EDUCATION DEVELOPMENT CENTER

HARGEISA SOMALILAND

Somali Interactive Radio Instruction Program

(SIRIP)

Report on the 2010 Study

Prepared By: Dr. Mohamed A. Dirir

October 2011

Connecticut, USA

Introduction

The Education Development Center (EDC) launched Somali Interactive Radio Instruction Program (SIRIP) for out-of-school children in three areas inhabited by Somali speakers (Somaliland, Puntland, and South Central) in 2008. The goal of SIRIP was to provide an opportunity to children at school age who were not fortunate to have access to formal education. The reasons for lack of education opportunity for these children include internal displacement, poverty, and lack of parental support. To offer education services to these less fortunate children, EDC created learning centers in Somaliland, Puntland, and South Central Somalia. The centers were staffed with EDC-trained teachers who are skilled in using Interactive Radio Instruction programs.

In order to evaluate the effectiveness of the services for the out-of-school children, EDC has prepared an assessment tool which consisted of mathematics and Somali literacy. The instrument was not intended to measure individual student achievement, or to evaluate schools. But the goal was to examine the effectiveness of the program in general. Hence the assessment tool was a survey that used samples of schools and students within these schools. The sampled examinees were tested twice during the course of a single school year; pretest and posttest. The same test was used at both sessions, and there was no concern on test security because the assessment was secure at EDC headquarters. Neither students nor their teachers were able to see the test material before or after the testing sessions. Education Development Center conducted the first administration of the test in 2009 with first grade students who attended the learning centers in two zones; Somaliland and Puntland.

The results of the 2009 study have shown favorable impact of SIRIP on the literacy and numeracy of the served population. Encouraged by the outcome of the 2009 study, EDC commissioned abroader study that addresses more questions in 2010. Among the enhancements of the new study were the additions of two more grades; grades 2 and 3, and the expansion to two more zones that were served by SIRIP: South Central and Galmudug. More importantly, a control group of students who attend formal schools in served zones was added to the 2010 study.

The pretest was administered in November and December 2010. The post-test was administered in May 2011 using the same instrument with the same samples of students. The following sections of the document explain the design of the assessment, research methodology, the sampling plan, the test administration, the data analyses, and the study results.

Test Development

Purpose of the assessment

The assessment in mathematics and Somali literacy were intended to examine the effectiveness of SIRIP services in helping children who do not have an opportunity to enroll regular elementary schools. Its utilities also include informing public policy on non-formal education for out-of-school children. In addition, it was intended to help decisions about resource allocation and teacher preparation to alleviate certain weaknesses in SIRIP that could be exposed by the test.

Somali Literacy test

The Somali literacy test was based on two documents: The EDC literacy learning objectives developed for the formal schools served by SIRIP, and the curriculum frameworks developed for Somali speaking elementary students. The two documents were similar in content, breadth, and depth. At Grade 1, for example, both EDC and the Somaliland Ministry of Education emphasize on reading and reciting the Somali alphabet. For instance, 39 of the 68 Somali literacy objectives at EDC are recognition of Somali alphabet. Based on the weights given to learning the alphabet by both EDC and the Somaliland Ministry of Education, the test for the out-of-school children was heavily represented by alphabet recognition and reciting.

The Somali literacy test was built with guidelines from the Early Grade Reading Assessment (EGRA). It was recognized early in the decade that the early grades are not tested in most of the national and international assessments (PISA, NAEP). Moreover, it has been found in many studies that reading proficiency at early grades (1-3) is highly correlated with reading achievement at later grades and age. These findings suggest that intervention to increase reading proficiency at a young age would be beneficial at later years of schooling. However, one has to assess reading at early grades in order to launch effective remedial interventions at early grades. As a result, the EGRA protocols and procedures were developed under the sponsorship of the USAID and the World Bank. The main basis of the EGRA procedures are speed (fluency) and comprehension. In other words, most of EGRA tasks are timed. That implies test administrators need to measure, with accuracy, the speed at which each student recognizes or reads the alphabet and simple words.

Countries such Kenya, The Gambia, Senegal, Egypt, and Guyana have thus far used EGRA procedures to assess reading at early grades. Furthermore, EGRA was adopted in languages other than the English language such as Arabic by Egypt, French by Senegal, Kiswahili by Kenya, and Spanish by Peru. In these multi-country studies, the reliability and the validity of the EGRA test were supported. In addition, the Research Triangle Institute carried out studies to validate the EGRA procedures and support its reliability. As a result of these findings, EDC has decided to adapt EGRA for the assessment of grades 1-3 out-of-school students in Somali literacy. The Following chart presents the composition of the Somali literacy test:

Chart 1: Scores on Somali Literacy Tasks

Learning objectives	**# of items**	**Number of score points**	**Assessment and expected responses**
Print orientation	3	3	Recognizes the flow of reading text; begins at top left, flows left to right, and top to bottom
Alphabet recognition	100 letters	LCPM*	Recites the Somali alphabet (both consonants & vowels) as many as she/he can in a minute
Phonemic awareness	12	12	Recognizes phonemes said by the examiner among 12 phonemes presented
Reading words	100 words	WCPM**	Reads 50 1- and 50 2-sylluable words , as many as he/she can in a minute
Invented words	50	WCPM	Reads non-words as many as possible in a minute
Listening comprehension	5	5	Answers comprehension questions from the paragraph read to him/her
Reading *** comprehension	5	5	Answers comprehension questions after reading a short paragraph

*Letters correct per minute ** Words correct per minute *** Not tested in grade 1

The Somali literacy test was administered as a structured interview on one-on-one basis. Most of the items were timed, and the test was administered by EDC staff and EDC-trained teachers at the learning centers. The test administrators used stop watches for the timed tasks, and followed guidelines and scoring rubrics that were developed by EDC consultants. The answers were recorded on a paper that contained the test items first and later transferred to a scoring grid that has other information about the examinees. The Somali literacy test took about 15 minutes per examinee.

