILITIES OF TEAM SUPERVISORS

The main responsibility of the team supervisors is to maintain the quality of the data collected ensuring that correct data collection procedures are followed. Specific responsibilities include:

1. Prepare all the equipment and resources needed for the school visit (see pre-visit checklist on Section 2.1, page 3).
2. Notify the Head Teacher and District Advisors responsible for the catchment area a few days in advance of the visits. DO NOT share the exact day you will be visiting schools.
3. Introduce the team to the head teacher and to the PI teachers and explain the purpose of the school visit (Please see a suggested script for the interchange with the head teacher on page 4).
4. Conduct the random selection of the class that will participate in the assessment. (Please see instructions in Section 2.5, page 6).
5. Conduct the random selection of the students to participate in the assessment. (Please see instructions for selecting students in Section 2.6, page 6).
6. Ensure that a witness follows the random selection process and signs a form stating that the team supervisor conducted the selection of classes and students randomly. (Please see the form included in Annex A).
7. Supervise the work of enumerators/assessors during the assessment, making sure that they administer the EGRA and the Student Context Questionnaire in an efficient, friendly, and technically correct manner. Observe one full assessment per day and provide feedback to the enumerator as to his/her performance. When observing the assessment, do not sit next to the enumerator, but instead stand nearby, observe, and listen. The presence of two adults may intimidate the child. (Please see Annex B for a sample school schedule).
8. Provide any necessary feedback to assessors to help them improve the quality of their work.
9. Help the assessors when you have finished your tasks and the assessors are still at work conducting assessments.
10. Ensure that all equipment and documentation is collected at the end of the visit (see End of Visit checklist on page 12).
11. Ensure all data is uploaded daily and work with the field supervisor and enumerators to clarify any inconsistencies. If there is no network access, notify your field supervisor.

12. Provide feedback to the field supervisors on challenges encountered and, when appropriate, give suggestions for future improvements of the data collection process.

Making Contact with Schools

Before each school visit, supervisors should do the following:

- Contact the District Advisor a few days before you arrive in a catchment area. Let them know what day you'll be arriving and for how many days you'll be in the district. They have a list of the schools you'll be visiting.
- Contact the Head Teacher as far in advance as possible, prior to visiting the school, to confirm you will be visiting at some point in the near future.
- DO NOT inform ANYONE at the school of the exact day you plan to visit – nobody should know which day the team will arrive as this could introduce bias.
- Ensure the Head Teacher understands the purpose and activities of the visit.
- Verify the materials needed for each school.
- Verify that tablets are charged and you have enough paper surveys in case a tablet crashes.
- If there are issues please notify your field supervisor immediately and work together with the central office to find a solution.
- Complete school information (e.g., name, region, date, etc.).

2.1 Pre-Visit Checklist

Team supervisors are responsible for ensuring that the team has all they need to conduct the assessment in the manner specified during training. Use this checklist when preparing for the school visit. We recommend that you do this the day before the school visit.

Item	Description	✓
1	Copy of notification letter from the Ministry of Education	
2	Your tablet and the tablets of the team fully charged, loaded with instruments	
3	A charger for your tablets	
4	Power pack for charging your tablet in the field	
5	2 printed copies of classroom observations forms. You will use the paper copy to conduct your classroom observation and then enter this information into the tablet after completing your observation. In addition, make sure to take 2 hard copies of EGRA and student questionnaires; 2 teacher questionnaires; and 2 head teacher questionnaires in case your tablet fails and you have to revert to paper.	
6	One master copy of each questionnaire for additional printing if necessary (in case technology fails)	
7	Clipboards, one per assessor (only for paper administration)	
8	Pencils, pens, sharpeners (a few for paper administration)	
9	Respondent incentives (one pencil and one eraser per student assessed)	
10	Two sets of number slips (one color for girls, one color for boys) for selecting boys and girls to be assessed and interviewed	
11	Extension cord/multi-prise with adapter	

I2	Modem & adapter for uploading data	
I3	1 or 2 spare tablets	
I4	Route Plan	

2.2 Key Contact Numbers

Field emergencies do happen and you may need to communicate with specific people to advise them on a problem. Examples may include, if your vehicle breaks down and you will be late arriving at the school, or you arrive at the school and the head teacher tells you that he/she has not been informed of your visit and cannot allow you to conduct the assessments. Or, for some reason the school is closed and you want to know whether there is another school in the area. For each school visit we will provide you with a sheet of contacts with names of the DEO, SEO, head teacher, etc.

Name	Position	Phone Number
Placide Simbizi	Monitoring, Evaluation, and Learning (MEL) Specialist, Soma	+250 788 737 749
Kenn Ndirangu	Field Manager, Incisive Africa	+250 788 305 422
Laura Harrington	Monitoring, Evaluation, and Learning (MEL) Director, Soma Umenye	+250 785 471 885

2.2 Briefing School Staff

When you and your team arrive at the school go straight to the head teacher's office. If the head teacher is not at the school, take note of the time and check the box in your report that is next to:

Head teacher at school when team arrived YES ☐ NO ☐

Afterwards, look for the Dean of Studies or Le Responsable or another administrative staff that may be replacing him/her.

Your first task is to brief the head teacher about the purpose of your visit. Head teachers should already have been informed that an EGRA team is visiting their school. This introductory briefing should not take long. We prepared the script below to assist you and save time. However, this is not for you to read directly to the head teacher. Please learn it by memory. Or, if you prefer, use your own words. This introduction should take no longer than ten minutes and should include the points listed below.

My name is ____ and I am the leader of the EGRA team assigned to conduct the reading assessment at your school. REB and the US Agency for International Development (USAID) are working together to develop ways to improve the reading skills of Rwanda's children. To do this, they have designed the Soma Umenye program that will be implemented all over Rwanda specifically to improve the reading skills of students in P1, P2, and P3. At this point it is important to know, on average, what the current level of reading at P1 is. This is our mission. We thank you in advance for your assistance.

This is what we expect you to do:

1. Introduce yourself and your team as the EGRA assessors.
2. Explain the purpose of the visit.
3. Outline the planned activities for the day: the team will administer the EGRA, administer the Student Context Questionnaire; interview one Kinyarwanda PI teacher; conduct one lesson observation (of a Kinyarwanda teacher teaching PI); and interview the head teacher.
4. Explain that the class to be assessed will be selected randomly by you, as supervisors of the EGRA team, in the presence of the head teacher and of a witness. (Witness could be the Dean of Studies/le Responsable or a PI teacher whose class is not selected.)
5. Explain that after the class is selected, students in that class will be selected randomly by you, (**not** by the teacher). Explain that 18 students will be selected: 9 boys and 9 girls.
6. Emphasize that all information you and your team collect is confidential and school, student and teacher names will not be used on any reports. The

information gathered is for a research study. It is not related to an inspection or performance assessments of students or teachers.

7. Ask the Head Teacher to help you identify a suitable location for EGRA administration. Respond to any questions from teachers or from the observer. For example, if the head teachers ask whether their schools will receive the program, you can say that all schools in Rwanda will participate in the program either in 2017, 2018, or 2019.
8. Make a list of all the PI classes that are functioning at the moment of the visit in any order, not listed alphabetically or following a specific order. For example, you DO NOT want a list such as G1, G2, G3, G4 or A, B, C, D, etc. Your list could be G2, G4, G1, G3. Please keep the list you used to randomly select the class to participate in the EGRA assessment because you will turn this in to your field supervisor at the end of data collection.

