USAID FROM THE AMERICAN PEOPLE | Basic Education Program

The Effectiveness of Interactive Radio English Instruction Programs in Grade One in Ethiopia

A Summative Evaluation Report – BESO/USAID

Teshome Nekatibeb

October, 2001
Addis Ababa, Ethiopia

Acknowledgements

The completion of this report could have been impossible had it not been for the cooperation of many individuals and institutions. At BESO Dr. Tom Tilson has been a great help in providing finances for positions required to complete data entry and analysis. Himself, an outstanding scholar in IRI, Dr. Tilson has also provided technical support in the research effort. Mr. Stuart Leigh, consultant to EMA, has been helpful in this project right from the very beginning. In addition to his activities in planning and implementing the project he has read the draft report of this evaluation and provided useful comments. Ato Demissew Bekele, General Manger of EMA, has been providing all the required support to make this evaluation a success. Ato Ejeta Negeri and W/t Tsehay Bekana, experts in English radio program production, were everywhere and everything in the project. From ICDR, special thanks go to Ato Deresse Mengistu who assisted in data processing and analysis. There are a large number of students, many teachers, Zone media supervisors and school directors who contributed all they could to make this evaluation fruitful. Thank you all!

Table of Contents

Acknowledgements ..... **i**
Executive Summary **1**
Overview 1
Main Findings in the Study 1
Conclusions 2
Part One **4**
1.1. Introduction 4
1.2. Aims and Objectives of the Study 4
1.3. Significance 4
Part Two **6**
Research Methodology **6**
2.1. Design 6
2.2. Instruments of Data Collection 7
2.3. Selection of Samples 8
2.4. Data Management and Analysis 9
Part Three **11**
Findings and Analysis of Data **11**
3.1. Effectiveness of IRI in Improving Educational Quality 11
3.1.1 Student Academic Gains in Control Schools without IRI Programs 11
3.1.2. Students' Academic Gains in Experimental Schools with IRI Programs 11
3.1.3. Students' Academic Gains in Experimental and Control Schools 12
3.2. Effectiveness of IRI in Closing Equity Gaps 13
3.2.1. Students' Academic Achievements in Control Schools by Location 13
3.2.2. Students' Academic Achievements in Experimental Schools by Location 13
3.2.3. Students' Academic Gains in Urban Control and Experimental Schools 14
3.2.4. Students' Academic Gains in Rural Experimental and Control Schools 15
3.2.5. Students' Academic Gains in Experimental and Control Schools by Location 16
3.2.6. Students' Academic Achievements by Sex in Control Schools 17
3.2.7. Students' Academic Achievements by Gender in Experimental Schools 17
3.2.8. Students' Academic Gains in Experimental and Control Schools by Sex 18
3.2.9. Students' Academic Achievements in Control Schools by the Organization of Learning in Grade One 19
3.2.10. Students' Academic Achievements in Experimental Schools by the Organization of Learning in Grade One 20
3.2.11. Students' Academic Gains in Experimental and Control Schools by the Organization of Learning in Grade One 20
3.2.12. Students' Academic Gains in Control Schools by Class Size 21
3.2.13. Students' Academic Gains in Experimental Schools by Class Size 22
3.2.14. Students' Academic Gains in Experimental and Control Schools by Class Size 22
3.2.15 Students' Academic Achievements in Control Schools by Region 23
3.2.16. Students' Academic Achievements in Experimental Schools by Region 24
3.2.17. Students' Academic Gains in Experimental and Control Schools by Region 25
Part IV **27**
Summary, Conclusions and Recommendations **27**
4.1. Summary 27
4.2. Conclusions 29
References **30**

List of Tables

Table 1. Additional Design Activities in Control and Experimental Schools *6*
Table 2. Research Settings and the Participants of Study *9*
Table 3. Students' Academic Gains in Control Schools *11*
Table 4. Students Academic Gains with IRI Programs in Experimental Schools *11*
Table 5. Students' Academic Gains in Control and Experimental Schools *12*
Table 6. Students' Academic Achievements in Control Schools by Location *13*
Table 7. Students' Academic Gains in Experimental (with IRI) Schools by Location *14*
Table 8. Students Academic Gains in Urban Experimental and Control Schools *14*
Table 9. Students' Academic Gains in Rural Experimental and Control Schools *15*
Table 10. Student Academic Gains in Experimental and Control Schools by Location *16*
Table 11. Students' Academic Achievements by Sex in Control Schools *17*
Table 12. Students' Academic Achievements in Experimental Schools by Sex *18*
Table 13. Student Academic Gains in Experimental and Control Schools by Sex *18*
Table 14. Students' Means Academic Gains in Control Schools by Organization of Learning in Grade One *19*
Table 15. Students' Academic Achievements in Experimental Schools by Organization of Learning in Grade One *20*
Table 16. Students' Academic Gains in Experimental and Control Schools by Organization of Learning in Grade One *21*
Table 17. Students' Academic Gains in Control Schools by Class Size *22*
Table 18. Student Academic Gains in Experimental Schools by Class size *22*
Table 19. Students' Academic Gains in Experimental and Control Schools by Class Size *23*
Table 20. Students' Academic Gains in Control Schools by Region *24*
Table 21. Student -Academic Achievements in Experimental Schools by Region *24*
Table 22. Student Academic Scores in Experimental and Control Schools by Region *25*

Executive Summary

Overview

This report constitutes the second part of the overall evaluation of the Interactive Radio English Programs for Grade One in Ethiopia. It is a summative evaluation of the programs and its overall aim is to find out whether or not the Grade One Interactive Radio Instruction English Programs were effective in Ethiopia. The main purpose of the study is to measure students' learning gains as shown by an achievement test comprising English listening skills. A secondary objective is to determine how IRI programs address equity issues in Ethiopia.

The main approach to the study is a basic pre-test/teach/post-test design. As such its aim is to indicate how much the Interactive Radio English Programs in Grade One contribute to or cause changes in student learning as compared to non-Interactive Radio Instruction. Participants in this study constitute control and experimental groups and they were randomly assigned to their respective groups. The dependent variable, student achievement score, was measured at two occasions: pre-test and post-test. At the pre-test both control and experimental students were given a listening comprehension test, using a tape-recorder. After an interval of seven months (which also constituted a treatment period), the same test was given to the same students in their respective groups. The principal independent variable was traditional instruction versus interactive radio instruction. Traditional instruction was given in the control schools and there was no exposure of the students to IRI in these schools. In experimental schools, students learned with IRI and the time allotted to English has been fully covered by IRI. This included 5 fifteen-minute radio lessons every week. The samples were assigned on random basis to both groups and the sample size was 1466.

Main Findings in the Study

- Interactive Radio English Instruction Programs were more effective in improving educational quality. Students who used these programs gained 22% as compared to 9% in control schools that did not use these programs.

- In traditional or non-IRI classroom instruction, learning outcomes continue to favour urban students over rural students. In interactive radio instruction, the students in urban and rural classes began with a slight differential in favour of the urban classes but there were no statistically significant differences in the gains of urban and rural IRI students with each gaining nearly an equal amount. Thus, IRI English programs were equally effective in increasing academic gains both in urban and rural schools.

- Within urban areas, a comparison of students' academic achievement in control and experimental schools indicate that students who use English IRI Programs significantly performed better than those that do not. Moreover, when academic gains in urban control and urban experimental schools are compared, it was found out that students in urban experimental schools gained

than double of those in control schools. This means in urban areas IRI English Programs are highly effective.

- Within rural areas, a comparison of control and experimental schools indicates that academic gains are much higher for students in experimental schools, even though the experimental schools started with notably lower performance in comparison to control schools as shown by pre-tests. This implies that the use of interactive radio in rural areas can help to eliminate differentials between urban and rural schools.

- Students' academic achievement did not significantly differ among males and females in control schools. In experimental schools, female students performed better than male students. Interactive Radio English Programs were more effective than traditional instruction for students of both sexes, but on aggregate favoured females over males.

- When it comes to the relationship of students' academic achievement to self-contained and non-self contained approach in grade one, the results show that higher gains are related to non-self contained approach in control schools. In experimental schools, higher students' academic achievement is positively related to the self-contained approach. In both control and experimental schools there has been a tendency for the differences in academic performance to narrow, but this tendency is much higher in experimental schools. This means that IRI is not only more effective overall, but also reduces gaps that may persist in traditional instruction.