Mathematics test

The mathematics test was based on EDC mathematics objectives developed for SIRIP, and curriculum frameworks developed for Somali speaking primary school students. The two documents overlap in content focusing on the same mathematics skills and learning objectives. Both contain a large portion of counting, as well as basic addition and subtraction. For example, some of the skills tested in grade 1 included counting, reading and writing numbers 1-9, ordering numbers 1-9, place-value, recognizing shapes, adding 1-and 2-digit numbers, and subtracting 1- and 2-digit numbers. The skills tested in grade 2 included addition of 2-digit numbers, multiplication, division, and numeric progression. In grade 3, the test included place-value, numeric progression, addition, multiplication, division, geometry, and measurement.

The mathematics test for the study was developed from the contents in EDC's mathematics objectives and mathematics textbook for Somaliland Schools. The composition of the grade 1 test and its scoring rubrics are presented in chart 2. Charts of tests for grades 2 and 3 are presented in Appendix B. The mathematics test was administered as a structured interview on one-on-one basis. It was conducted by trained examiners from EDC staff and SIRIP partners, and took about 15 minutes per examinee. The test examiners followed guidelines and scoring rubrics that were developed by EDC. The results from the interview were recorded on a scoring grid, in which other student information was also recorded. The actual mathematics test is shown in Appendix A.

Chart 2: Scores on Mathematics Tasks in Grade 1

Learning objectives	# of items	Maximum score points	Assessment and expected responses
Counting numbers	2	7	Counts 1-20; starts at 9 and counts thru 15
Writing numbers	2	6	Writes single- and two-digit numbers
Addition	2	4	Adds one- and two-digit numbers without carrying over, and up to 99
Subtraction	2	4	Subtracts one- and two-digit numbers without carrying over, and up to 99
Place value	2	2	Indentifies tens and ones in two-digit numbers
Geometric shapes	2	2	Recognizes two simple geometric shapes
Number sequence	1	2	Recognizes number sequence and fills in missing numbers in a series with 20 at the highest

The Sample

The target population in the study was children who do not have chances to attend formal school, and are taught through SIRIP. At the time of the study, close to 30,000 children in Somaliland, Puntland, Galmudug, and South Central Somalia were enrolled in the program.Practically, it was not feasible to assess each child in the program so a sample was taken from each zone and region. Logically, this choice would make the cost of the assessment affordable, without any loss of information and generalization. Ten regions from the four zones were selected for the study, with number of regions from each zone varying from one in Galmudug to four in Somaliland. Moreover, each region was represented by a single town that has several formal schools and SIRIP centers. The only exception was at the Lower Shabbelle region, which was represented by three towns: Merka, Carbiska, and Ceelasha.Table 1 presents a summary of the number of students pre-tested in each region and town, along with the number attended in formal schools and the number enrolled SIRIP center.

Table 1: Examinee Counts by Region/Town and Type of School

Zone	Town	Grade 1		Grade 2		Grade 3		Total
		Formal	IRI	Formal	IRI	Formal	IRI	
Galmudug								
	Galkaio					128	131	**259**
Puntland	Bosaso	40	40	40	40	39	40	**239**
	Galkaio	29		60	82	38	39	**248**
	Garowe	63	30	64	58	35		**250**
Somaliland	Borama		80	61	61	60	42	**304**
	Burco	40	40	60	61	40	62	**303**
	Hargeisa	42	72	137	77	480	31	**839**
	Laascaanood	64		32	64	100		**260**
South Central	Carbiska				55		54	**109**
	Ceelasha				70		58	**128**
	Merka				41		45	**86**
	Mogadishu	144		146		146		**436**
Total		**422**	**262**	**600**	**609**	**1066**	**502**	**3461**

As shown in the tables, a total of 3461 examinees were tested in December 2010; 684 in grade 1, 1209 in grade 2, and 1,568 in grade 3. The table also highlights that the number of students tested in formal schools were markedly higher than the number of students tested in SIRIP learning centers in grades 1 and 3. Additionally, some of the regions/towns contain empty cells, which could make some of the comparisons among school types difficult. For example, it will not be possible to compare Formal and SIRIP schools in all regions in South Central, or grades 1 and 3 in Laascaanood, or grade 3 in Garowe, or grades 1 and 2 in Galmudug.