Once the classroom has been selected, ask the head teacher when the Kinyarwanda lesson is taught in that particular classroom. Depending on when the Kinyarwanda lesson is taught, you as team leader will need to arrange the order of the Head Teacher Interview, and Teacher Interview.

9. Thank the head teacher on behalf of the team and point out that we could not do the work without his or her assistance.

Now that you are on good terms with the head teacher and he/she knows what you and your team will be doing at the school you can ask to see the place for EGRA administration, check that the location is adequate for the purpose of the assignment.

2.2 Location to Administer the EGRA

Ideally, you would communicate with the head teacher prior to your team's arrival at the school and ask for his/her assistance to identify a suitable place for the assessment. Usually, this will be a spare classroom or library, as shown in Figure 1 below. As the supervisor, you will need to inspect the designated area for the administration of EGRA.



Features of a good EGRA assessment location:

- Enough light for children to read with desks spaced out;
- Away from distractions (outside noise, visual distractions, etc);
- Desks completely clear of papers and other distractions; and
- Mobile phones on silent.

Figure 1: EGRA testing in a spare classroom

2.5 Random Selection of the Class for EGRA Assessment

It is essential that in each school, one PI class be selected at random. To ensure the selection is random, first make a list of all PI classes in any order. You can copy it from a list that the head teacher already has, you can write it down as the head teacher dictates the names, or you can write only the numbers of the classes. For example, if there are five PI classes, your list of classes could be 5, 3, 1, 2, and 4. When you have the list of all PI classes in the school, you are ready to start the selection. Remember that you do not want G1, G2 or Class 1, Class 2, Class 3, etc. You want this list mixed (for example, G4, G2, G3, G6, etc.). The random selection of the class to be assessed should be conducted in the presence of the head teacher. If there is a community representative at the school to serve as a witness, that person should observe the class selection. If there is no representative from the community ask another member of the school staff to act as a witness—head master, deputy head master, assistant, another teacher, or any other staff member. The process of selecting the class is described below.

If there is only **one** PI class, that one class will be assessed.

If **two** classes:

- o Day of the month is 01-15, select the first from the list.
- o Day of the month is 16-31, select the second.

If **three** classes:

- o Day of the month is 01-10, select the first from the list.
- o Day of the month is 11-20, select the second.
- o Day of the month is 21-31, select the third.

If **four** or more classes:

- o Day of the month is 01-07, select the first from the list.
- o Day of the month is 08-14, select the second.
- o Day of the month is 15-21, select the third.
- o Day of the month is 22-31, select the fourth.

This procedure will effectively eliminate any bias in selection of the class to be assessed that is based on a head teacher's or teacher's knowledge of classroom performance or teacher characteristics.

This is the class from which you will randomly select 9 boys and 9 girls for the assessment. Once the class has been selected ask the head teacher: *"At what time will this PI class be receiving the Kinyarwanda lesson?"* Note that you may have to wait for the Kinyarwanda class to start. For example, if you arrive at the school at 8:00am, and randomly select the PI class but the Kinyarwanda class only starts after the break. Select the students for assessment, start the EGRA assessment, and after the break go to the class to observe the Kinyarwanda class and interview the teacher. While the EGRA assessment is being conducted, interview the head teacher and observe or help the assessors.

When it is time for the PI Kinyarwanda class to start, ask the head teacher to introduce you to the teacher. Explain your mission in a few words. Thank the teacher for his/her contribution. Ask the teacher to see the attendance roster for the class. Take a photo of the complete roster with your tablet. If the class has no attendance roster, make a note of it. Write on your Supervisor report *Teacher of selected class could not show an attendance roster*. Check how many students are registered in the class (those who are included in the roster). Write on your report that you will know the number of students present by checking how many number cards you distributed when you randomly selected students for assessment.

2.6 Random Selection of Students for Assessment

Not every child in the class will be assessed. Teachers do not select students—YOU select students. Explain to the teacher the procedure you will use to randomly select children for assessment. If a witness is present to observe the selection process, he/she will follow the process and sign a form confirming that the process was followed (Please see annex A). Before you start the selection you must explain to students the objective of the assessment. See the brief script we prepared to assist you or use your own words.

Good morning. My name is ____ and these are my colleagues ____, ____, and _____. I am here today to ask for your help. We are going to many schools and asking children like you to answer some questions because we are working with the government of Rwanda in its efforts to make children read better. This is NOT a test but we will ask you to show some things that you have learned. I am going to give everyone in this class a number and then I will select some numbers (like a lottery) to participate. You do not HAVE to participate. If you do not want to participate say NO when I call out your number and I will ask another student.

When selecting the children for assessment our objective is to eliminate teacher or enumerator bias. You will select 18 students (nine boys and nine girls). The steps for randomly selecting students are detailed below.

1. Separate the class in two groups: boys on one side and girls on the other side. Please note that in the pilot, teams found it easier to first sample girls, then sample boys (or vice versa) so that each group stands and is sampled, then move on to the next.
2. If there are 9 or fewer boys, there is no need to distribute cards with numbers since all of them will be assessed. If there are 9 or fewer girls, there is no need to distribute cards with numbers since all of them will be assessed. If fewer than 9 boys or 9 girls **DO NOT** take students from another PI class to complete 9.

If 10 or more boys or 10 or more girls follow the instruction below:

3. Give each student in the group a card with numbers from 01 to the total number of students in that group. For example, if there are 11 boys they will receive numbers/cards 1 – 11 (the order in which you give the numbers does not matter). If there are 15 girls they will receive numbers/cards 1 – 15 (again, in any order). Make a note of this.
4. In the left-hand column of the table below you have the number of students in the group. For example, if there are 13 boys in the group, look for the number 13 in the left-hand column. Look at number 13 in the left-hand column. Select for assessment the 9 boys that

hold the numbers that appear in the right-hand column (**07 05 01 03 04 08 09 02 06**).

5. If there are 16 girls in the group, look for the number 16 in the left-hand column. Select for assessment the 9 girls that hold the numbers shown in the right-hand column (**15 09 16 06 05 03 08 10 13**). To avoid taking 18 students out of class all at once and having them wait to be assessed, we prefer that you call the numbers of three students, take them to EGRA assessment and then, return to the classroom to call another three students.

NUMBER OF GIRLS/BOYS PRESENT IN CLASS	In each of the two groups select the boys or the girls that hold these numbers
10	07 05 01 03 04 08 09 02 06
11	07 11 06 10 09 03 08 02 01
12	01 04 05 08 06 11 03 07 12
13	06 13 01 04 03 12 09 05 08
14	12 13 01 07 02 10 09 11 04
15	01 03 06 07 08 04 02 09 10
16	15 09 16 06 05 03 08 10 13
17	12 09 15 14 17 11 07 03 02
18	06 04 10 09 17 07 15 12 11
19	09 03 15 16 05 17 01 14 19
20	02 01 03 17 15 09 14 20 19
21	05 04 11 12 13 10 20 02 17
22	11 02 17 16 19 15 22 14 09
23	05 15 02 12 13 19 20 03 14
24	03 08 15 06 14 10 22 16 12
25	17 09 25 02 12 07 24 06 18
26	10 18 20 21 17 03 11 07 25
27	08 14 11 20 19 03 15 27 23
28	21 28 14 12 27 17 26 19 07
29	23 12 09 08 11 03 28 20 06
30	30 22 24 09 02 27 08 06 19
31	01 25 07 26 04 10 31 30 23
32	24 23 26 17 13 04 12 11 21
33	12 33 04 16 21 14 02 31 32
34	13 18 11 32 04 03 07 22 26
35	25 09 12 21 19 08 17 33 23
36	17 11 29 03 34 23 10 12 26
37	26 03 25 05 18 16 12 23 02
38	32 17 34 10 28 18 13 27 15
39	17 08 24 10 20 28 15 05 34
40	26 31 15 35 24 36 39 01 14
41	33 12 35 18 06 02 34 09 36
42	33 15 28 01 17 32 19 03 10
43	13 42 31 20 18 28 40 25 39
44	24 36 25 13 11 10 34 18 33
45	16 42 10 27 44 11 02 29 37
46	10 05 21 18 38 12 33 26 23
47	06 12 08 19 34 35 07 09 17
48	14 05 34 42 43 41 46 32 25
49	37 43 39 01 17 18 38 40 49
50	31 42 36 04 37 09 44 10 06

This table of 9 random numbers was produced according to the following specifications: Numbers were randomly selected from within the range of 1 to 50. Duplicate numbers were not allowed. This table was generated on 8/01/2017.