- When it comes to the relationship of academic gains and class size, in control schools students in large classes significantly achieved more than smaller classes. This is against the common assumption that relates lower student performance with large classes. In experimental schools, students in larger classes performed better than those in smaller classes at the pre-test, but with the introduction of IRI these gaps were eliminated and equality was achieved.

- With reference to the relationship between academic scores and regions, students gained significantly in all regions (except in Harari) from traditional instruction in control schools. However, student academic gains are more than double greater in experimental schools. This suggests that IRI should be introduced in regions where the innovation thus far has not been applied.

In summary, the results of the summative evaluation show that the IRI English lessons had a significant impact on academic achievement. These positive results occurred in all regions, in urban and rural schools and among male and female students. There was a slight advantage in terms of female students.

Conclusions

This study has demonstrated that *English for Ethiopia*- IRI has been more effective in improving student learning than has been the traditional instruction. This suggests that the innovation must continue and be adopted on a wider scale. This includes all regions and both urban and rural schools.

The fact that IRI has provided more opportunities to female students to perform well is a very positive tendency in Ethiopia. When it comes to teachers, students having male IRI teachers benefited more than those having female teachers. Although, this study has not controlled for the size of male and female teachers, the overall tendency suggests that training and other mechanisms may be followed to encourage female teachers to use radio as much as male teachers.

There is no question that IRI is going to continue under heterogeneous groups as far as age is concerned. The fact that it has managed to address these groups is a very strong feature. The strength of IRI to address the needs of self-contained classes has to be encouraged since this will be the future classroom structure in Ethiopia. IRI English Programs in Grade One demonstrated that it is possible to meet the interests of less experienced and experienced teachers. This creates a unique opportunity particularly for rural and inexperienced teachers to get a curricular material that they may not easily find or produce in their schools. The potentials in IRI English Programs to eliminate differences among urban and rural schools and among regions call for more attention to strengthen these programs in the context of Ethiopian primary education.

.

Part One

1.1. Introduction

This report constitutes the second part of the overall evaluation of the Interactive Radio English Programs for Grade One in Ethiopia. The main aim of the program is to improve the quality of English language instruction in Ethiopian primary schools. Interactive Radio Instruction (IRI) in Ethiopia is both a new approach and at a pilot stage. As such, it is being evaluated if it should be adopted on a wider scale. This program is supported by BESO/USAID and conducted by EMA. The first report, a formative evaluation, has been completed and submitted to BESO/USAID as well as to EMA. This study constitutes the summative evaluation of the programs in order to describe the effectiveness of Interactive Radio Instruction Grade One English Programs in Ethiopia. The main subject of the study is students' learning gains measured by an achievement test comprising listening skills. This study also presents how IRI programs address equity issues in Ethiopia as a secondary objective. The current report describes the design of the study, the testing experience, and the outcomes of the process.

1.2. Aims and Objectives of the Study

The overall aim of this study is to analyse the learning gains resulting from the exposure of Grade One students to Interactive English Radio Instruction Programs in Ethiopia. As a secondary objective this study also reflects upon the impact of IRI programs on equity issues. The objectives of this study are:

- To compare the academic scores and gains of students exposed to IRI English Programs in Grade One with students of the same grade having no exposure to these programs.
- To describe the academic scores and gains of students with and without IRI programs in terms of school location, students' sex, students' age, learning structure in Grade One, class size, teachers' experience and teachers' sex , and
- To draw some conclusions that may contribute to the adoption of the program on a wider scale in Ethiopia.

In general, the current study is an attempt to examine the contribution of Interactive Radio Instruction Programs to improve quality and equity in Ethiopian education particularly in English language at the initiation of primary schooling.

1.3. Significance

English is the most widespread language around the world, and as such it is the number one international language in Ethiopia. It is the language of instruction beginning from secondary school, or even earlier in some regions. It is taught as a subject from Grade one is one of those that determine whether or not a student qualifies for higher education. Providing quality learning in English is, therefore, vital to enable students to remain in formal schooling as well as in the world of work. IRI is one of the experimental approaches to meet these ends. Studying the outcomes and effectiveness of this experimental approach is important because it provides vital

information to decision makers as to whether or not the programs should continue, expand to a wider scale or be improved.

Equity is a major problem of Ethiopian education. The issues comprise several dimensions including gender, location, school facilities, teachers etc. Experiences from several countries suggest that IRI has an important impact in closing equity gaps in education. This study includes an approach to show the potential of IRI to narrow inequality in education while simultaneously improving quality.

Although Ethiopia has many years of experience in supporting primary education with radio, only limited research has been conducted as to the effectiveness of the programs. Thus, the scope of development of educational radio research and the areas investigated were very limited. From these perspectives, the current study enriches the possibilities for conducting research on radio programs and raises some important issues for future investigation.

1.4. Limitations of the Study

Although language skills can be assessed in a form of listening, speaking, reading and writing, this study is limited only to the assessment of listening comprehension. A variety of reasons including logistical, experiential and pedagogical have contributed to this limitation. The assessment of other skills is essential because they are included in the curriculum.

The fact that IRI is new in Ethiopia has contributed to some problems in data collection. While reliance on the South African experience has solved many design issues in test development, there remain limitations on school information. Obtaining more information will be important to carry out more in-depth analysis of data in the future.

There was a considerable attrition of participants from the study. Some of the reasons have to do with technical problems in post-test production on tape recorders while others are normal dropout effects.

The administration of the tests using a tape recorder limited the testers from solving some problems in terms of instructions and making corrections during the tests. In fact, the use of the mother tongue for instructions in some places might have produced some misconceptions.

Conducting the tests in large classes with inadequate furniture and facilities may mean that copying answers from one another could not be adequately controlled.

The use of pictures as test instruments might need some level of experience and maturity from children. Difficulties in interpreting or reading pictures might interfere with assessment of students' listening skills. Pilot tests must be conducted to find out and solve potential problems in the future.

Part Two

Research Methodology

2.1. Design

This study is a basic pre-test/ teach/ post-test design, which aims at indicating how much the Interactive Radio English Programs in Grade One contribute to changes in student learning as compared to non-radio instruction. Participants in this study constitute control and experimental groups and they were randomly assigned to their respective groups. The Educational Media Agency (EMA) requested Regional Educational Bureaux (REBs) to identify one zone in which there are urban and rural schools and where a Zonal Media Officer or another professional education worker is available to assist. Each REB was asked to provide 6 rural and 6 urban schools from the selected zone. The 12 schools were also required to be within 50 km of the Zonal Office to be accessible to Zonal Media Officer by public transportation. From a pool of 98 schools EMA identified 16 urban and 16 rural schools at random. Finally, it was from this list that EMA randomly assigned schools to four groups: urban control (N=8), urban experimental (N=8), rural control (N=8) and rural experimental schools (N=8).

The dependent variable, student achievement score, was measured at two occasions: pre-test and post test. At pre-test both control and experimental students were given a listening comprehension test, which was developed and recorded on a tape-recorder. After an interval of seven months, which also constituted the treatment period, the same tests were again given to the same groups. The two principal independent variables were traditional instruction and interactive radio instruction. Traditional instruction was used in the control schools and there was no exposure of the students to IRI in these schools. In experimental schools students learned using IRI and the time allotted to English was 5 periods of radio lessons every week. One of the lessons at the end of each week constituted a "review lesson". The main purpose of the review lesson was to provide the summary lessons conducted in the week. Each IRI lesson was structured into segments and each program lasted for 15 minutes. These segments comprise standard opening, structured practice, first teacher-led activity, song, another structured practice, physical activity and the second teacher led-activity (instructions only). There were other crucial procedures that were performed to ensure that the independent variables were the ones that really created differences. The table below shows some of these activities in both groups of schools.

Table 1. Additional Design Activities in Control and Experimental Schools

Control Schools	Experimental Schools
• Text books were supplied. • Pencils were provided. • Few visits were carried out. • ICDR teachers' guides were printed, but not distributed because they were already available.	• Text books were supplied. • Pencils were provided. • 2-day teachers' radio utilization training was conducted. • Radio, radio cassettes, teachers' guides both radio and ICDR were provided. • Few visits were carried out.