In the posttest, which was administered in May 2011, the numbers tested in some locations have changed. In the towns where the differences in counts existed, the number of students in the pretest sample was more than the number of students in the posttest sample.The numbers of children in the posttest were 569, 1064, and 1317 respectively for 1st grade, 2nd grade, and 3rd grade. Expectedly, the formal schools have shown slightly better retention rates than the SIRIP learning centers. The changes in the numbers of enrolled or served students could be attributed to political instability in the areas where the schools/centers were located. There could have been continual insecurity and persistent internal displacement in these areas. The largest changes occurred in the South Central zone while the changes in Somaliland were negligible.

The ages of the target population varied because each child might have left formal school at different point in time due to unexpected circumstances or had never attended school. The distribution of the age of the control group was similar to that of the target group. The median ages of the sample (both groups) were 9 years for grade 1, 10 for grade 2, and 11 for grade 3. The comparable median ages in developed countries are 6, 7, and 8 respectively for grades 1, 2, and 3.

In general, the livelihood might be different between children who attend formal schools and children who attend SIRIP centers. Some of the children served by SIRIP might be engaged in some kind of work for living. Some might be street peddlers, some might be in the shoe shinning service, and others might be waiters. The proficiency levels in numeracy and Somali literacy of the children in SIRIP centers might also varied. Theymight have includedsome who have never attended school and others who may have left school at different grades and at different ages. Hence, some would have spotty knowledge in basic mathematics and Somali literacy, while others might have none. Due to these differences among the group served, extra caution was used in sampling this heterogeneous population.

Finally, the circumstances under which the SIRIP students were in and the settings in which they were taught were markedly different from their peers in regular schools. The SIRIP population faces a host of challenges in schooling, most daunting of which is lack of permanent home. These students might have been displaced from their homes one or more times, and face new beginnings in living conditions and in education each time they migrate. And the detrimental effect of mobility on education and student achievement is well documented. Logically, many of the students in SIRIP might have been displaced from homes where they had attended regular schools with normal teaching-learning environment. They would reasonably need time and effort in adapting to the new setting in SIRIP, and the success in the adaptation could vary among students. Also, this process could have adverse effects on their academic performance and achievement.

Data Analysis

The purpose of the study was to examine the effectiveness of SIRIP in teaching out-of-school children Somali literacy and numeracy. The study was not set to proof that SIRIP is superior to regular classroom learning, but was intended to assess if the program has delivered learning experiences that was not inferior to the regular classroom learning experiences. In other words, the objective of the study was to proof that SIRIP is as effective as regular teaching. To that end two sets of analyses were conducted. First the average statistics and percentages for the whole sample were computed for the pretest and the posttest, and performances by type of school, zone, age group, and gender were examined. Second, Repeated Measures Analysis of Variance (ANOVA) was employed to assess the significances of differences, if any, in performances in pretest and posttest for students who had both scores. The main comparison was by school type, which reflects the instruction delivery mode (Formal or SIRIP.) Rationally, the null hypothesis in the ANOVA was that the instruction time of formal schools was more effective than the instruction time of SIRIP. Rejecting this null hypothesis would indicate that SIRIP is at least as good as the formal school. In addition to the school type, the significance of differential effects by gender and age, as well as their interactions with school type, were also examined.

For the Somali literacy, the scoring procedure on Chart 1 was used. As children at the different grades are expected to differ in literacy proficiency levels, it might be practical to focus at different levels of literacy results for each grade. At 1st grade, for example, it might be sensible to focus easiest parts of the test such text orientation, letters correct per minute (LCPM), phonemes. For the 3rd grade, on the other hand, the focus could be on 1- and 2-syllable words (WCPM), non-words, listening, and reading.

For the mathematics test, the scores for each grade were grouped into math domains. In grade 1, for example, five clusters were formed: Number proficiency, addition /subtraction, placing value, geometric shapes, and number sequence. For number proficiency and addition /subtraction, the average scores were computed. For the other three domains, which were scored in 0-2 rubric, the percent of students who received a score of 1 or 2 were computed. These five indicators were used to evaluate the performance of the 1st grade students on the mathematics test.In grade 2, the mathematics item scores were grouped into four sections: Addition, multiplication, date, and number sequence. Average scores were computed for the first two sections while percent of 1 and 2 scores were computed for the last two. Similarly, grade 3 mathematics scores were grouped into four sections: Numbers, addition, multiplication, and geometry and measurement. The average scores were computed in all

sections of the grade 3 mathematics. Finally, average mathematics total scores were computed in each grade.

Results

Comparisons of average scores

In this section, the analyses results of the pretest and posttest data are presented.Table 2 presents the summary of the Somali literacy test for 1st grade by zone. As the test was developed following the EGRA procedures, the results are presented in EGRA units.The South Central zone did not have SIRIP schools. The last row (All) represents the performance of all students regardless of school type or zone. It shows that students made achievement progress between the pretest and posttest administrations. At this grade, the gains were greater in the easy tasks (LCPM, Phonemes) than in the harder section (WCPM, Non-words). We can examine the relative performance of the SIRIP and Formal schools in the row for the school type. About 66 percent of the students in SIRIP knew the direction to which the text flows during the pretest. That number went up by 14% during the posttest. Further, the average scores for the listening went up more for SIRIP than for formal schools.

