

PC-AAA-750

**FORMATIVE EVALUATION STUDY OF
INTERACTIVE RADIO INSTRUCTION
ENGLISH PROGRAM**

INTERIM REPORT

MARCH 1994

DIRECTORATE OF PRIMARY EDUCATION

NWFP

TABLE OF CONTENTS

EXECUTIVE SUMMARY..	111
1 PROBLEM.	1
2 PROGRAM BACKGROUND	1
3 STUDY PURPOSE.....	3
3.1 Purpose....	3
3.2 Indicators of success.	3
4 EVALUATION DESIGN.	4
4 1 Phases..	4
4.2 Sample.	4
4 3 Introduction of the program	5
4 4 Instruments and data collection	5
4 5 Baseline and summative evaluations	5
5 FINDINGS...	6
5 1 Testing sample.....	6
5.2 General findings	6
5 3 Specific findings..	6
6 CONCLUSIONS AND RECOMMENDATIONS.	10
6.1 Answers to evaluation questions....	10
6.2 Recommendations..	11
ANNEX A FINDINGS FROM THE FIRST IRI TRIALS--NOVEMBER 1992	13
ANNEX B: LIST OF TEHSILS AND IRI CLASSES TESTED-- DECEMBER 1993	22
ANNEX C. EVALUATION FORMS	24
ANNEX D: BASELINE/SUMMATIVE EVALUATIONS	29
ANNEX E: ENGLISH TEST I AND ITEM SCORES	32

INTERIM REPORT
FORMATIVE EVALUATION STUDY OF IRI ENGLISH PROGRAM
EXECUTIVE SUMMARY

BACKGROUND

In 1992, the Provincial Secretary of Education asked the Directorate of Primary Education to develop an English program for the primary level. Because most primary teachers had little command of English, it was decided to adapt existing Interactive Radio Programs in English to the Pakistani context. Two limited field trials showed that major changes were necessary before the program would be appropriate to schooling conditions in NWFP. After modifications based on these field trials were complete, the program was ready for broadcast to a larger number of experimental schools in the fall of 1993.

The development of the IRI program has been the responsibility of Curriculum Bureau Branch (CBB) staff detailed to the Instructional Materials Development Cell (IMDC). They have worked under the supervision of a curriculum consultant provided by the Primary Education Development (PED) Program. One CBB staff member prepares the scripts based on a set of curriculum objectives while another prepares the teachers' guides and supervises the development of workbooks by IMDC staff. The lessons are prerecorded in the studios of Radio Pakistan with the help of two producers and an engineer especially assigned to the job. The staff of the broadcasting station has also been trained under the supervision of an expert provided by PED. The lessons are recorded on cassettes.

In mid-September 1993 Radio Pakistan began broadcasting Level I English lessons. The student lessons were timed to coincide with the end of the break so an accompanying teacher's program could be broadcast before the lesson and during the time when the teachers were free. Thus, the teacher program was broadcast at 10:45 for 15 minutes, there was 5 minutes of news followed by the IRI lesson from 11:05 until 11:25, after which there was a 10 minute workbook practice organized by the teacher. Eventually there would be three levels of English, each including 120 lessons, introduced to children in Classes Three, Four and Five. In multi-classes, students at all levels listen to the programs.

Study purposes The main purpose of the present formative evaluation is to test whether interactive radio meets the standards of the developers in providing effective English instruction to primary school children. The following are the main evaluation questions

Do students develop English skills in the three areas of listening comprehension, reading and writing?

Does the IRI program produce satisfactory results in all kinds of NWFP schooling environments?

A positive response to these questions was to be taken as evidence that the IRI materials provided effective English instruction in the Pakistani context. A negative response to either question was to be taken as meaning that the program needed further review and possible modification.

The indicators of success The standards of success set by the developers to constitute evidence that the program objectives had been met were the following:

English skills To indicate that the materials effectively produced the skills of listening, reading and writing, it was expected that more than half the sample classes in the province would receive a class average score of 75 percent or higher on a test of English skills. The test was prepared by the staff of the Curriculum Bureau Branch (CBB) located in the Instruction Materials Development Cell (IMDC). The test was composed of items similar to those found in the radio lessons and workbooks at the mid year mark.