Tests that last for only 2 hours were designed so that students may not be bored or tired. It was believed that an interval of seven months was adequate in order to eliminate the effects of the pre-tests from being carried over to the post-tests. Moreover, since students were from schools randomly assigned to experimental and control groups it was believed that any effects from testing were controlled through equalization among the groups. The study included only Grade One students and there were no non-equivalent groups. Students who took pre-tests but did not appear for the post-tests were eliminated from analysis. Academic gain was the only dependent variable used in this study. However, in order to see what radio contributes to closing equity gaps, other independent variables were built into the design. Dane (1990) argues that such variables differentiate participants, but cannot be manipulated and are not subject to random assignment. Seen from these perspectives, the additional variables in this study represent participant characteristics as well as the characteristics of their working conditions that cannot be manipulated.

2.2. Instruments of Data Collection

The main instrument used to collect data in this study was a listening comprehension test. Arnott, Manfield and Mentis (1993) indicate that language skills can be assessed at four levels: listening, speaking, reading and writing. In the current study, the choice of listening comprehension has been influenced by a variety of factors. Primarily, there was an appropriately validated test instrument obtained from South Africa's "English in Action" Radio Learning Programs. Secondly, the age and educational experience of students was believed to be appropriate for the type of test currently adopted. These children were too young for reading or writing tests. Thirdly, since this was for the first time that such tests have been conducted it was believed that testers might also lack experience and expertise. The choice of verbal comprehension has also to do with the pedagogical principle which notes that verbal comprehension is more important in the development of a language in a child than the ability to verbalise thought through sound (Lennberg, 1970; Thurston and Thurston, 1963).

The testing situation was conducted on group basis (in a class) for two hours in two sessions with an interval of ten minutes break. In order to control other variables that may influence the experiment other than the independent variable the test was pre-recorded and administered via a tape recorder. Testing took place in two stages: Pre-test and post-test.

Pre-Test: The tests developed by South African "English in Action" Radio Learning Program provided a solid background from which the Ethiopian version was developed. In fact this test has also been a major instrument in Kenyan Radio Language Arts Project (Arnott, Manfeild and Mantis, 1993). With this background, EMA English language experts and external consultants reworked the test as follows:

- Adjusted all the test items according to the curriculum,
- Changed items that were culturally inappropriate,
- Redesigned pictures that lacked clarity,
- Refined the instruction, and
- Analysed the final version of the test before administration. The test was administered at the beginning of November 2000.

Post-test: There was an implicit understanding to evolve the pre-test to a more complete post-test after the result of a statistical item analysis as well as feedback from the testing experience. Astonishingly only two items were found to have weak discriminating powers and these were believed to have little impact on the process. Consequently, the post-tests were conducted in June, 2001 with no alteration of the pre-test.

In both phases, tape-recorded tests were administered to the testees by zonal media co-ordinators. These coordinators had previously taken training as to how to conduct these tests in schools.

The other instrument for data collection was a questionnaire, which was completed by zonal media supervisors and the school directors. This questionnaire required information on participant characteristics teacher characteristics, school characteristics and in which the project was conducted. This information was included in order to address equity issues.

2.3. Selection of Samples

A stratified cluster sampling procedure with three strata was used. These were region, school type and school Group (status). Each stratum was defined into the following clusters.

- Regions: Addis Ababa, Oromia, Amhara, Tigray, Harari, Gambella, Dire Dawa and SNNP.
- School Location: Urban and Rural Schools.
- School Type (Group): Experimental and Control

The schools selected were all government primary schools. Region was used as a stratifying variable because there are differences in the delivery of primary education among them. School location was considered for inclusion in order to reflect the demographic characteristics of the population. Moreover, there was an assumption that urban and rural schools differ in facilities, teachers' experience, and exposure to English, shift system etc. School group was another stratification criteria in order to find out or isolate the impact of the new approach as compared to the standard instructional situation. For the current phase, 8 regions were selected based on the desire to work with a widely representative group within the practical limits or capacity of the project in its first year. Within experimental and control schools, grade one students were randomly selected on section basis. Table 1 below summarizes the research settings and the participants of the study.

Table 2. Research Settings and the Participants of Study

Duration of Tests	Region	Status (group) of Schools		Participants of the Study		Other Participants
		Control	Experimental	Control	Experimental	
Pre-Test	Tigray	2	2	143	149	Total number of schools which provided data on participant characteristics was 26
	Amhara	2	2	124	97	
	Oromia	2	2	143	149	
	SNNPR	2	2	143	178	
	Harariii	2	2	106	77	
	Dire Dawa	2	2	129	106	
	Gambella	2	2	127	147	
	Addis Ababa	2	2	166	136	
	Total	16	16	1081	1039	
Post-Test	Tigray	2	2	117	137	
	Amhara	2	2	65	102	
	Oromia	2	2	116	114	
	SNNPR	2	2	120	130	
	Harariii	1	1	43	56	
	Gambella	2	2	87	102	
	Addis Ababa	2	2	131	146	
	Total	13	13	679	787	

With the exception of Harari, the total number of experimental schools and that of control schools was 2 each from each of the regions. This constitutes a total of 32 schools with 16 control and 16 experimental schools for the whole project. The total number of students who originally took the pre-test in these schools was 2,120. Of this total, the number of pupils who took the post-test were 1,466. The reasons for the reductions were the irregularities in the administration of examination and the dropout of pupils from the examinations. Irregularities in examinations during the post-test happened in Dire Dawa and Harari. It was reported from Dire Dawa that poorly recorded cassettes were the sources for not conducting post-tests as required while in Harari the tests were conducted by untrained Zonal Media Supervisors. For Dire Dawa, it was decided to totally disregard results from schools (N=235) while for Harari results from only one control and one experimental school was accepted (N=99). The total population for which complete results were obtained, thus, amounts to 679 in control schools and 787 in experimental schools. The total number of sample schools for which pre-tests and post-tests were obtained amounted to 26.

2.4. Data Management and Analysis

The administration of the test at different times required the preparation of different data files for both the pre-test and post-test. In both cases after items were coded data were entered into the SPSS program. The questionnaire, which provided data on the characteristics of participants as well as their working environment, was entered to the same program after the separate files created for pre-tests and posts-tests were merged. Careful cleaning of data followed the entry since it was discovered that there were some mismatches. In some regions students who did not sit for the pre-test were seated for post-tests while in others those who were present on the pre-tests have dropped out.

Data analysis for the tests as well as for the questionnaires was performed using SPSS with means, t-tests and ANOVA as the main tools for summarizing data and testing the null-hypothesis. Tables were used for demonstrating results and providing a comparative analysis of student mean scores and academic gains between different groups of schools included in the study.

Part Three

Findings and Analysis of Data

3.1. Effectiveness of IRI in Improving Educational Quality

3.1.1 Student Academic Gains in Control Schools without IRI Programs

Educational quality in the current study was operationalized to mean student academic gains as measured by the listening comprehension test. As part of the experimental design, students in the control schools were exposed to both pre-tests and post-tests. Results indicate that there are indeed improvements of achievements by students due to the educational treatments other than radio. The results are summarized in Table 3 below.

Table 3. Students' Academic Gains in Control Schools

Measurement (N=679)	Mean Scores	Standard Deviation	Significance
Post-test	66.0825	13.8539	
Pre-test	56.3292	19.4502	
Academic Gain (Post test- Pre test)	9.7533	18.6253	.000

The above table indicates that the mean of students' academic scores for the post-test was 66.08 percent while it was 56.32 percent for the pre-test. The mean academic gain of students was 9.75 percent. A 2-tailed t-test indicates for equality of means that the difference between post-test and pre-test scores was highly significant at 0.05 level (df=678, t=12.993). This implies that the traditional intervention produces significant improvements in students' learning.

3.1.2. Students' Academic Gains in Experimental Schools with IRI Programs

Students in the experimental schools who were exposed to IRI programs were given the same pre-test and post-test at the same time with control schools. Results indicate the academic improvement is more than twice the improvement achieved in control schools. Table 4 below shows the results.