Table 2: Average Literacy Scores for 1st Grade

		ORIENTATION: % CORRECT ALL 3		LETTERS/MIN		PHONEMES: % CORRECT		1-SYLLABLE WORDS/MIN		2-SYLLABLE WORDS/MIN		NONWORD/MIN		LISTENING	
		PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Puntland	IRI	54	72	14	40	16	37	8	14	4	10	5	8	2.9	3.7
	Formal	83	89	29	49	63	97	11	23	8	18	7	15	4.0	4.6
Somaliland	IRI	70	86	38	42	53	81	24	23	22	18	21	16	3.0	4.1
	Formal	96	97	36	70	70	84	16	30	11	36	10	30	3.8	4.5
South Central	Formal	97	96	46	64	74	83	22	30	25	30	22	26	4.2	4.7
All Zones	IRI	66	80	30	41	40	57	20	20	18	15	17	13	3.0	3.9
	Formal	93	95	38	64	70	86	16	28	14	30	12	25	4.0	4.6
All Students		**82**	**90**	**35**	**58**	**59**	**78**	**18**	**26**	**15**	**25**	**14**	**21**	**3.6**	**4.4**

Table 3 summarizes the literacy results for 2nd grade. All students in the sample made gains in achievement in all units. The gains were similar for both subgroups; SIRIP and Formal. This could be an indication that the two teaching methods are equally efficient.

Table 3: Average Literacy Scores for 2nd Grade

		ORIENTATION: % CORRECT ALL 3		LETTERS/MIN		PHONEMES: % CORRECT		1-SYLLABLE WORDS/MIN		2-SYLLABLE WORDS/MIN		NONWORD/MIN		LISTENING		READING	
		PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Puntland	IRI	81	90	32	70	83	81	15	34	12	27	10	24	4.0	4.5	1.9	3.2
	Formal	86	98	40	65	85	98	17	31	15	27	12	20	3.7	4.7	2.0	3.7
Somaliland	IRI	94	99	71	71	84	91	37	36	30	29	27	45	4.4	4.8	2.3	4.2
	Formal	94	99	60	70	86	87	30	33	24	29	23	29	4.4	4.8	2.1	4.3
South Central	IRI	67	61	64	74	78	80	29	36	24	27	17	19	3.7	3.3	2.6	3.1
	Formal	99	92	70	75	88	90	35	33	30	28	20	19	4.0	4.0	4.6	3.9
All Zones	IRI	83	87	58	72	82	86	27	35	23	28	19	33	4.1	4.3	2.3	3.6
	Formal	93	97	57	69	86	90	28	32	23	28	19	25	4.1	4.7	2.7	4.1
All Students		**88**	**92**	**57**	**71**	**84**	**88**	**28**	**34**	**23**	**28**	**19**	**29**	**4.1**	**4.5**	**2.5**	**3.9**

Table 4 shows similar results for the literacy test in 3rd grade. At this grade, it might be sensible to focus the comparative results at the high end or more difficult sections such as reading and listening. In all other sections, SIRIP students gained a little while formal school students did not gain at all or slightly lost ground. In the reading and listening sections, however, both groups gained scores. But the SIRIP gains were much higher than the gains for the other students.

Table 4: Average Literacy Scores for 3rd Grade

		ORIENTATION: % CORRECT ALL 3		LETTERS/MIN		PHONEMES: % CORRECT		1-SYLLABLE WORDS/MIN		2-SYLLABLE WORDS/MIN		NONWORD/MIN		LISTENING		READING	
		PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Galmudug	IRI	78	100	35	74	63	58	31	37	25	29	23	27	3.0	4.5	3.7	3.1
	Formal	95	99	74	93	98	0	46	47	30	42	29	38	4.3	4.9	4.2	4.5
Puntland	IRI	79	100	49	63	80	91	24	27	19	21	17	19	3.3	4.3	0.5	3.9
	Formal	100	95	54	69	95	98	27	36	23	30	18	27	3.5	4.6	4.2	4.1
Somaliland	IRI	99	96	66	76	88	90	29	36	25	30	23	26	4.3	4.8	2.1	4.4
	Formal	98	99	84	72	83	90	41	33	35	25	29	24	4.2	4.7	3.6	4.5
South Central	IRI	88	80	79	81	92	95	47	51	33	32	27	24	4.1	4.5	3.6	4.1
	Formal	99	80	87	56	89	64	40	25	34	22	26	16	3.8	3.6	4.6	3.6
All Zones	IRI	90	94	63	75	86	88	35	39	27	29	23	25	3.9	4.6	2.5	4.0
	Formal	98	96	79	72	88	87	40	34	32	28	27	25	4.1	4.5	3.9	4.3
All Students		**95**	**95**	**72**	**73**	**87**	**87**	**38**	**36**	**30**	**28**	**26**	**25**	**4.0**	**4.6**	**3.4**	**4.2**

The results for 1st grade mathematics are presented in table 5. In all sections in math, the gains for the SIRIP schools were higher than the gains for the formal schools. The gain in total math scores, for example was 6.0 points (10.2 – 16.2) for SIRIP and 2.8 points for formal schools. That implies the SIRIP schools gained more than twice as much as the formal schools gained. Similarly, the scores for the

Sequence section went up by 33 percent for SIRIP schools and 11 percent for formal schools. The pattern of the differential gains between the two groups was evident in all math domains in 1st grade. We will test the significance of these differences later in the ANOVA section.