If a student who has his/her number selected says, “*NO, I don’t want to participate*” or indicates that he/she does not want to participate, simply say “*fine, no problem.*” Thank the student, be friendly, do not show displeasure, and do not communicate to the student in words or body

language that you disapprove of his/her refusal to participate. Then choose another student that has the number before or the number after. For example, look for 20 in the left-hand column and see which numbers correspond to 20: **02 01 03 17 15 09 14 20 19**. If the student who was holding card # 17 declines to participate, immediately call the student holding card # 18 first and if that student does not wish to participate, call the student holding card number #16. Now, let's say that the student holding card # 20 says he or she does not want to participate. Note that student #19 is already part of the sample, so go to student the student holding card # 18.

Once the children have been selected, please thank the teacher and witness/observer, ask her/him to sign the statement that s/he has observed the selection process and believes that it was conducted properly, and s/he is free to leave. Take the students to the assessment location with the assessors and assessments starts. If a student who was selected, changes his/her mind as the EGRA assessment is initiated, the enumerator should let you know so that you can select another student of the same gender. You do not have to collect the card, students can keep them but remember that we need the number of boys and the number of girls who were present that day.

2.7 During the EGRA Assessment

Once you have made the selection, bring the assessors and the students to the designated EGRA area. A short break between assessments is needed to allow the assessors to complete administrative tasks. Remember to replenish supplies and provide drinks and snacks for the assessors.

Ensure that assessors administer the EGRA in accordance with the EGRA administration guidelines, as explained and practiced during training. After they complete the EGRA assessment they will administer the Student Context Questionnaire. When a child has finished, he/she should be given the incentive and the enumerator should guide the student back to the classroom, with thanks for his/her participation.

While the assessment is being administered, you, the team leader, can conduct the head teacher interview. When finished, monitor and observe the assessors throughout the assessment Try to make regular spot checks of student responses and look out for common issues, including:

- o Are the assessors building a rapport with the child?
- o Have the assessors remembered to ask each child if they agree to participate?
- o Is the assessor using the tablet properly?
- o Is one assessor taking too long to administer the EGRA or complete the Student Context Questionnaire?

2.8 EGRA Assessor Workload

On average, application of the EGRA instrument takes about 10-20 minutes per child. We estimate that it should take another 15 - 20 minutes to administer the Student Context

Questionnaire. This means that one assessor can complete 2 assessments per hour. If there are three assessors, they will take around 4 hours to complete 18 assessments. When you have finished observing the lesson and interviewing the head teacher and the Kinyarwanda teacher, you should help with the student assessment. Remember that you also must complete the checklist of EGRA administration.

2.9 Other Instruments being used for Data Collection

In addition to the EGRA assessment and the Student Context Questionnaire, other information about the school context is being collected. Before you arrive at the school you should decide whether you will collect this information or whether another team member selected by you should do it instead. Please examine the table below.

Instrument	Sample	Administered by
Head Teacher Interview	- All head teachers of the schools in the sample. - If the head teacher is absent on the day of the visit, make a note of the absence on your report and interview the deputy head teacher or another staff member who is responsible for the school in the absence of the head teacher. Check below the script we have developed to assist you. Use your own words if you prefer.	Team supervisor
Kinyarwanda Lesson Observation	- After you finish the classroom selection ask the head teacher when the Kinyarwanda lesson for that class takes place. You will need to determine the timing of the observation and interviews. If there is no Kinyarwanda lesson for the selected class for that day ask the head teacher to take you to the next PI Kinyarwanda lesson being taught by that teacher.	Team supervisor or other trained member of the team

Good morning. My name is ____ and I am the team supervisor of the group of assessors who are administering the EGRA assessment to children registered at this school. We are working with the government of Rwanda in its efforts to make children read better. We are collecting information in 304 schools on the reading ability of children enrolled in PI. Because you teach Kinyarwanda to PI students you could help us greatly. We would like to observe one of your classes from beginning to end when you are teaching PI. This information is totally confidential and your name will not appear on any reports or documents. We just need examples of how a Kinyarwanda teacher teaches PI classes. Just do what you are used to doing. I will sit in the back and just observe and write some notes.

Teachers, the same as all of us, do not like to be observed when they are doing their work. They may feel threatened or they may think that you will tell others they do not do their job well. It is up to you, to put the teacher at ease and make him or her trust you. Emphasize the

need for the information to design teacher training programs and other activities to help children learn to read.

You (or the team member you selected) must observe a whole class. So if the teacher is halfway through his/her class, ask for the time the next PI Kinyarwanda class starts and return during time.

If you have explained your purpose and mission and the teacher does not allow you to observe his/her class, thank the teacher and see if there is another Kinyarwanda teacher on your list that will teach PI. If there is no other teacher teaching Kinyarwanda to PI, then write on your report that the only Kinyarwanda teacher declined to participate and move on to your other activities.

2.10 Typical Problem Scenarios during Data Collection

Every data collection effort involving so many schools and people will sooner or later run into problems. We have prepared a brief list of problem-scenarios you might encounter when you are visiting schools. Deal with any problems immediately. If you see any common mistakes during EGRA assessment, correct it right away and then raise the issue with the whole assessor team when you review the visit later in the day.

PROBLEM

A child does not want to take part in the EGRA assessment.

SOLUTION

Children do not have to participate if they don't want to. If a child does not agree to take part or to continue the assessment even after it has started, thank the student and send him/her back to the class and select another child following the procedure you have learned. If a student cannot, or will not, respond to a specific sub-test, but is continuing the test, this should be recorded as auto stop on the task.

PROBLEM

Noise and interruptions make it difficult to carry out the assessment.

SOLUTION

If this is a persistent problem, it may be necessary to identify another location at the school site. If children are causing noise by playing too closely to the assessment area and there is no other adequate location ask a teacher to move the noisy children away.

PROBLEM

An assessor's tablet does not work or has run out of power.

SOLUTION

Check the tablet functions yourself. If it does not work, ask the assessor to use the paper version of the instrument. The assessors will be trained in using the paper version of EGRA.

PROBLEM

The child appears tired and is unwilling to continue the assessment.

SOLUTION

If a child looks tired and unwilling to continue the assessment, be friendly, smile, offer praise for the child's effort and send him/her back to the classroom. Sometimes it may be necessary to discontinue the assessment if a child cannot stay on task. For ethical reasons, if a child refuses to carry on, the test cannot continue. If the child appears unwell, stop the test, send him/her back to the class and inform the teacher.

PROBLEM

The data collection team arrives but the school is closed.