Table 4. Students Academic Gains with IRI Programs in Experimental Schools

Measurement (N=787)	Mean Scores	Standard Deviation	Significance
Post-test	76.5851	13.5936	
Pre-test	54.1773	17.8909	
Academic Gain (Post test-Pre-test)	22.4079	19.3370	.000

Table 4 above indicates that the mean academic achievement of students in the post-test was 76.58 percent while in the pre-test it was 54.17 percent in experimental schools. The mean gain in student achievement scores was 22.40 percent. A 2-tailed t-test indicates that the difference between the pre-test and post-test scores was highly significant at 0.05 level (df=786, t=32.509). The implication of this finding is that interactive radio instruction indeed had a major impact on students learning of English.

3.1.3. Students' Academic Gains in Experimental and Control Schools

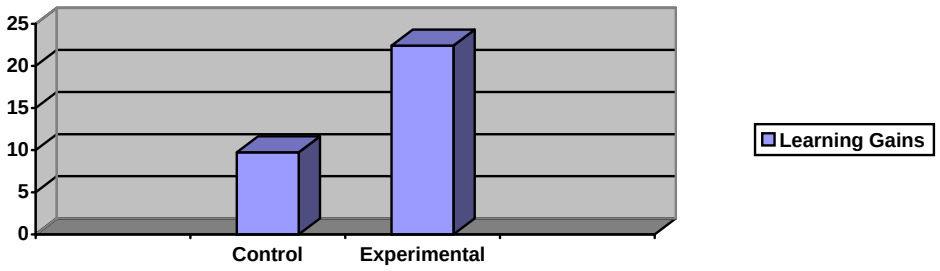
From the above two tables it is clear that students in both the control and experimental groups were exposed to pre-tests and post-tests. In control schools, the students learned using traditional methods and curricula while learning was supplemented by IRI in experimented schools. Results show that students who learned with IRI scored about 13 percent more than those who learned without Interactive Radio programs.

Table 5. Students' Academic Gains in Control and Experimental Schools

Status (group) of the School	Measurement		Academic Gain (Post test-Pre test)	Sig.
	Post- test	Pre-test		
Control (N=679	66.0825	56.3292	9.7533	.000
Experimental (N=787)	76.5851	54.1773	22.4079	.000
Mean Difference	10.5027	2.1519	12.6546	

Table 5 above also indicates that when the mean scores of students in percentages at the post-test is considered between experimental and control groups, there is a gain of about 10 points in favour of experimental schools. This is in contrast to a gain of 2 point advantage of control schools over experimental schools at the pre-tests. A 2-tailed significance test for the equality of means indicates that there was a highly significant statistical difference between control and experimental schools during the post-test (Sig. 2-tailed=.000, df=1464) and less statistically significant difference during the pre-test (Sig.2-tailed=.028, df=1464). This finding indicates that students in control schools had better standards of English knowledge as measured by listening skills (an average of 56 % against 54%) just sometime after they joined Grade One or at the pre-test. The following graph shows these results.

Graph 1. Learning Gains in Control and Experimental Schools



The introduction of IRI in experimental schools has brought a significant difference in learning gains over the control schools. That is, experimental schools gained a mean of 22 % against 9% in control schools after seven months of teaching English. As it

will be indicated later in this report, an advantage of about 13% produces about 0.65 effect size. This finding is consistent with other findings where higher learning gains for students using IRI were repeatedly reported in comparison to control schools. For instance, Bosch (1997) presented data from six countries in which IRI schools consistently scored better than the control schools. In that report, it has been indicated that the effect sizes of IRI programs ranged between 0.24 and 0.94.

3.2. Effectiveness of IRI in Closing Equity Gaps

3.2.1. Students' Academic Achievements in Control Schools by Location

The previous finding has shown that traditional instruction has an impact on students learning. However, this difference varies depending on the location (e.g urban/rural) of the school. In the current study, academic gains from traditional instruction have been studied to verify to what extent traditional instruction reduces the gaps (differences in students' academic performance between pre-tests and post-tests) between rural and urban students. Table 6 summarizes the results.

Table 6. Students' Academic Achievements in Control Schools by Location

Measurement	School Location	Number of Participants	Mean Scores	Mean Difference	Std. Deviation	Sig.(2-tailed test)
Post-test	Urban	405	67.4691	3.4363	12.6885	0.001
	Rural	274	64.0328		15.2099	
Pre-test	Urban	405	57.0802	1.8613	19.2690	0.221
	Rural	274	55.2190		19.6977	

The above table indicates that the mean scores in the post-test for urban control schools was 67.46 percent while this amounts to 64.03 for rural schools. In other words, urban schools scored more than rural schools by a mean percentage of 3.43 points and this difference is statistically significant at 0.05 level. Students of urban and rural background had narrower differences in the pre-test. Those from urban areas had advantages by mean difference of 1.8613 percentage points and that is not statistically significant. Two trends are clear from these results. First, urban students scored better than rural students in the tests both before and after lessons. Second, although not significant, the gap between the two groups is not decreasing, and traditional educational intervention and urban schools continue to gain more points or marks.

3.2.2. Students' Academic Achievements in Experimental Schools by Location

It has already been indicated that learning with IRI has significant impact or effect upon students. However, there was also a hypothesis that learning with IRI closes the differences in academic performance between learners from urban and rural areas. This hypothesis finds its root in previous studies which show that rural students gain much more total scores than their urban counter parts (see the formative evaluation report). Another root in the above hypothesis may also be the potential of IRI to

provide equal opportunity for learners irrespective of their locations. What this trend looks like is summarized in the table below.

Table 7. Students' Academic Gains in Experimental (with IRI) Schools by Location

Measurement	School Location	Number of Participants	Mean Scores	Mean Difference (Urban-Rural)	Std. Deviation	Sig.(2-tailed test)
Post-test	Urban	404	80.7426	8.5428	11.1389	.000
	Rural	426	72.1997		14.5518	
Pre-test	Urban	404	57.3700	6.5607	18.8099	.000
	Rural	426	50.8094		16.2230	

Table 7 above shows that the post-test mean score of urban school students was 80.74 while it was 72.19 for the rural ones. The mean difference for the post-test was 8.54 in favour of the urban schools, which was also statistically significant. Similarly, urban school students scored better than rural school students in the pre-test. As shown in the table, the mean score of urban students at the pre-test was 57.37 percent while it was 50.56 percent for rural students on the same test. The mean difference between the two groups in pre-test was 6.56 and this was also statistically significant. When the difference between urban experimental schools in post and pre-test is considered the gain amounts to 23.37 in the post-tests, and for rural schools the gain was 21.39.

As it is clear from the above findings, interactive radio instruction has not reduced the differences in students' academic performance between urban and rural schools, but contributed to increased academic gains in both environments. At the same time neither the academic gains in favour of urban schools nor the mean differences in urban-rural experimental schools between the two tests was significant. In the Ethiopian context, therefore, radio has to be taken as an instrument for raising quality in both environments.

3.2.3. Students' Academic Gains in Urban Control and Experimental Schools

Although urban schools share similar characteristics in many ways, the reactions of their students to innovations may vary. From this perspective, the impact of IRI in relation to experimental schools and control schools in urban areas is summarized in the following table.

Table 8. Students Academic Gains in Urban Experimental and Control Schools

Measurement	Status of the School	Number of Participants	Mean Scores	Mean Difference	Std. Deviation	Sig.
Post-test	Control	405	67.4691	13.2734	12.6885	.000
	Experimental	404	80.7426		11.1389	
Pre-test	Control	405	57.0802	0.829	19.2690	.829
	Experimental	404	57.3700		18.8099	

The above table indicates that the mean academic scores of urban control schools in the post-test was 67.46 percent while it was 80.74 for urban experimental schools.

This shows a mean difference of 13.27 in favour of the experimental schools. A 2-tailed t-test for equality of means indicates that the mean gains for control and experimental schools were statistically significantly different from each other. In the pre-test, students in the urban control group achieved a mean of 57.08 percent while students in experimental schools got 57.37 percentage points. This result indicates that prior to radio intervention students had statistically insignificant differences in their English knowledge. Therefore the increments observed in post-tests could be related to educational interventions, but the attainment is much higher for experimental schools (with IRI schools) than schools without radio. Also important in this finding is that with radio instruction students' achievement gaps were rather widening in urban centres. This should not be astonishing since adding something different in schools having the same standard has to lead to increased results. Thus, radio still has a positive role in maximizing performance where it is introduced.