Table 5: Average Mathematics Scores for 1st Grade

		MEAN SCORE						PERCENT SCORED 1 OR 2					
		NUMBERS		ADDITION		TOTAL		PLACE VALUE		RECOGNIZE SHAPE		SEQUENCE	
		PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Puntland	IRI	5.6	9.1	1.4	3.0	8.2	14.8	27	52	29	70	19	28
	Formal	7.9	9.7	3.2	4.1	14.0	17.7	62	80	72	96	43	46
Somaliland	IRI	7.7	9.1	1.7	3.8	11.0	17.3	23	70	56	74	19	71
	Formal	10.6	10.3	3.5	4.1	17.3	18.7	51	76	97	98	41	58
South Central	Formal	9.4	10.2	3.4	4.5	15.5	18.8	34	53	76	88	62	74
All Zones	IRI	7.1	9.1	1.6	3.4	10.2	16.2	24	62	49	73	19	52
	Formal	9.4	10.1	3.4	4.3	15.7	18.5	48	70	83	95	49	60
All Students		**8.5**	**9.8**	**2.7**	**4.0**	**13.5**	**17.8**	**39**	**67**	**69**	**88**	**37**	**58**

The mathematics results for 2nd grade are summarized in table 6. The score gains for SIRIP and formal schools were the same or close in all areas. This suggests that the effects of the two instructional methods might have been similar.

Table 6: Average Mathematics Scores for 2nd Grade

		MEAN SCORE				PERCENT SCORED 1 OR 2				TOTAL SCORE	
		ADDITION		MULTIPLICATION		MONTHS		SEQUENCE			
		PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Puntland	IRI	2.6	4.5	1.7	3.0	45	78	15	37	5.3	9.3
	Formal	3.3	4.1	0.9	2.2	25	76	18	30	4.7	7.9
Somaliland	IRI	2.8	2.8	1.3	1.5	29	39	13	17	4.5	5.1
	Formal	2.7	2.9	1.0	1.8	32	48	6	22	3.9	5.6
South Central	IRI	3.4	3.1	2.1	2.2	49	79	29	43	6.6	6.7
	Formal	4.2	3.7	2.1	2.4	88	77	40	49	8.4	7.8
All Zones	IRI	2.9	3.4	1.7	2.1	40	60	19	29	5.3	6.7
	Formal	3.2	3.4	1.3	2.0	46	61	20	28	5.2	6.6
All Students		**3.1**	**3.4**	**1.5**	**2.0**	**43**	**60**	**19**	**29**	**5.3**	**6.6**

The third grade mathematics results are shown in Table 7. Similar to grade 2, the gains in math scores for SIRIP and formal schools were close in most of the subsections tested. The gains in total math, however was notably larger for SIRIP schools (6.1 to 11.3) than for formal schools (6.6 to 9.9).

Table 7: Average Mathematics Scores for 3rd Grade

		NUMBERS		ADDITION		MULTIPLICATION		GEOMETRY & MESUREMENT		TOTAL SCORE	
		PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Galmudug	IRI	1.0	2.1	1.3	2.5	0.7	1.7	0.1	2.9	2.3	9.3
	Formal	2.8	3.4	3.3	3.4	1.5	1.8	0.3	0.9	7.6	9.6
Puntland	IRI	1.4	1.3	1.6	2.1	1.4	0.9	0.5	0.5	4.9	4.7
	Formal	2.4	2.9	2.5	3.0	1.2	1.4	1.1	1.8	7.3	9.2
Somaliland	IRI	2.6	3.7	2.5	3.8	0.4	2.2	0.6	4.1	5.9	13.8
	Formal	2.1	3.3	2.3	2.9	0.6	1.6	0.9	2.1	5.7	9.7
South Central	IRI	2.5	3.0	2.8	2.8	1.5	2.1	2.3	4.1	9.2	12.0
	Formal	3.1	3.1	3.1	3.4	1.7	2.0	0.5	3.2	8.3	11.7
All Zones	IRI	2.3	2.9	2.3	3.1	0.9	1.9	1.2	3.4	6.1	11.3
	Formal	2.4	3.2	2.6	3.1	1.0	1.6	0.8	2.1	6.6	9.9
All Students		**2.3**	**3.1**	**2.5**	**3.1**	**1.0**	**1.7**	**1.0**	**2.6**	**6.4**	**10.4**

The results for all grades and both subjects indicate that the achievement gains between the pretest and the posttest were close. In fact, the gains by SIRIP schools were slightly higher in many instances. This implies the instructional method implemented in SIRIP was not less effective than the instructional methods used in formal schools. Next, we test the significance of some of the differences in score gains between SIRIP and formal schools.

Repeated Measures Analysis of Variance

As was mentioned earlier in the report, the second phase of the analysis was to test the significance of the score differences between pretest and posttest. Repeated Measures ANOVA was used for this purpose, and only students with scores at both times were included in the analysis. The number of matched examinees varied among grades; 1061 in grade 3, 871 in grade 2, and 512 in grade 1. The main goal in this analysis was to examine whether students achievement gains in SIRIP schools are lower than the achievement of students in formal schools. This was captured in instruction or school type crossed with time or repeated test. The time indicator represents the re-test condition, and examines whether students' scores have increased between the pretest and the posttest. In other words, time assesses the effectiveness of the instruction between the two testing times for both types of schools.