Effective in all schooling environments To indicate that the program was effective in all schooling environments, achievement scores were expected to meet the standard that a majority of classes would obtain 75 percent correct answers or higher in each type of schooling environment (male and female, urban and rural, large and small classes, single and multi-classes) and with teachers having varying characteristics (academic and professional training, service experience, English fluency, etc.)

Components tested. The program to be tested included approximately 2 months of radio lessons, one 20 minute lesson a day, student workbooks and a set of teachers' instructions.

In reality, an evaluation of this kind cannot isolate and test radio lessons and their associated materials only, especially on so large a scale as the experimental sample used in this study. Rather it tests an instructional system for introducing, monitoring and maintaining a program--only one component of which is the lessons, workbooks and guides around which the system is organized.

STUDY DESIGN

Sample A sample of 300 schools was selected for the first extensive use of the IRI program. The schools included approximately 10 (6 male and 4 female) schools in each of the major

tehsils of NWFP (see ANNEX B) The schools selected were ones where IMDC materials had already been introduced successfully, and where a system existed for introducing new programs. Schools were taken from most of the major tehsils (sub-districts) to give local officers experience with the program before extending it to all schools in NWFP. Radios were given to the 300 schools in September, and soon thereafter approximately 30 extra radios were given to additional schools. Besides the radio, each school received two sets of batteries, the teachers' guide, student workbooks, and cassettes of the lessons if radio reception was poor.

Introduction of the program Master trainers (ASDEOS) from the selected tehsils were brought in groups to the IMDC in Peshawar for training in how to introduce the program in their districts. They in turn trained the designated experimental teachers, other ASDEOs and LCs in how to use IRI in the classroom. The ASDEOs and LCs formed a chain of command to supervise the use of IRI and collect routine observation data. Radio Pakistan began broadcasting the daily lessons in mid-September 1993.

Instruments and data collection In addition to achievement tests, two proformas were used in the evaluation (see ANNEX C).

Form 1 An observation proforma was filled out each week in each class by an LC for the first two months of the program and thereafter once a month if the teachers were having no difficulty with the lessons.

Form 2 An interview/teacher characteristics form was filled out twice yearly at the same time that achievement tests were administered by the IMDC/CBB staff.

Achievement tests Achievement tests were administered before the winter break in mid-December 1993 and before the summer break in June 1994. (This paper reports the results of the December assessment only).

CONCLUSIONS AND RECOMMENDATIONS

The IRI program has been well received in the schools as well as among education officials both at the provincial and federal levels. The Federal Curriculum Wing, after observing the program's success, has organized a workshop to write a new English curriculum based on the IRI lessons. The Wing is tentatively planning to extend the program to all of Pakistan. If Pakistan Broadcasting Corporation in Islamabad broadcasts the programs to the whole of the country, this would help resolve the problem of poor reception in some areas of NWFP because of the stronger capacities of the federal radio facilities.

Conclusions after the first IRI extensive testing cycle The main program questions have been answered positively in the first testing cycle

Do students develop skills in the three areas of listening comprehension, reading and writing?

The overall scores on the achievement tests were listening comprehension 87 percent, reading 84 percent, and writing 82 percent. 84 percent of the sample classes obtained scores of 75 percent or more, much higher than the 51 percent majority required to meet the IMDC/CBB standard

Does the IRI program produce satisfactory results in all kinds of NWFP schooling environments?

Classes in all schooling environments (male and female, urban and rural, large and small, single and multi-classes) and having teachers with a variety of characteristics (academic and professional training, service experience, English fluency, etc) all obtained achievement scores of 75 percent correct answers or higher

As a consequence, the positive response to the two evaluation questions means that the IRI program in the first extensive testing trial meets the expectation of the developers. By the criteria defined in the achievement test, and by ratings of the teachers, the program is capable of providing primary school children in NWFP with effective English instruction