3.2.4. Students' Academic Gains in Rural Experimental and Control Schools

Whether or not the use of radio in rural schools brings about change in students' performance and if that change is significant is an important issue. It is also equally important to understand if the differences in students' performance at pre-test stage were deceasing or increasing at the post-test stage.

Table 9. Students' Academic Gains in Rural Experimental and Control Schools

Measurement	Status of the School	Number of Participants	Mean Scores	Mean Diff. (Exp-Cont.)	Std. Dev.	Sig. (2-tailed test)
Post-test	Control	274	64.2100	8.1669	15.2099	0.000
	Experimental	383	72.1997		14.5518	
Pre-test	Control	274	55.2190	4.4096	19.6977	0.002
	Experimental	383	50.8094		16.2230	

Table 9 above shows that the mean academic scores for rural control schools was 64.21 percent while it was 72.19 percent for rural experimental schools at the post-test stage. The mean difference in favour of rural experimental schools is 8.16 percent in post-test and this difference is statistically significant. As far as results from the pre-tests are concerned, the above table shows that rural control schools have shown an academic mean score of 55.21 percent while the mean for rural experimental schools was 50.80 percent. The mean difference between the two groups of schools was 4.40 in the pre-test. A 2-tailed t-test for equality of means indicated that the difference between academic mean scores in rural control and experimental schools is significant. In other words, the use of radio helps not only to eliminate previous gaps in students' performance, but also maximizes students learning to the highest extent that it surpasses those who do not use it. This observation particularly becomes obvious when the academic gains of experimental schools (21.39) is compared to those of control schools (8.88 points between the post-tests and pre-tests).

3.2.5. Students' Academic Gains in Experimental and Control Schools by Location

What do the differences in academic gains look like when experimental and control groups are compared by location?

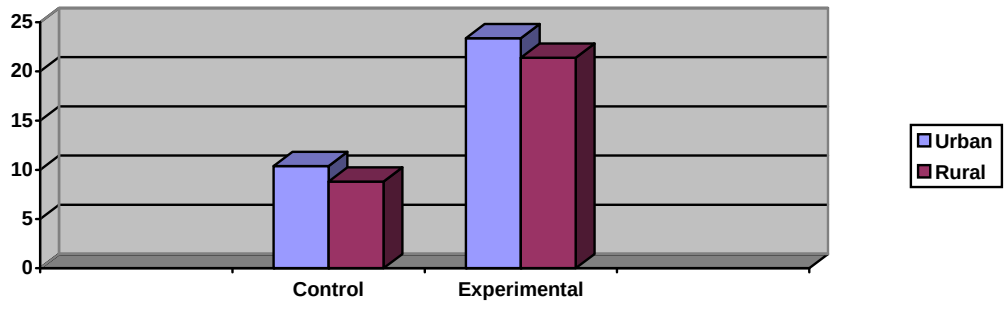
Table 10. Student Academic Gains in Experimental and Control Schools by Location

Measurement	School Type	Mean Academic Scores		Mean Difference	Sig.
		Urban	Rural		
Post-test	Control	67.4691	64.0328	3.4363	.001
	Experimental	80.7426	72.1997	8.5428	.000
Pre-test	Control	57.0802	55.2190	1.8613	.221
	Experimental	57.3700	50.8094	6.5607	.000
Mean Gain (Post test-Pre test)	Control	10.3889	8.8138	1.5751	.304
	Experimental	23.3726	21.3903	1.9823	.322

In Urban Control schools N= 405. In Urban Experimental Schools N=404. In Rural Control Schools N=274. In Rural Experimental Schools N=383. *effect size = 0.65

From the above table it becomes clear that the mean difference in post-tests between urban control and rural control schools was 3.43 implying better achievements for urban schools than the rural ones after the academic year with traditional instruction was completed. Secondly, when results in post-tests for urban and rural experimental schools were compared, urban experimental schools achieved better by a mean difference of 8.54 %. Third, when the mean gains between urban control and urban experimental schools were compared, it was to found out that the urban experimental school gains were more than twice (23.37%) the gains of urban control schools (10.38). This means that radio instruction has been more effective than traditional instruction in urban centres, and the difference in student performance has increased in post-test than in the pre-tests. Fourth, when mean comparison between rural control and rural experimental schools was made, it became clear that rural experimental schools performed better than rural control schools by a wide margin. This indicates once again that radio has not only been more effective than traditional instruction, but also could narrow the gaps between rural experimental and rural control schools. Graph 2 below shows learning gains by location of schools.

Graph 2. Learning Gains in Urban and Rural Schools



Overall, more learning achievements both in traditional and radio instruction have been observed when results in pre-tests were compared with results in the post-test. Nevertheless, these gains have been observed more in experimental schools than

control schools. In fact Table 10 above shows that the effect size of IRI is 0.65. In comparison to results reported from other countries by researchers like Bosch (1997), this result has shown a fair success. Within that radio seems to have benefited urban experimental schools than rural experimental schools, although that was statistically insignificant. This implies that radio could produce improvements without creating imbalances between urban and rural schools. In Control Schools statistically significant imbalances were not observed too, but quality improvement in relation to IRI was very low.

3.2.6. Students' Academic Achievements by Sex in Control Schools

Whether females or males gain more from schooling is a matter of interest as far as closing equity gaps are concerned. This is so because several studies indicate disparities among boys and girls in academic achievement. The following table indicates the results from control schools.

Table 11. Students' Academic Achievements by Sex in Control Schools

Measurement	Sex of the student	Number of Participants	Mean Scores	Mean Difference	Std. Deviation	Sig. (2-Tailed test)
Post-test	Male	373	66.1461	0.1707	14.7199	0.873
	Female	305	65.9754		12.7528	
Pre-test	Male	373	55.4223	1.9466	19.9337	0.195
	Female	305	57.3689		18.8130	

Table 11 above indicates that grade one male students had a mean score of 66.14 percent in the post-test whereas females had 65.97 percent. A t-test for equality of means indicates that this difference has not been statistically significant. In the pre-test the mean score of males is 55.42 and 57.36 for females. The t-test indicates that the observed differences were not statistically significant. This finding, in fact, is not consistent with various statistical reports that show the advantages of boys over girls in academic achievement in Ethiopia. In fact if absolute values in mean gains are considered female students were the ones who improved or learned more than male students since they gained 10.72% against 8.60% for males.

3.2.7. Students' Academic Achievements by Gender in Experimental Schools

The potential of IRI to narrow gaps between the performance of male and female students may not be considered so significant from the above table. This is because the performance gaps were not statistically impressive or were not simply apparent. However, findings in the table below show that female students who were taught using IRI have significantly achieved more than males. In the post-tests males had an average of 74.52 percentage points while females had 78.81. A t-test indicates that this difference is statistically significant. What is interesting is the fact that both males and females had nearly the same standard in the pre-tests. If the gains of male students between the post and pre-tests are considered it becomes 20.99 mean points while that of females amounts to 23.94 mean points and this is statistically significant.

Table 12. Students' Academic Achievements in Experimental Schools by Sex

Measurement	Sex of the student	Number of Participants	Mean Scores	Mean Difference	Std. Deviation	Sig. (2-tailed test)
Post-test	Male	432	74.5232	4.2929	14.4915	0.000
	Female	398	78.8161		12.1811	
Pre-test	Male	432	53.5330	1.3413	17.2572	0.294
	Female	398	54.8743		18.5499	

The significance of the above finding is the fact that IRI is useful to eliminate academic disadvantages of girls in schools. In most cases, it is well known that female students either lack access to education or drop out of the process before completion. It appears that IRI has the capacity to enhance girls' education given that it is strategically planned and practiced.

3.2.8. Students' Academic Gains in Experimental and Control Schools by Sex

Females have gained significantly more than males in the post-test in experimental schools. However, this difference does not show up in the traditional or control schools. When the mean differences between males and females in control schools in the post-test is compared, the difference is statistically insignificant. However, in experimental schools females have scored more than males by 4.29 %, which is statistically significant. Table 13 below summarizes the results.