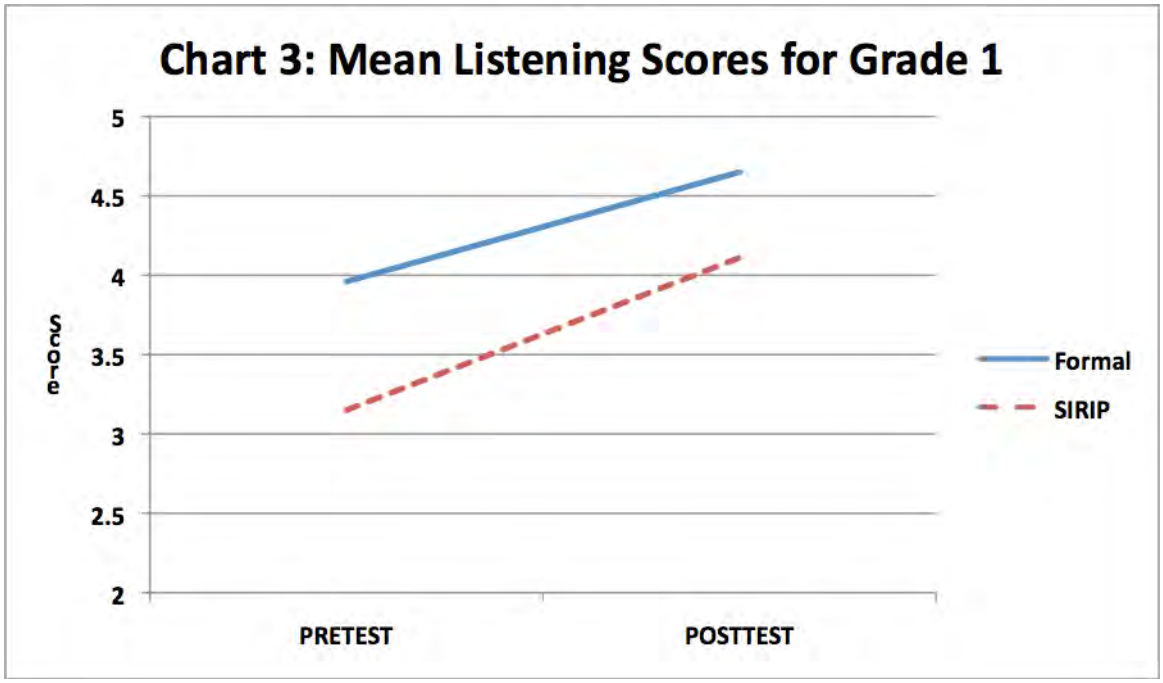
The school type obviously assesses the effects of the instructional program on scores. It tests the significance of the differences in mean scores between SIRIP students and the control sample. The interaction between school type and time enables to test whether the instructions received by the two groups between the pretest and posttest were differentially effective. Two other indicators were added to this analysis: student's gender and age. For the age, two age groups were formed in each grade. The older half was called 'Old' and the younger half was called 'Young'. The significance of differences between the groups was also examined along with their interactions with time and school type. The ANOVA results for main variables such as reading, listening, some EGRA scores, and total math are presented in body of the report. The ANOVA tables for the rest of the scores are reported in Appendix B.

The first ANOVA table, which contains the results of listening test for grade 1, is presented below. Generally ANOVA tests the significance of differences in mean scores among groups. The first row of the table is the significance test of the mean score difference between school types, and there is significant difference as indicated by the F value. The second row is the significance test of the score gains for all subjects. Put it differently, TIME effect assesses the significance of score change due to instruction between the pretest and the posttest. Expectedly, the TIME effect (instruction) is significant in most of scores. The third line tests whether the TIME between the two exams has affected the school type differently. This answers the question: Have the formal school instruction produced better results than the SIRIP instruction? The F Value indicates that the interaction was not significant. Hence we can conclude that the two instruction modes did not result in significant differences in listening scores.

Table 8: Repeated Measures Analysis of Variance
Listening Grade 1

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TYPE (Between Subjects)	1	86.54	86.54	44.52	<.0001
TIME (Within Subjects)	1	128.64	128.64	137.15	<.0001
TIME*Type	1	3.49	3.49	3.72	0.0544
TIME*Gender	1	0.38	0.38	0.41	0.5236
TIME*AGEGROUP	1	0.86	0.86	0.92	0.3378
TIME*Type*Gender	1	1.60	1.60	1.71	0.1921
TIME*Type*AGEGROUP	1	0.65	0.65	0.69	0.4067
Error(TIME)	494	463.36	0.94		

Although the interaction was not significant, it might be helpful to compare the means of the groups in the pretest and the posttest. Chart 3 highlights the changes in listening scores of SIRIP and non-SIRIP students. As can be seen at the slopes of the lines in the chart, the SIRIP students were gaining at a rate higher than the students in the formal schools.