SOLUTION

Call the designated person at Soma Umenye (for the intervention schools) or at Incisive Africa (for the control schools) to inform them and they will direct you to another school. If the school is open but the head teacher is not expecting the visit, show them the notification letter from REB (REB has granted permission

for EGRA to be carried out in schools). However, if the head teacher still does not give permission for the assessments to be carried out in their school, call the designated person at Soma Umenye or at Incisive Africa to inform them and they will direct you to another school.

2.11 End of Visit Checklist

Before you and your team leave the school, please use the checklist below to make sure that you have completed all tasks and have left nothing behind.

	Activities Completed	✓
1	Assessed 18 students—9 boys and 9 girls (unless there were fewer than 9 boys and 9 girls in the class) randomly selected for assessment including EGRA reading tasks and student context questionnaire.	
2	All teacher, classroom, and head teacher data have been collected—Teacher Interview, Head Teacher Interview, and Lesson Observation.	
3	Collected all equipment used.	
4	Ensured that data on tablets is uploaded and backed up.	
5	Returned the furniture in the assessment room to its original state	
6	Had a quick debrief with the assessors to identify any potential issues before leaving the school.	
7	Thanked the head teacher and his/her staff for their cooperation.	

We hope that this manual will help you and your team to collect the quality data that we need. Thank you in advance for your efforts. Please check the annexes that follow. Annex A is the form to sign by the person who witnessed that you conducted the random sampling of the class and of the students to be assessed. Annex B is a brief report that you must complete after each school visit and then integrate into one WEEKLY report. Annex C (EdIntersect form used during data collection).

THANK YOU AND GOOD LUCK!

Annex A

Form to be signed by the class/students random selection witness ¹

Date: ____/____/2017

School Name

Sector District

I, _____ as the supervisor of the EGRA data collection team that visited the above school herewith confirm that I briefed the school director and 3rd party witness about the class and student sampling protocol as outlined in the briefing manual and as trained, with my signature I hereby confirm that I am personally liable for any and or mistakes or omissions that I might have committed when briefing the school director and witness briefing.

Signature _____

Date: ____/____/2017

WITNESS STATEMENT

I, _____ herewith confirm that after being fully briefed about the class and student selection process for the EGRA assessment. Do confirm that the procedure was adhered to as I was briefed.

Signature _____

Date: ____/____/2017

Telephone number: _____

Name of school director _____

Signature _____

Date: ____/____/2017

¹ The witness could be a community representative, the head teacher, a teacher or any school administrative staff.

ANNEX B

Sample School Schedule

	Enumerator 1:	Enumerator 2:	Enumerator 3:	Supervisor
07:00	Arrival at school			
07:00 to 08:00	Meeting with head teacher / selection of PI teacher and sections, sample students, identify place to conduct EGRA			
08:00 to 08:30	EGRA & Student Q	EGRA & Student Q	EGRA & Student Q	Class Observation
08:30 to 09:20	EGRA & Student Q EGRA & Student Q	EGRA & Student Q EGRA & Student Q	EGRA & Student Q EGRA & Student Q	Observation
Pause/recreation				
09:40 to 10:10	EGRA & Student Q	EGRA & Student Q	EGRA & Student Q	Teacher Q
10:10 to 11:30	EGRA & Student Q EGRA & Student Q	EGRA & Student Q EGRA & Student Q	EGRA & Student Q EGRA & Student Q	Head Teacher Q
11:30	1 st Shift closes Data collection should be finished by 11:30 or shortly thereafter			

ANNEX C

Daily Team Supervisor Report

Date ____/____/____ Name of school:

Head teacher was at the school when team arrived

Head teacher had been informed that EGRA assessment would take place

Location for assessment was adequate

Number of boys assessed using EGRA

Number of girls assessed using EGRA

Number of boys who did not want to participate in the EGRA

Number of girls who did not want to participate in the EGRA

Number of schools where Kinyarwanda teacher was present as the team started the random selection of classes

Number of schools where head teacher was present when team arrived

Number of Kinyarwanda classes with rosters

Number of students registered in the Kinyarwanda class observed

Number of students present on the day of the visit in the Kinyarwanda class observed

SUMMARY OF ACTIVITIES

Please indicate the exact numbers below:

--	--

SCHOOL RELATED PROBLEMS/SOLUTIONS

(difficulty making contact with the school, accessibility of the school, school not open, selection of classes and students, absence of the head teacher, teacher or community representative, methodology, instrument administration

Problems	Decisions/Solutions

General problems encountered on our end (for example, problems with tablets, materials missing, transportation and logistics/operations).

--

Next Steps

ANNEX D

Weekly Team Supervisor Report

FIELD SUPERVISOR DATA COLLECTION REPORT CHEMONICS/EDINTERSECT • IBTCI/INCISIVE AFRICA

PERIOD

30 September 26 - October 03 - 07 October 10 - 14

Please indicate the exact numbers below:

Total number of schools completed		Number of Lesson Observations conducted	
Number of intervention schools completed		Number of interviews with head teachers conducted	
Number of control schools completed		Number of schools where the head teacher was present when the team arrived	
Number of EGRA assessments administered		Number of schools where all PI teachers were present as the team started the random selection of classes	
Number of boys assessed using EGRA		Number of Kinyarwanda classes with rosters	
Number of girls assessed using EGRA		Number of students registered in the Kinyarwanda class observed	
Number of teacher interviews conducted		Number of students present on the day of the visit in the Kinyarwanda class observed	

SUMMARY OF ACTIVITIES

SCHOOL RELATED PROBLEMS/SOLUTIONS

(difficulty making contact with the school, accessibility of the school, school not in function, selection of classes and students, absence of the head master, teacher or community representative, methodology, instrument administration

Problems	Decisions/Solutions

General problems encountered on our end (for example, problems with tablets, materials missing, transportation and logistics/operations).

--

Next Steps

ANNEX E

Checklist of Required Materials

CHECKLIST OF REQUIRED MATERIALS

		YES	NO
1	Map of the school location (map of the District and the sector of the school location)		
2	Name of the school and school contacts		
3	Work plan (with the specific sequential schools to be visited)		
4	Student incentives (pencil and eraser)		
5	Plastic folder to keep papers, maps, etc.		
6	Checklist for a daily control/verification		
8	Latest version of the EGRA & survey tool/charged tablets		
9	Latest version of the supervisor's manual		
10	Sample tickets (supervisors)		
11	Modem or talktime for using phone as hotspot (supervisors)		
12	Batteries (always have backup batteries)		

CHECKLIST OF VERIFICATION TO GUARANTEE THE QUALITY OF THE DATA

		YES	NO
1	The evaluation team arrived at school on time/		
2	All the enumerators had the complete evaluation kit with them in the field (Check the kit with the checklist. If there is any item missing, report NO and indicate the number of those who did not have the complete kit. Report the fault to the supervisor and ask to proceed towards its immediately correction)		
3	As the head of the team or the supervisor arrived at the school, he/she went to the principal/director's office to inform and explain the enumerators' work		
4	The Head Teacher was at the school on time		
5	The teachers started the classes on time according to the official schedule of the school (Confirm the official schedule for each school).		
6	The process of selecting a sample of classrooms was quick and efficient. The enumerators knew exactly what to do.		
7	The process of selecting a sample of students for the EGRA assessment was quick and efficient. The enumerators knew exactly what to do.		
8	The enumerators knew exactly what their functions were: manage the		