Table 13. Student Academic Gains in Experimental and Control Schools by Sex

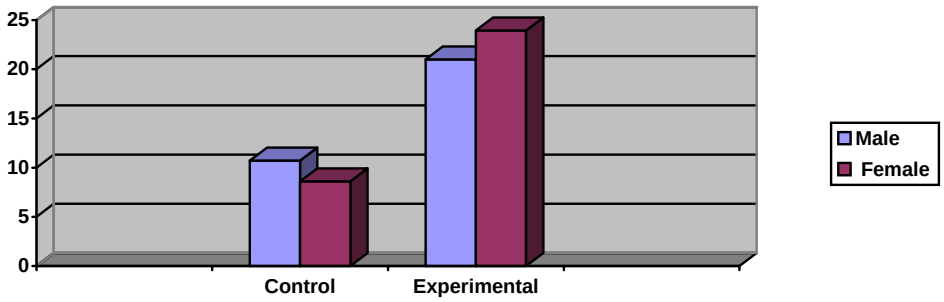
Measurement	School Type	Mean Academic Scores		Mean Difference	Sig.
		Male	Female		
Post-test	Control	66.1461	65.9754	0.1707	.873
	Experimental	74.5232	78.8111	4.2929	.000
Pre-test	Control	55.4223	57.3689	1.9466	.195
	Experimental	53.5330	54.8743	1.3413	.294
Mean gains	Control	10.7239	8.6066	2.1173	0.161
	Experimental	20.9902	23.9418	2.9516	0.032

Males in Control schools N= 373. Females in control schools N=305. Males in experimental schools N=409. Females in experimental schools N=378.

The above table suggests that in pre-tests, both male and female students had similar academic standards. Table 13 also indicates that in control schools the academic gain of male students was 10.72 % points when results of the post-test are compared with that of pre-tests. In experimental schools, the net academic gain of male students was 20.98 %. This is statistically significant. The academic gains of female students between the pre-test and post-test in control schools amounted to 8.60% but was 23.94% points in experimental schools. The gains were much greater in experimental schools than in control schools. Here, it was female students who achieved more than the male students as they started slightly behind males but finished significantly better than them at the post- test. This suggests once more that IRI is more effective than

traditional instruction for students of both sexes, but on aggregate favoured females over males. The following graph shows learning gains by gender of the students.

Graph 3 Learning Gains by Sex of Students



3.2.9. Students' Academic Achievements in Control Schools by the Organization of Learning in Grade One

The self-contained approach is one of the new instructional technologies used to implement the new Ethiopian Policy of Education and Training. However, not all schools have adopted the approach and so there are many schools with the previous linear or non-self contained approach.

Table 14. Students' Means Academic Gains in Control Schools by Organization of Learning in Grade One

Measurement	Status of Grade One	Number of Participants	Mean Scores	Mean Difference	Std. Deviation	Sig. (2-tailed test)
Post-test	Self-Contained	356	64.3680		13.2296	.000
	Linear	296	69.0963	4.7283	13.2562	
Pre-test	Self-Contained	356	60.1334		13.6493	.000
	Linear	296	54.2736	5.8598	23.2630	

When the mean gains were considered between the two approaches in control schools at grade one, it appeared that more learning takes place in the linear (non-self contained) classrooms.

Table 17 shows that at pre-test, the mean academic gains of self-contained students (60.13%) were significantly higher than the mean scores of linear students (54.27%). However, this has changed significantly in favour of linear students at the post-test where by self-contained students scored an average of 64.36 % and the linear students scored 69.09 %. One of the explanations for such changes can be linked to the novelty of the new approach to teachers and to the insufficient knowledge of the new approach in English classes.

3.2.10. Students' Academic Achievements in Experimental Schools by the Organization of Learning in Grade One

In experimental schools, students in both self-contained and linear classes have certainly gained from IRI as there are huge gains in post-tests over pre-tests. At the same time, significant differences in academic gains persisted in favour of the self-contained classes both at the post-test and pre-test.

Table 15. Students' Academic Achievements in Experimental Schools by Organization of Learning in Grade One

Measurement	Status of Grade One	Number of Participants	Mean Scores	Mean Difference	Std. Deviation	Sig. (2-tailed test)
Post-test	Self-contained	254	80.2165	5.3619	10.7547	.000
	Linear	533	74.8546		14.4467	
Pre-test	Self-contained	254	67.4747	19.7007	15.6752	.000
	Linear	533	47.8189		15.1587	

The above table also indicates that students' gains are significantly better in linear classes (27.03%) than in the self-contained classes (12.74%). Because of this, the mean difference between the two groups at the pre-tests (19.70%) has gone down to (5.36%) in the post-tests. In general, it is possible to conclude that in schools that used IRI, learning gaps have narrowed down between self-contained and linear approaches, although it has not been totally closed.

3.2.11. Students' Academic Gains in Experimental and Control Schools by the Organization of Learning in Grade One

When mean academic gains in experimental and control schools by the organization of learning in Grade One were compared, the following results were found. First, in post-test self-contained experimental and linear experimental students gained more results than their peers in control schools. This means that IRI has been more effective than traditional instruction.

Second, traditional instruction has been much more effective with students in linear control schools (69.09%) than in self-contained schools as indicated in the post-test (64.36%). Third, in pre-tests self-contained students in both types of schools performed better than linear students in both types of schools, but experimental self-contained students were the best performers of all groups.

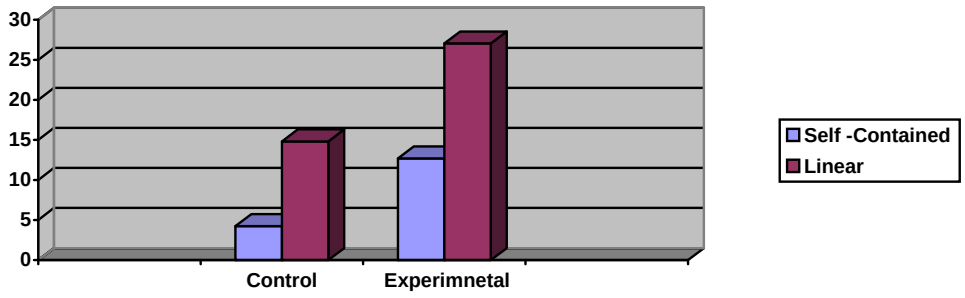
Table 16. Students' Academic Gains in Experimental and Control Schools by Organization of Learning in Grade One

Measurement	School Type	Mean Academic Scores		Mean Difference	Sig.(2-tailed)
		Self-Contained	Linear		
Post-test	Control	64.3680	69.0963	4.7283	.000
	Experimental	80.2165	74.2165	5.3619	.000
Pre-test	Control	60.1334	54.2736	5.8598	.000
	Experimental	67.5197	47.8189	19.7007	.000
Mean Gains	Control	4.2346	14.8226	10.5881	.000
	Experimental	12.6968	27.0356	14.3388	.000

In control self-contained N=356. In control linear N=296. In Experimental self-contained N=254. In Experimental Linear N= 533.

Fourth, academic gains were higher in experimental schools than in control schools, but the highest gains were in non-self contained experimental schools. Fifth, gaps were narrowed down in experimental schools more than they did in control schools when pre-tests and post-tests are compared. This shows that IRI plays more of a role in closing gaps between schools having different approaches to learning than traditional schooling. Graph 5 below demonstrates learning gains by organization of learning.

Graph 4. Learning Gains by Organization of Learning



3.2.12. Students' Academic Gains in Control Schools by Class Size

There is a common assumption that students do better in smaller classes than in larger classes, because there is more individualized attention, behavioural control, enriched curriculum, cohesion and sense of self-control by students. Stockard and Mayberry (1992) indicate that several large-scale reviews of literature in this area have noted very inconsistent results. The studies sometimes show that students have higher achievements in smaller classes, sometimes in larger classes, but most often the results are insignificant. However, when classrooms are smaller than about 15 students it is possible to consistently see an increase in achievement. In Ethiopia, the target classroom size by the Ministry of Education is 55 students. In the current study, the available data showed that the size of classes ranged between 45 and 155 students. Of these, the total proportion of classrooms having between 45-60 students was only 3%. Since this proportion has been insignificant, classroom size was defined into small and large categories. Small classes were defined as those that cater for students

from 45 to 75 in size while large classes are those with 76 to 155 students. Table 20 below summarizes the results.