Tables 9 through 11 present the ANOVA results for different scores and grades. In table 9 the difference between means for the school types was significant, and the means between testing times was significant. The only other significant effect was the interaction among time x type x gender. This implies the changes in scores between pretest and posttest was different between boys and girls, and also

Table 9: Repeated Measures Analysis of Variance
One-Syllable words for Grade 2

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TYPE (Between Subjects)	1	2887.53	2887.53	8.17	0.0044
TIME (Within Subjects)	1	12613.58	12613.58	36.25	<.0001
TIME*Type	1	46.44	46.44	0.13	0.715
TIME*Gender	1	997.15	997.15	2.87	0.091
TIME*AGEGROUP	1	1.14	1.14	0	0.9544
TIME*Type*Gender	1	2195.28	2195.28	6.31	0.0123

TIME*Type*AGEGROUP	1	2.84	2.84	0.01	0.9281
Error(TIME)	647	225143.96	347.98		

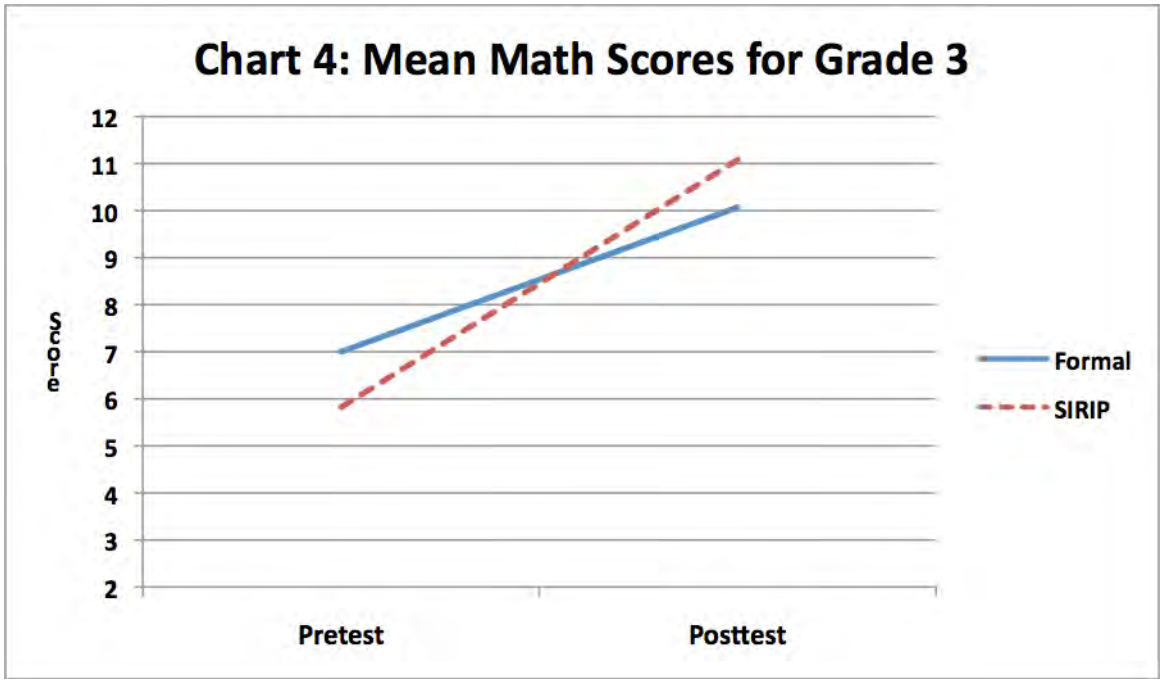
depended on which school type they have attended. A closer look of the data reveals that scores for boys and girls in formal schools were close in the pretest. But boys scored higher than girls in the posttest. For SIRIP students, on the other hand, the boys scored slightly higher than girls in the pretest, and this difference decreased in the posttest.

In table 10, all independent variables and their interactions, except TYPE and TIME x Gender, were significant. This suggests that instruction time had notable effect on score change, and its interaction with other variables (except gender) was also significant. Hence we could say the mathematics instruction was effective, and had differential effects on school type and age group.

Table 10: Repeated Measures Analysis of Variance
Mathematics for Grade 3

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TYPE (Between Subjects)	1	3.59	3.59	0.12	0.7256
TIME (Within Subjects)	1	8535.66	8535.66	375.47	<.0001
TIME*Type	1	591.32	591.32	26.01	<.0001
TIME*Gender	1	27.77	27.77	1.22	0.2693
TIME*AGEGROUP	1	130.68	130.68	5.75	0.0167
TIME*Type*Gender	1	112.69	112.69	4.96	0.0262
TIME*Type*AGEGROUP	1	93.80	93.80	4.13	0.0425
Error(TIME)	1064	24188.09	22.73		

The results in table 10 are presented graphically on chart 4. The SIRIP students were performing lower than their peers at the pretest, but performed higher after five months of instruction. There are other findings in the data that portray similar situations where SIRIP students gained ground on non-SIRIP students.