	EGRA evaluation, observe the classrooms or interview the teacher		
9	The members of the team (supervisors and enumerators) worked in harmony		
10	The enumerators developed activities and actions to motivate the students to participate and respond to the EGRA evaluation without pressure or unfriendliness		
11	The enumerators master the mother tongue or the daily/usual language of the students		
12	The data collection supervisor constantly monitored the field team, intervening whenever necessary either in a preventive or proactive manner		

CHECKLIST OF ACTIONS TO GUARANTEE THE QUALITY OF THE DATA

		YES	NO
1	Check if the team has all the required material before going to the evaluation field (see checklist of required materials)		
2	Proceed to the quality control of the field activities as established (see checklist of verification)		
3	Visit all the schools selected for that specific date and monitor the assessment procedures to guarantee the proper implementation of the methodology		
4	Control and guarantee the quality of the daily selection of samples		
5	Interview the Head Teacher		
6	Conduct daily/weekly meetings with the field supervisor to update the status of the assessment, plan the work and to discuss challenges, constraints, solutions, etc. Report all the key aspects of the agreed decisions		
7	Upload the complete data of the selected schools daily		

ANNEX F: ASSESSOR OBSERVATION CHECKLIST

Supervisor: _____ **Assessor:** _____

Date: _____

PRIOR TO TEST ADMINISTRATION	Yes	No	Remarks
Assessor prepares the assessment space appropriately (no materials on the table/desk except the pupil stimuli) and all materials are ready when the child sits down. Switches off			
Assessor is relaxed, makes the child feel comfortable and establishes a good rapport.			
Assessor reads the assent verbatim. In the case where the child does not wish to participate, the assessor thanks the child and escorts them back to their class.			
ADMINISTRATION OF SPECIFIC SUB-TASKS	Yes	No	Remarks
Sub-task 1: Listening Comprehension			
Assessor reads the story once, in a clear, loud voice, at a good pace, and without making mistakes.			
Posture: Assessor holds the tablet in a way that allows for efficient marking and so that the child cannot see the screen.			
Assessor asks comprehension questions and leaves a 10 second response time before moving to the next question.			
Assessor marks student response as “correct” or “incorrect” or “no response” on the tablet.			
Assessor poses all relevant questions to the child without making any mistakes.			
Sub-task 2: Letter Name Identification			
Assessor turns to correct page in the stimulus book and places it in front of the child before beginning to read the			
Instructions: Assessor reads instructions to the child verbatim without adding or skipping words.			
Assessor discontinues the sub-task if the child’s responses to the first 10 letters (first line) are incorrect.			
Posture: Assessor holds the tablet in a way that allows for efficient marking and so that the child cannot see the screen.			
Timer: Assessor starts the tablet timer at the moment when the child starts to say the first item (not before or after) and stops it if a child has read all of the items BEFORE the allowed time is up.			
Marking of responses: Assessor records student responses quickly, accurately, and using the appropriate markings and			
Pace: If a child hesitates for 3 seconds on a given item, the assessor tells the child to “go on” and points to the next letter.			
Last item attempted: assessor accurately marks the last item attempted without delay, and without needing to change his/her			

Transition: At the end of the sub-task, the assessor thanks the child and quickly continues to the next sub-task, without making comments to the child.			
Sub-task 3: Syllable Sound Identification			
Assessor turns to correct page in the stimulus book and places it in front of the child before beginning to read the			
Instructions: Assessor reads instructions to the child verbatim without adding or skipping words.			
Assessor discontinues the sub-task if the child's responses to the first 10 syllables (first line) are incorrect.			
Posture: Assessor holds the tablet in a way that allows for efficient marking and so that the child cannot see the screen.			
Timer: Assessor starts the tablet timer at the moment when the child starts to say the first item (not before or after) and stops it if a child has read all of the items BEFORE the allowed time is up.			
Marking of responses: Assessor records student responses quickly, accurately, and using the appropriate markings and			
Pace: If a child hesitates for 3 seconds on a given item, the assessor tells the child to "go on" and points to the next letter.			
Last item attempted: assessor accurately marks the last item attempted without delay, and without needing to change his/her			
Transition: At the end of the sub-task, the assessor thanks the child and quickly continues to the next sub-task, without making comments to the child.			
Sub-task 4: Familiar Words			
Assessor turns to correct page in the stimulus book and places it in front of the child before beginning to read the directions.			
Assessor reads the instructions.			
Posture: Assessor holds the tablet in a way that allows for efficient marking and so that the child cannot see the screen.			
Assessor puts on the table the materials needed before starting the sub-task			
Assessor quickly and accurately marks the child's responses and moves to the next item.			
Sub-task 5a: Oral Reading Passage			
Assessor turns to correct page in the stimulus book and places it in front of the child before beginning to read the			
Instructions: Assessor reads instructions to the child verbatim without adding or skipping words.			
If the learner hesitates for 3 seconds before attempting the next item, assessor says "please go on" and points to the next word.			
Assessor stops task if no words correct in first line <u>of the stimuli</u>			
Posture: Assessor holds the tablet in a way that allows for efficient marking and so that the child cannot see the			

Timer: Assessor starts the tablet timer when the child starts to say the first item (not before or after) and stops it if a child has read all of the items BEFORE the allowed time.			
Marking of responses: Assessor records pupil responses quickly, accurately, and using the appropriate markings and			
Last item attempted: For timed sub-tasks, assessor accurately marks the last item attempted without delay, and without needing to change his/her mark.			
Transition: At the end of the sub-task, the assessor thanks the child and quickly continues to the next sub-task, without making comments to the child.			
Sub-task 5b: Reading Comprehension			
Assessor removes the pupil stimuli before posing comprehension questions.			
Posture: Assessor holds the tablet in a way that allows for efficient marking and so that the child cannot see the screen.			
Assessor reads questions in a clear, loud voice, at a good pace, and without making mistakes.			
Assessor asks comprehension questions and leaves a 10 second response time before moving to the next question.			
Marking of responses: Assessor records pupil responses quickly, accurately, and using the appropriate markings and			

Student Context Questionnaire (if applicable for your)			
Assessor reads the questions as written clearly and accurately.			
Assessor quickly and accurately marks the child's responses and moves to the next item.			
Assessor does not read response options to the child.			
TABLET USE	Yes	No	Remarks
Orientation: Assessor changes the orientation of the tablet from vertical to horizontal depending on which is best for a given sub-task.			
Data input: Assessor quickly and accurately selects information from drop-down menus and selects other data and information.			
Use of tablet: Assessor uses the finger with ease and can scroll down the screen at the same pace as the child.			

ANNEX G. THE EGRA INSTRUMENT

The EGRA Exam

Because Soma Umenye may still be using this version of the EGRA, it is not included in the DEC (public) version of the report.

ANNEX H. STUDENT CONTEXT FORM

6. Pupil context interview/ Ikiganiro ku mibereho y’umunyeshuri	
	✖ Baza buri mwana ikibazo mu ijwi riranguruye nk’uganira na we. Wisoma uranguruye ibisubizo binyuranye byatanze. Tegereza ko umwana asubiza noneho wandike igisubizo yatanze mu mwanya wabugenewe cyangwa ushyire akamenyetso X mu kazu kanditsemo ibihwanye n’igisubizo umwana yatanze. Read each question out loud to the child, as in an interview. Do not read the response options out loud. Wait for the child to respond, then write their response in the space provided, or put an X in the box of the option that corresponds to the child’s response.