Table 17. Students' Academic Gains in Control Schools by Class Size

Measurement	Category of Class Size	Number of Participants	Mean Scores	Mean Difference	Std. Deviation	Sig. (2-tailed test)
Post-test	Small Large	343 309	63.9504 69.3608	5.4101	12.9238 13.4484	.000
Pre-test	Small Large	343 309	57.1356 57.8479	0.7123	15.3091 22.1487	.630

Table 20 shows that the mean academic post-test score of students in small size classrooms was 63.95% while that of those in large size classrooms was 69.36%. The mean difference was 5.41% and this difference was statistically significant. In pre-test these group of students had similar scores. One possible explanation for this finding may be that most of the students in control schools are urban students. These students have better opportunities in covering the curricula, getting special services as well as teacher quality.

3.2.13. Students' Academic Gains in Experimental Schools by Class Size

Students' academic scores in pre-test clearly indicated that large classrooms related to higher gains than small classes. This result may be related to various particularities of schools that may need further investigation.

Table 18. Student Academic Gains in Experimental Schools by Class size

Measurement	Category of Class Size	Number of Participants	Mean Gains	Mean Difference	Std. Deviation	Sig. (2-tailed test)
Post-test	Small Large	200 587	76.3250 76.6738	2.6923	11.5368 14.2346	.754
Pre-test	Small Large	243 587	51.2346 55.1789	1.1151	18.0678 17.7346	.007

However, this trend changed with IRI since at post-test both groups scored similar results with small classes gaining more. Table 21 above summarizes these results. The fact that small classes have been favourable for working with IRI has also been found in the formative evaluation of the current project. In many occasions teachers, directors and zonal media supervisors have noted how difficult it was to use IRI in large classes.

3.2.14. Students' Academic Gains in Experimental and Control Schools by Class Size

When students academic gains between control and experimental schools by class size was considered, it was found out that the mean difference in academic achievement in control schools increases from 0.71 in pre-test to 5.41 in the post test

in favour of large classes. In Experimental schools the mean difference in academic gains has decreased from 3.9414 to an insignificant 0.3488 in favour of large classes.

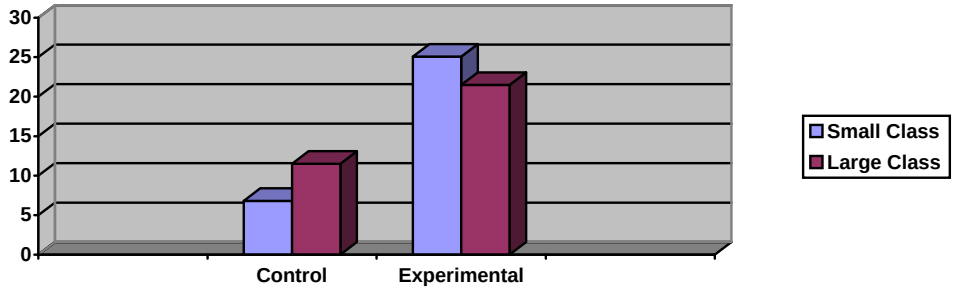
Table 19. Students' Academic Gains in Experimental and Control Schools by Class Size

Measurement	School Type	Mean Academic Scores		Mean Difference	Sig.(2-tailed)
		Small Class	Large Class		
Post-test	Control	63.9504	69.3608	5.4104	.000
	Experimental	76.3250	76.6738	0.3488	.754
Pre-test	Control	57.1356	57.8479	0.7123	.630
	Experimental	51.2375	55.1789	3.9414	.007
Mean Gains	Control	6.8149	11.5129	4.6981	.002
	Experimental	25.0875	21.4949	3.5926	.023

Students in control small size classes N=343. Students in control large classes N=309. Students in small experimental classes N=200.Students in large experimental classes N= 587.

This shows that radio instruction has been effective in reducing differences between differing class sizes while traditional instruction continued maintaining differences. On the other hand, differences in students' academic gains were higher in experimental schools than in controls schools, but mean differences in student performance have been less for experimental schools than control schools. Again this shows that IRI has been closing the gaps between students learning in large classes and small classes. The following graph shows learning gains by class size.

Graph 5. Learning Gains by Class Size



3.2.15 Students' Academic Achievements in Control Schools by Region

Students' academic scores in control schools at regional level indicates that in all regions post-tests were better than pre-tests with the exception of Harari. For Harari, perhaps issues other than academic conditions might have influenced the outcomes. This assumption is normal since students in all regions have shown improvements except in this region. In fact the highest academic gains from the regions was observed in Gambella with 31.45 mean percentage points which in most cases is classified with Harari as an emerging region. Overall the total mean academic score in the pre-tests for all regions was 56.32 while this was 66.08 in the post-tests. These results are summarized in the table below.

Table 20. Students' Academic Gains in Control Schools by Region

Measurement	Region	No.of Participants	Mean scores	Std. Error	Std. Deviation	Sig. (ANOVA
Post-test	Tigray	117	58.867	1.1538	12.4806	.000
	Amhara	65	63.538	1.4283	11.5150	
	Oromia	116	76.422	0.8328	8.9699	
	SNNPR	120	65.437	1.2499	13.6918	
	Harari	43	63.081	3.0247	19.8344	
	Gambella	87	60.201	1.6320	15.2223	
	Addis Ababa	131	70.114	.8064	9.2293	
	Total	**679**	**66.082**	**0.5317**	**13.8539**	
Pre-test	Tigray	117	51.559	1.1211	12.4806	.000
	Amhara	65	61.384	1.5319	12.3509	
	Oromia	116	60.883	1.9537	21.0415	
	SNNPR	120	61.354	1.6342	17.9017	
	Harariii	43	67.209	2.3710	15.5476	
	Gambella	87	27.758	1.6227	15.1352	
	Addis Ababa	131	64.847	.9508	10.8829	
	Total	**679**	**56.329**	**0.7464**	**19.4502**	

From the above table, there is no doubt that students had different standards when they joined grade one English program. With traditional instruction added to these differences, the gaps created continue to significantly persist by the end of grade one.

3.2.16. Students' Academic Achievements in Experimental Schools by Region

The table below shows that there were huge academic gains in the post-tests over the pre-tests results in experimental schools in all regions.

Table 21. Student -Academic Achievements in Experimental Schools by Region

Measurement	Region	No.of Participants	Mean Scores	Std. Error	Std. Deviation	Sig. (ANOVA
Post-test	Tigray	137	77.974	.9773	11.4385	.000
	Amhara	102	80.906	1.0524	10.6287	
	Oromia	114	77.631	1.2533	13.3819	
	SNNPR	130	73.826	1.1107	12.6644	
	Harariii	56	68.434	1.5014	14.9386	
	Gambella	102	65.122	1.7622	17.797	
	Addis Ababa	146	83.458	0.8544	10.3235	
	Total	**830**	**75.887**	**0.959**	**14.2864**	
Pre-test	Tigray	137	45.018	1.0077	11.7944	.000
	Amhara	102	47.892	13.1525	1.3023	
	Oromia	114	49.649	1.3920	14.8625	
	SNNPR	130	50.250	1.3499	15.3917	
	Harari	56	67.575	1.5277	15.2004	
	Gambella	102	49.779	1.8511	18.6954	
	Addis Ababa	146	72.020	1.2503	15.1068	
	Total	**830**	**54.852**	**0.6249**	**18.0023**	

From the table, it can be seen that the mean score of students for all regions amounted to 75.88 in the post-test and 54.85 in pre-test. A difference of 21.03 mean percentage point is in deed, remarkable and radio influence is very vivid.