The AVOVA result grade 3 for reading is presented in table 11. The school type, time and their interaction were all significant. As explained earlier in the report, this indicates that SIRIP and non-SIRIP school means were different in general, and changes in scores from pretest to posttest were different

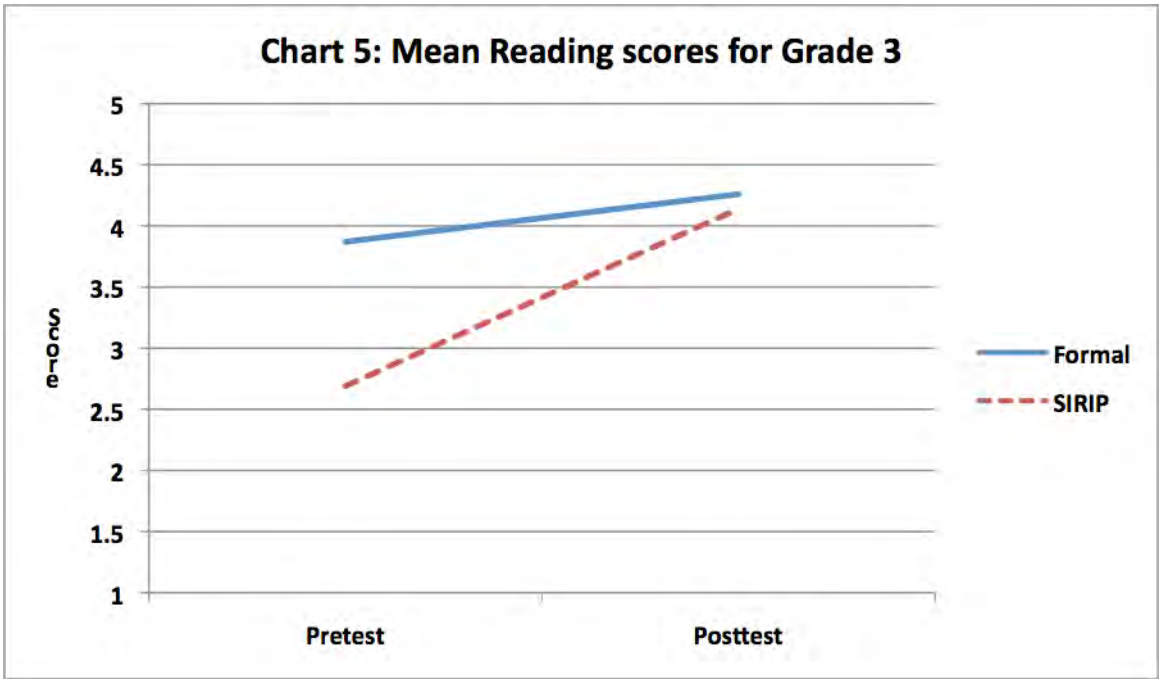
Table 11: Repeated Measures Analysis of Variance
Reading for Grade 3

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TYPE (Between Subjects)	1	167.14	167.14	43.60	<.0001
TIME (Within Subjects)	1	336.25	336.25	171.47	<.0001
TIME*Type	1	113.01	113.01	57.63	<.0001
TIME*Gender	1	1.30	1.30	0.66	0.4158
TIME*AGEGROUP	1	23.39	23.39	11.93	0.0006
TIME*Type*Gender	1	0.10	0.10	0.05	0.8197
TIME*Type*AGEGROUP	1	5.39	5.39	2.75	0.0976

Error(TIME)	894	1753.16	1.96		

for the two groups as well. This could imply that the effects of the instructions were different. The interaction between TIME and age group was also significant, which could mean TIME had different effects on age.

The results in table 11 are shown in a graphical form on chart 5. Again, average reading for SIRIP schools in grade 3 got closer to average reading for non-SIRIP school.



Conclusion

The purpose of this study was to examine the effectiveness of SIRIP as a tool of instruction for children who were not able to attend formal schools. A sample of 3500 students from SIRIP schools and formal schools in grades 1, 2, and 3 were selected for the study. All students took a mathematics test and a Somali literacy test in December 2010. After five months of instruction, the same test was administered to all students in May. The SIRIP and non-SIRIP students did not have many things in common except grade level and perhaps age. Their livelihoods, homes, learning environments, security and safety, and social wellbeing might all have been different. Additionally, SIRIP was not the choice for those who enrolled but the only alternate to an unavailable formal school.

The literacy test was based on EGRA tools, which mainly focus on fluency in early grades. The mathematics test was based on EDC's math objectives and Somaliland's curriculum and textbooks. Students were assessed in eight section of ERGA, and five sections of math problems. The EGRA test was structured interview, while the math test was a mixture of interview and work-alone formats.

The results show that SIRIP students have gained from learning more than non-SIRIP students. Tables 2 through 7 present data that support the average gains for SIRIP students were greater or equal to those of their peers in formal schools. The ANOVA tables (8-11) show the significance tests of the differences in gains between SIRIP and non-SIRIP students. In all cases presented (and many in Appendix B), the score gains for SIRIP students were greater than the score gains for non-SIRIP students. Charts 3 through 5 highlight the differential gains between the two groups. In all cases presented, the SIRIP students were performing lower the beginning of the school year, but closed the gap in performance between them and the non-SIRIP children within five months. It is tempting to imagine how gain scores of the two groups would stack up if the comparisons were for full school year. In fact, this could be predicted with statistical extrapolation.

At the onset of the study, the goal was to evaluate SIRIP, and see if it is less effective than formal schools. But the outcome of the study portrays SIRIP as an effective instructional tool that would help the less fortunate children who could not get into regular schools. It might serve well other groups of children in addition to internally displaced. For example, families who live in remote areas in the country side might benefit from a program like SIRIP. Although this is an education policy issue beyond the scope of the study, but the evidence to support such policy is presented here.