Section 1: Languages				
Ikiciro cya 1: Indimi				
1.	At home, who do you live with as your guardian or guardians?/ Mu rugo ubana na bande bakurera/nande ukurera ?	Mother and/or Father/ Mama na/cyangwa Data/ Papa		Grandparents / Sogokuru na Nyagokuru
		Extended family/ Abandi bavandimwe		Other. Specify/ Abandi. Bavuge
2.	What language (s) do you speak at home? [Multiple responses are allowed] Mu rugo, muvuga uruhe rurimi/izihe ndimi? [Ibisubizo binyuranye birashoboka]	Kinyarwanda/Ikinyarwanda		English/ icyongereza
		French / Igifaransa		Kiswahili /Igishwahili
		Other, Specify Izindi. Zivuge. _____		Don’t know/ No answer / Simbizi
3.	What language (s) does your Kinyarwanda teacher speak in class? [Multiple responses are allowed] Umwarimu wanyu w’ikinyarwanda akoresha uruhe ururimi /izihe ndimi/ mu ishuri? [Ibisubizo binyuranye birashoboka]	Kinyarwanda		English/ icyongereza
		French / igifaransa		Other, Specify/Izindi. Zivuge _____
		Don’t know / No answer / Simbizi		

Section 2: Socioeconomic Status Ikicro cya 2: Urwego rw'imibereho mu by'ubukungu		Yes / Yego	No / Oya	Don't know/Simbizi. No answer / Nta gisubizo
4.	Does anyone at your house have a mobile /telephone? Ese mu rugo iwanyu hari umuntu waba atunze terefone igendanwa?			
5.	<i>If yes to question 4: You told me someone in your family has a mobile phone. Do you ever read anything on it?</i> [Niba ari yego ku kibazo cya 4]: Wambwiye ko hari umuntu mu murungu wawe ufite terefoni igendanwa. Waba hari icyo wigeze usomera kuri icyo terefone?			
6.	Do you have piped water (robinet) at your home? Ese mufite robinet y'amazi mu rugo iwanyu?			
7.	Do you have a radio at your home? Ese mufite radiyo mu rugo iwanyu?			
8.	Do you have a television at your home? Ese mufite tereviziyo mu rugo iwanyu ?			
9.	Does anyone at your house have a bicycle? Ese mu rugo iwanyu hari uwaba atunze igare?			
10.	Does anyone at your house have a motorcycle? Ese mu rugo iwanyu hari uwaba atunze ipikipiki?			
11.	Does anyone at your house have a car? Ese mu rugo iwanyu hari uwaba atunze imodoka?			
12.	What type of light do you use at home? <i>[Multiple responses are allowed] do not read response options</i> Ni ubuhe bwoko bw'amatare mukoresha mu rugo? [Ibisubizo byinshi biremewe]?	Candles/ buji		Electric light bulb/Itara ry'amashanyarazi
		Paraffin lamp/Itara rya peteroli		Solar panel lamp/Itara rikoreshwa imirasire y'izuba
		Biogas lamp/itara rya biyogaze		Torch/flashlight / Itoroshi
		Other/ Ikindi		No answer/Don't know/ Simbizi
Section 3: Hunger, attendance and repetition Ikicro cya 3: Ibijyanye n'amafunguro, ubwitabire ku ishuri no gusibira		Yes / Yego	No / Oya	Don't know/ No answer / Simbizi
13.	Did you eat something before			

	coming to school today? (at home) <i>Uyu muni waba hari cyo wariye (mu rugo) mbere y'uko uza ku ishuri?</i>			
14.	Did you drink something before coming to school today? (at home) <i>Uyu muni waba hari cyo wanyoye (mu rugo) mbere yuko uza ku ishuri?</i>			
15.	Did you go to school before P1? (nursery, pre-school, kindergarten) Waba warize mu ishuri ry'inshuke (materineri, garidiyene) mbere yo gutangira umwaka wa mbere?	Yes/ Yego	No/Oya	Don't Know. No answer/ Simbizi. Nta gisubizo
16.	What class were you in last year? <i>Umwaka ushize wigaga mu mwaka wa kangahe?</i>		Nursery School/ ishuri ry'inshuke (materineri, garidiyene)	P1
			P2	P3
			Not in school/ Sinigaga	Don't Know. No answer/ Simbizi
17.	Were you late to school yesterday (or the last school day you had)? Ejo hashize cyangwa umunsi uherukira ku ishuri niba ejo utarize, waba warakererewe kugera ku ishuri?	Yes / Yego	No / Oya	Don't know/ No answer / Simbizi
18.	If yes to Q17 why were you late for school? [Multiple choices are allowed] Niba ari yego ku kibazo cya 17, kuki wakererewe? [Ibisubizo byinshi biremewe]		Need to do chores (Gukenerwa mu gukora imirimo yo mu rugo)	Go work in the field (Kujya gukora mu murima)
			Waiting to eat (Gutegereza kurya)	Go to market (Kuntuma ku isoko)
			Help care for other children (Kurera barumuna bange)	Sleep (kuryamira)
			Went to play with my friends (Mba nagiye gukina n'inshuti zange)	Sick/not feeling well (Nari ndwaye. Ntabwo nari meze neza.)
			Long distance to school (Urugendo rurerure kugera ku ishuri)	Tend Livestock (Kuragira)
			Other (ikindi)	Don't Know. No answer/ Simbizi

19.	Last week, how often were you absent from school?		0 days. Not absent Sinigeze nsiba.		1-2 days Hagati y’umunsi umwe n’iminsi ibiri.
			3-4 days. Hagati y’iminsi 3-4		5 days –everyday Iminsi 5- Buri munsi
	Mu cyumweru gishize wasibye ishuri kangahe?	Don’t Know. No answer/ Simbizi			
Section 4: Access to books and reading practices Ikiciro cya 4: Kubasha kubona ibitabo no kwimenyereza gusoma		Yes / Yego	No / Oya	Don’t know/ No answer / Simbizi	
20.	Do you take Kinyarwanda books home from the classroom? Ese ujya utahana mu rugo ibitabo by’Ikinyarwanda ubivanye mu ishuri?				
21.	Is there a place in the community where you can go to read/borrow Kinyarwanda books/ Story books? Aho mutuye, haba hari aho ushobora gusoma cyangwa gutira ibitabo by’ Ikinyarwanda cyangwa iby’inkuru?				
22.	Do you ever lend or borrow a book or other learning materials to/from other students? Ese waba ujya utiza cyangwa utira ibitabo byo gusoma cyangwa ibindi bikoresho by’ishuri abandi banyeshuri?				
23.	Do you participate in any reading activity after school? Ese ujya witabira ibikorwa ibyo aribyo byose byo gusoma nyuma yo kuva ku ishuri?				
24.	Do you enjoy reading in a group with other children? Ese waba unezezwa no gusomera hamwe n’abandi bana?				
25.	At home, does someone read a story to you?				

	Mu rugo, haba hari umuntu ujya ugusomera inkuru?			
26.	At home, do you read to someone out loud?/ <i>Mu rugo, haba hari uwo ujya usomera uvuga cyane?</i>			
27.	Do you read on your own at home?/ Iyo uri mu rugo ujya usoma ibitabo?			
28.	Do you have a favorite book? / Hari igitabo ukunda by'umwihariko?			
29.	If yes, What's the title/ Niba ari yego, kitwa gute?			