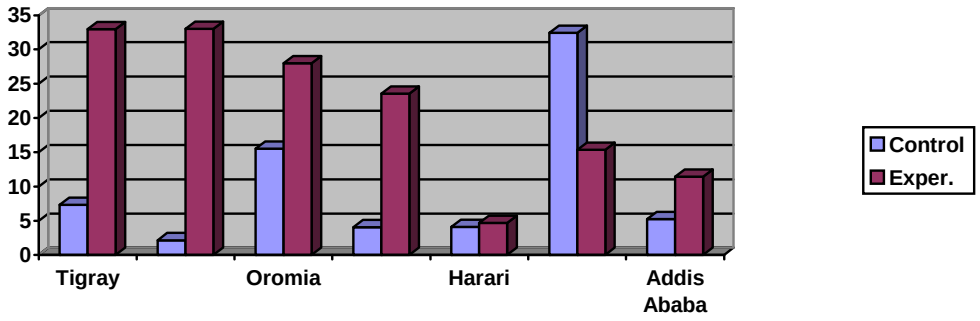
3.2.17. Students' Academic Gains in Experimental and Control Schools by Region

When a comparison of students' scores in experimental and control schools is made by region in post-test, it becomes clear from the table below that in all regions students' scores were higher in experimental schools than in control schools. Overall, experimental schools attained more than 22 percentage points differential over control schools. Between pre-tests and post-tests, overall results indicate a gain of about 10 percentage points for control schools and this is significant. From the table it is clear that both traditional instruction and IRI are effective in producing learning gains, but IRI has produced 22% gains while traditional instruction has produced a gain of nearly 10%. Overall, the use of radio instruction inverted the gaps between control and experimental schools at a regional level. Table 31 below summarizes the results.

Table 22. Student Academic Scores in Experimental and Control Schools by Region

School Type	Region	Mean Academic Scores		Mean Gains	Sig.
		Post-test	Pre-test		
Control	Tigray	58.867	51.559	7.3077	.000
	Amhara	63.538	61.384	2.1538	
	Oromia	76.422	60.883	15.5388	
	SNNPR	65.437	61.354	4.0833	
	Harai	63.0814	67.2093	4.1279	
	Gambella	60.201	27.758	32.4425	
	Addis Ababa	70.114	64.847	5.2672	
	Total	66.285	56.329	9.9560	
Experimental	Tigray	77.974	45.018	32.9562	.000
	Amhara	80.906	47.892	33.0147	
	Oromia	77.631	49.649	27.9825	
	SNNPR	73.826	50.250	23.5769	
	Harai	72.5446	67.575	4.6875	
	Gambella	65.122	49.779	15.3431	
	Addis Ababa	83.458	72.020	11.4348	
		76.887	54.1173	22.4079	

Graph 6. Learning Gains by Region



Part IV

Summary, Conclusions and Recommendations

4.1. Summary

This report constitutes the second part of the overall evaluation of the Interactive Radio English Programs for Grade One in Ethiopia. It is a summative evaluation of the programs and its overall aim is to describe the effectiveness of Interactive Radio Instruction Grade One English Programs in Ethiopia. The main purpose of the study is to measure students' learning gains as shown by an achievement test comprising English listening skills. A secondary objective is to determine how IRI programs address equity issues in Ethiopia.

The main approach to the study is a basic pre-test/teach/post-test design. As such its aim is to indicate how much the Interactive Radio English Programs in Grade One contributes to changes in student learning as compared to non-radio instruction. Participants in this study constitute control and experimental groups and they were randomly assigned to their respective groups. The dependent variable, student achievement score, was measured at two occasions: pre-test and post-test. At the pre-test both control and experimental students were given a listening comprehension test, which had been developed and recorded on a tape-recorder. After an interval of seven months (which also constituted a treatment period), the same test was given to the same students in their respective groups. The principal independent variable was traditional instruction versus interactive radio instruction. Traditional instruction was given in the control schools and there was no exposure of the students to IRI in these schools. In experimental schools students learned using IRI and the time allotted to English was fully covered by IRI. This included 5 fifteen-minute radio lessons every week. The samples were assigned on random basis to both groups and the sample size was 1466.

Main findings in the study are the following:

- Interactive Radio English Instruction Programs were more effective in improving educational quality. Students who used these programs gained 22% as compared to 9% in control schools that did not use these programs.

- In traditional or non-IRI classroom instruction, learning outcomes continue to favour urban students and inequity persists. In interactive radio instruction, the students in urban and rural classes began with a slight differential in favour of the urban classes but there were no statistically significant differences in the gains of urban and rural IRI students with each gaining nearly in equal amounts. Thus, radio was equally effective in increasing academic gains both in urban and rural schools.

- Within urban areas, a comparison of students' academic achievement in control and experimental schools indicate that students who use English IRI Programs performed significantly better than those in schools that do not use

English IRI programs. Moreover, when academic gains in urban control and urban experimental schools are compared, it was found that students in urban experimental schools gained more than two times the gain of students in control schools. This means in urban areas IRI English Programs are highly effective.

- Within rural areas, a comparison of control and experimental schools indicates that academic gains are much higher for students in experimental schools, even though the experimental schools started with a notably lower performance in comparison to control schools as shown by pre-tests. This implies that the use of interactive radio in rural areas can help to eliminate differentials between urban and rural schools.

- Students' academic achievement did not significantly differ among male and female students in control schools. In experimental schools, female students performed better than male students. Interactive Radio English Programs were more effective than traditional instruction for students of both sexes, but on aggregate favoured females over males.

- When it comes to the relationship of students' academic achievement to self-contained and non-self contained approach in grade one, the results show that higher gains are related to non-self contained approach in control schools. In experimental schools, higher student academic achievement is positively related to the self-contain approach. In both control and experimental schools, there has been a tendency for the differences in academic performance to narrow, but this tendency is much higher in experimental schools. This means that IRI is not only more effective overall, but also reduces gaps that may persist in traditional instruction.

- When it comes to the relationship of academic gains and class size, in control schools students in large classes significantly achieved more than those in smaller classes. This is against the common assumption that relates lower student performance with large classes. In experimental schools, students in larger classes had an advantage in the pre-test, but with the introduction of IRI this gap was eliminated and equality was achieved.

- With reference to the relationship between academic scores and regions, students gained significantly in all regions (except in Harari) from traditional instruction in control schools. However, student academic gains are more than two times greater in experimental schools. This suggests that IRI should be expanded in regions where the innovation thus far has not been introduced.

In summary, the results of the summative evaluation show that the IRI English lessons had a significant impact on academic achievement. These positive results occurred in all regions, in urban and rural schools and among male and female students. There was a slight advantage in terms of female students. Finally, the Interactive Radio Programs tended to reduce the differences in some areas such as class size, and the classroom structure.

4.2. Conclusions

This study has demonstrated that *English for Ethiopia*- IRI has been more effective in improving student learning than traditional instruction. This suggests that the innovation must continue and be adopted on a wider scale. This includes all regions and both urban and rural schools.

The fact that IRI has provided more opportunities to female students to perform well is a very positive tendency in Ethiopia. When it comes to teachers, students having male IRI teachers benefit more than those having female teachers. Although, this study has not controlled for the size of male and female teachers, the overall tendency suggests that training and more mechanisms can be followed to encourage female teachers to use radio as well as male teachers.

There is no question that IRI is going to continue under heterogeneous groups as far as age is concerned. The fact that it has managed to address these groups is a very strong feature it has to continue to consider. The strength of IRI to address the needs of self-contained classes has to be encouraged since this will be the future classroom structure in Ethiopia. IRI Programs in Grade One demonstrated that it is possible to meet the interests of less experienced and experienced teachers. This creates a unique opportunity particularly for rural and inexperienced teachers to get a curricular material that they may not easily find or produce in their schools. The potentials in IRI English programs to eliminate differences among urban and rural schools and among regions call for more attention to strengthen these programs in the context of Ethiopian primary education.

References

Arnott, J., Mansfield, J. and Mentis, M. (1993). A Summative Evaluation of OLSET's " English in Action" Radio Learning Program.

Bosch, A. (1997). Interactive Radio Instruction: Twenty –Three Years of Improving Educational Quality. Education and Technology Notes, Vol.I, No.1: World Bank.

Dane, F.C. (1990). Research Methods. Pacific Grove, CA:Brooks/Cole.

Lenneberg, E.H. (1970). Speech as a Motor Skill With Special References to Nonaphasic Disorder in Cognitive Development in Children. London: Chicago Press.

Stockard, J. and Mayberry, M. (1992). Effective Educational Environments. Newbury Park/ Califorina: Sage.

Thruston, L.L. and Thruston, T.G. (1993). Primary Mental Ability Manual for Grades K to One. Science Research Associates, Incorporated.