Section 5: Family/work responsibilities Ikiciro cya 5: Imirimo yo mu rugo umwana akora

30.	At home, what chores do your parents/caregivers expect you to do regularly? [Multiple responses are allowed] Ni iyihe mirimo yo mu rugo ababyeyi bagusaba gukora kenshi? (Ibisubizo byinshi birashoboka)		Wash clothes/ Kumesa imyambaro	Go work in the field/Gukora mu murima (guhinga)
			Collect firewood/Gutashya	Prepare food/ Guteka
			Go to the market/ Kuja guhaha ku isoko	Tend livestock / Kuragira
			Help with other children in the family/ Kurera barumuna bange	Fetch water/ Kuvoma
			Study or read/ Kwiga/Gusoma	No response /Nta gisubizo
			Other /ikindi	

**Section 6: School work feedback and interaction
Ikiciro cya 6: Kuganira ku mukoro wakorewe mu ishuri**

		Yes / Yego	No / Oya	Don't know/ No answer / Simbizi/ Ntagisubizo
31.	Does your Kinyarwanda teacher check the work that you do during class? <i>Umwari mu w'ikinyarwanda ajya areba imyitozo ukorera mu ishuri?</i>			
32.	Do you show your parents/ caregivers your homework? <i>Ababyeyi bawe/abakurera bajya bagenzura umukoro wawe wakoreye mu rugo?</i>			

ANNEX I. LESSON OBSERVATION FORM

Soma Umenye**Lesson Observation Form**

Instructions for the observer: This observation form should be completed during a lesson. Discuss with the teacher where to be positioned in the class to keep distraction to a minimum and to be able to hear and see the teacher and the learners. Please read the appropriate lesson plan and /or teacher's guide prior to observation. Refer to the lesson plan as you observe the lesson. Please take notes while observing the teacher and the students.

Look at the students' written work to appreciate the students' understanding. The observation should last through the entire lesson. Read the statements and check the box that best describes your observation. Check one box for each statement.

A. School Information		B. Date:	
Name:	Code:	Province:	District:
Sector:	Cell:	Village:	
GIS coordinates:			
C. Teacher and Class Observed			
1. Family Name: _____		First Name: _____	Telephone # 1: _____
		Telephone # 2: _____	
2. Class: ☐ P1 ☐ P2 ☐ P3		Class group/shift : _____	
4. Term:		Unit:	Lesson Number:
D. Observer' information			
2. What time did the lesson started?			
3. What time did the lesson ended?			

Lesson preparation	Yes	No
Does the teacher use the teacher's guide?		
Does the lesson plan have an objective?		
Is the student attendance register filled on the day of the lesson?		
Teaching methodology		
Students were given an opportunity to do individual work		
All students actively participated in the lesson activities		
The teacher provided positive verbal reinforcement/appreciation to learners		
The teacher equally encouraged boys' and girls' participation		
Teaching and learning aids were used during the lesson		
Learners' learning was scaffolded using the "I do, we do, you do" teaching method		
Teacher respected the times as described in the lesson plan		
Literacy technical skills		
The lesson reflected at least two core reading skills		
phonological awareness		
phonics		
reading fluency		
comprehension		
vocabulary		

Students in the class had the opportunity to handle/use books during the lesson (including learners with SEN)		
Learners were motivated and concentrated on the lesson		
Assessment		
Instructions for evaluations were clearly explained and understood by the learners		
Learners were prepared to answered questions during the lesson:		
More than half the learners		
Half the learners		
Less than half		
The teacher asked children what they have learned from the lesson		
Learners' learning progress was systematically and effectively monitored		
Evaluation activities were fully inclusive for learners with SEN		
Learners were asked to continue to practice what they have learned at home		

Use of teaching and learning materials

72. To deliver this lesson, the teacher and learners used (tick all that apply):

- ☐ The teacher's guide provided by Soma Umenye and REB
☐ The chalkboard
☐ Read Aloud Books
☐ Student textbooks provided by Soma Umenye and REB
☐ The supplementary reading materials provided by Soma Umenye and REB
☐ Alphabet Chart
☐ Pocket board

☐ Class library
☐ Class dairy
☐ Teacher's exercise book
☐ Other pedagogical documents

☐ Other (please specify): _____

CLASSROOM INVENTORY

Boys present in this class/shift at the time of your observation? [COUNT THE BOYS]	Boys	
Girls present in this class at the time of your observation? [COUNT THE GIRLS]	Girls	
Attendance register available and completed daily	☐ Yes ☐ No	
Total enrolled Learners in this class/shift	Boys:	Girls:
Learners with SEN	Boys:	Girls:
Repeaters in this class/shift	Boys:	Girls:

Age range of learners in this class/shift	Between _____ and _____ years old
To determine the number of Kinyarwanda textbooks available, please ask the learners to hold their Kinyarwanda textbook. [IF BOOKS ARE IN CLASSROOM LOCK/STORAGE, REQUEST THAT TEXTBOOKS BE DISTRIBUTED TO CHILDREN] Number of learners with Kinyarwanda textbook (approved by REB) Do Kinyarwanda textbooks look used?	# of children ☐ No ☐ Yes

ANNEX J. ADDITIONAL RESULTS

ANNEX J. ADDITIONAL NOTES AND RESULTS

Details on Student-level Outcomes

The student-level analysis utilizes student scores from all the sub-tests of the EGRA administered to students, as described in Section 4.2. For all the sub-tests, we generated several outcomes in line with the protocol outlined in the Early Grade Reading Assessment (EGRA) Toolkit, Second Edition (RTI International, 2015). We describe these outcomes in Table Annex I.I below.

Table Annex J.I. Outcomes related to each EGRA sub-tests and definitions

Score	Definition
Raw score	Sum of correct responses
Percent correct	Raw score divided by the number of items in the sub-test
Zero score	If scored zero or not
Percent correct attempted	Raw score divided by the number of items attempted
Fluency score (for timed sub-tests only)	Raw score divided by the time given subtracted from the time remaining for the sub-test, multiplied by 60 seconds.

To estimate impacts of Soma Umenye on student performance on all of the outcomes listed in Table Annex J.I in each sub-test will require us to make a large number of comparisons between the treatment and the control group. As the number of estimated impacts increases, it becomes more likely that we will find a false positive, meaning that we will conclude there is an effect when there is none. This fact is often referred to as the **multiple comparisons problem**. When only one hypothesis test is conducted, it is common to conclude that an impact is statistically significant when the probability that it is due to chance (and not to the program) is 5 percent. However, when using a 5 percent significance threshold and more than one test is conducted at the same time, the probability of a false positive—that is the probability that a significant impact is due to chance and not to the program—is larger than 5 percent. For example, if we estimate impacts on all outcomes related to the letter identification sub-test—fluency score, raw score, zero score, and percent correct out of attempted—and do not adjust for multiple comparisons, the probability of a false positive is 22 percent.¹ In addition to analysis of multiple outcomes, multiple comparison concerns apply to analysis of multiple subgroups as well.

Our strategy for addressing multiple comparison concerns was to focus the primary analysis of program effectiveness in each sub-test, such as letter identification or oral reading fluency on a single key outcome identified prior to beginning the analysis.² Using a small set of outcomes within each domain makes it less likely that statistically significant findings will emerge by chance. Limiting the primary analysis to a single outcome means that it will not be necessary to do statistical corrections for multiple comparisons. Additionally, because we selected the primary outcomes before beginning the analysis, it prevented us from inappropriately focusing the assessment of program effectiveness on outcomes that happen to emerge as statistically significant (or the perception that this may have been the case).

¹ The probability of finding a false positive when conducting n independent two-tailed t tests, given the null hypothesis is true for each test, is found using the formula $1 - (1 - \alpha)^n$, where α is the significance threshold used in the test.

² This multiple comparison framework is consistent with that suggested by Schochet (2009).

For the four timed sub-tests, we selected the fluency score or the correct items identified/read per minute as the primary outcome for each sub-test, as adequate competency in these skills is a prerequisite for competency in the successively more advanced skills. For untimed sub-tests, we selected the percentage of correctly answered questions as the primary outcome. We list the primary and secondary outcomes for each of the sub-tests in Table Annex J.2.

Table Annex J.2. Primary and secondary outcomes for each EGRA sub-test

Domain	Outcomes
Letter identification (Timed)	
Primary outcome	Correct letters identified per minute
Secondary outcomes	Raw letter score Percent of letters correct Zero letter score Percent of letters correct of attempted
Syllable identification (Timed)	
Primary outcome	Correct syllables identified per minute
Secondary outcomes	Raw syllable score Percent of syllables correct Zero syllable score Percent of syllables correct of attempted
Familiar word reading (Timed)	
Primary outcome	Correct familiar words read per minute
Secondary outcomes	Raw familiar words score Percent of familiar words read correct Zero familiar words read score Percent of familiar words correct of attempted
Reading fluency (Timed)	
Primary outcome	Correct words read per minute
Secondary outcomes	Raw words score Percent of words read correct Zero words read score Percent of words correct of attempted
Listening Comprehension	
Primary outcome	Percent of items correctly answered
Secondary outcomes	Raw score Zero score Percent of items correct of attempted
Reading Comprehension	
Primary outcome	Percent of items correctly answered
Secondary outcomes	Raw score Zero score Percent of items correct of attempted

Impact estimates for the primary student-level outcomes are presented in the main report. We present impacts on the secondary outcomes for each sub-test below.

Table Annex J.3. Impacts of Soma Umenye on Student Letter Identification Skills

Scores	Treatment	Control	Impact
Raw Score	23.21	20.36	2.85**
Percent Correct	23.21	20.36	2.85**
Zero score (%)	23.9	32.1	-8.3***
Percent correct of attempted	57.30	51.08	6.22***
Sample size	2,455	3,011	

Sources: EGRA administration data, 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools.

*** Impact (difference between treatment and control group means) is statistically significant at the .01/.05/.10 level.

Table Annex J.4. Impacts of Soma Umenye on Student Syllable Identification Skills

Scores	Treatment	Control	Impact
Raw Score	13.08	11.02	2.07**
Percent Correct	13.08	11.02	2.07**
Zero score (%)	36.6	45.3	-8.7***
Percent correct of attempted	38.72	33.45	5.27**
Sample size	2,455	3,011	

Sources: EGRA administration data, 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools.

*** Impact (difference between treatment and control group means) is statistically significant at the .01/.05/.10 level.

Table Annex J.5. Impacts of Soma Umenye on Student Familiar Words Identification Skills

Scores	Treatment	Control	Impact
Raw Score	4.92	3.95	0.96**
Percent Correct	9.84	7.91	1.93**
Zero score (%)	59.4	66.7	-7.3***
Percent correct of attempted	27.93	22.57	5.36***
Sample size	2,455	3,011	

Sources: EGRA administration data, 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools.

*/** Impact (difference between treatment and control group means) is statistically significant at the .01/.05/.10 level.

Table Annex J.6. Impacts of Soma Umenye on Student Oral Fluency Skills

Scores	Treatment	Control	Impact
Raw Score	5.44	4.52	0.92
Percent Correct	14.31	11.89	2.41
Zero score (%)	63.4	70.0	-6.6**
Percent correct of attempted	27.52	23.32	4.20
Sample size	2,455	3,011	

Sources: EGRA administration data, 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools

*/** Impact (difference between treatment and control group means) is statistically significant at the .01/.05/.10 level.

Table Annex J.7. Impacts of Soma Umenye on Listening Comprehension Skills

Scores	Treatment	Control	Impact
Raw Score	3.91	3.78	0.13**
Percent Correct	78.26	75.61	2.65**
Zero score (%)	1.4	2.1	-0.7
Percent correct of attempted	69.3	60.21	9.07
Sample size	2,455	3,011	

Sources: EGRA administration data, 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools

*/** Impact (difference between treatment and control group means) is statistically significant at the .01/.05/.10 level.

Table Annex J.8. Impacts of Soma Umenye on Reading Comprehension Skills

Scores	Treatment	Control	Impact
Raw Score	0.69	0.60	0.09
Percent Correct	13.86	12.04	1.81
Zero score (%)	67.6	72.6	-5.0**
Percent correct of attempted	70.10	74.21	-4.10
Sample size	2,455	3,011	

Sources: EGRA administration data, 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated. The analysis accounts for clustering of students within schools

**/* Impact (difference between treatment and control group means) is statistically significant at the .01/.05/.10 level.

Details on Teacher-level Outcomes

As in the student-level analysis, we selected this single composite standardized classroom score as our primary outcome to prevent a multiple comparison problem. Table Annex J.9 shows the primary and secondary outcomes for the teacher-level analysis.

Table Annex J.9 Primary and secondary outcomes for classroom observations

Primary outcome	Secondary outcomes
Composite standardized classroom observation score	Raw total classroom observation score Raw lesson preparation score Raw teaching methodology score Raw literacy technical skills score Raw assessment score Raw use of materials score

Impact estimates for the primary teacher-level outcome is presented in the main report. We present impacts on the secondary outcomes below.

Table Annex J.10. Impacts of Soma Umenye on Different Classroom Observation Scores (Raw Scores)

Dependent Variables	Treatment	Control	Impact
Raw total classroom observation score	20.07	17.88	.45**
Lesson preparation raw score	3.01	3.02	.19
Teaching methodology raw score	5.5	5.34	.27
Literacy technical skills raw score	4.28	4.23	.06
Assessment raw score	3.81	2.85	.96**
Use of teaching materials raw score	3.47	2.54	.92**
Sample size	129	152	

Sources: Classroom Observation 2017.

Note: Differences between treatment and control group means were tested using two-tailed t-tests unless otherwise indicated.

* Impact (difference between treatment and control group means) is statistically significant at the .05 level.

Table Annex J.11. Control Variables for Student-Level Analysis

<i>Student demographics</i>
Student age
Gender
Repeater
Overage for grade
Speak Kinyarwanda at home?
Has piped water
Household asset index
<i>Sector and School Characteristics (at baseline)</i>
Urban/rural status
Number of schools in sector
Has HGSE program (%)
Whether Public school
Has library
Has nursery
Number of PI students*
Number of Kinyarwanda teachers*
Has teacher training curriculum
<i>Sector and School Characteristics (at follow-up)</i>
More than half the students have basic school supplies (pen/pencil/notebook/uniform)
School partnering with another NGO
Distance from district office (km)

Table Annex J.12. Control Variables for Teacher-Level Analysis

Variable description
<i>Teacher background</i>
Teacher age
Gender
Highest level of education
Years of experience
<i>Classroom Characteristics - Sector and School Characteristics (at baseline)</i>
Urban/rural status
Number of schools in sector
Has HGSF program (%)
Whether Public school
Has library*
Has nursery*
Number of PI students*
Number of PI Kinyarwanda teachers*
Has teacher training curriculum
<i>Classroom Characteristics - Sector and School Characteristics (at follow-up)</i>
More than half the students have basic school supplies (pen/pencil/notebook/uniform)
School partnering with another NGO
Distance from district office (km)

U.S. Agency for International Development
1300 Pennsylvania Avenue, NW
Washington, DC 20523
Tel: (202) 712-0000
Fax: (202) 216-3524
www.usaid.gov