Recommendations The following are recommendations coming from observations during the trial

o **Speaking** Speaking is an essential skill that was not tested in the evaluation cycle. Informal tests of this skill show the children respond very well to simple questions like "What is this?" and other questions to which they have been exposed in the radio lesson. However, no systematic test was conducted of this skill. Though difficult to conduct with large numbers of children, some sort of test needs to be devised to test this skill in the IRI classrooms

o **Reading and writing** Although children scored high in the two skill areas of reading and writing, IRI by its nature is weakest in developing these skills. These skills depend more heavily on the capabilities of teachers who in the primary system of NWFP are weak in English. Careful monitoring of these skill areas should be conducted routinely with special attention reserved for workbooks and teachers guides to ensure that they adequately build these skills. Also teacher training for IRI needs to stress these two areas

o **Field support** As with all programs placed in the field, the essential requirement in effectiveness is to have strong systems to support their use. The number of schools, while not large, which could not be tested because something went awry in the implementation of the IRI program, suggests the need for continuing emphasis on these support systems. The support systems include initial training in use of the radio, the workbooks and the teachers' guides, routine supervisory visits, and periodic assessment and feedback to the teachers. In this way, it may be possible to prevent a teacher, while there is still time to do something about it, from putting the radio in the cupboard and ignoring the program.

o **Next phases.** In the next phase, when the program is extended to all schools in September 1994, the Education Department will have to make decisions about how to handle the "radio question". Will teachers or communities be willing to provide their own radios or will the government provide a radio to each school? Is there a place in the recurring budget for replacing batteries? This issue needs to be discussed and resolved before the program starts. Also the recurring budget will need to allocate resources for radio air time, which over the next years will increase as more English levels are introduced.

o **Other subjects.** The success of the IRI program in developing language skills where teacher capability is weak, suggests that similar programs may be the answer for poor skill development in other subjects. For example, existing school programs are not now able to produce high levels of Urdu competency in most primary school children. The problem is compounded in Pashto areas where the mother tongue is very different from Urdu. Many teachers themselves have difficulty with Urdu and fall back on their mother tongues when teaching in the classroom. Other subjects like math and science where teachers score very low on tests of Class Five competency are also areas where IRI could supplement existing programs or substitute for them altogether. IRI programs could provide quality instruction to nonformal classes of over-school-age children, to out-of-school girls and to adults, or to preservice or inservice teachers to improve their training.

INTERIM REPORT
FORMATIVE EVALUATION STUDY
INTERACTIVE RADIO INSTRUCTION (IRI) MATERIALS
IN NWFP SCHOOLS

1 PROBLEM

Educators, teachers, students and parents have complained that English instruction in government primary schools of NWFP is weak. They say that the available materials are difficult for teachers to use and for students to understand. Many teachers who are themselves weak in English fall back on rote memory work to simplify the task of instruction. The English they teach is poorly pronounced, heavily accented and poorly comprehended. Interactive Radio Instruction (IRI) has proved an effective means of teaching English in other countries. The Directorate of Primary Education was asked by the provincial Secretary of Education to adapt this method for use in NWFP schools.

2 PROGRAM BACKGROUND

In November 1992, a field study was conducted in 11 classes of 5 schools of the Peshawar area to test the suitability of using an Interactive Radio Instruction Program in NWFP. Among the questions this early trial was to answer was a) whether it was possible to adapt an existing IRI program to the Pakistani schooling environment, b) what kinds of changes would be required, and c) at which schooling levels should the program be introduced. The IRI program used was one designed for Kenya and rerecorded using Pakistani voices in compliance with local preference. The study indicated that changes needed to be made before the IRI program could be successfully adapted for use in Pakistan. Most significant in terms of developing the program were three points: a) that music would have to be removed in keeping with local sensitivities, b) that the radio program could not assume English competency in primary school teachers, and 3) that the radio lessons were more effective when used with older primary children who could respond to the radio cues without much teacher direction. Also, because primary students have to cope with learning two or three languages in the early primary years, it seemed better to wait until they were well grounded in these languages before introducing English.

As a result of this trial, it was decided to produce three levels of English, for Classes Three, Four, and Five. Each level would

include radio lessons, student workbooks and teachers' guides. A complete level would consist of 120 lessons, one 20 minute lesson per day with 10 minutes of workbook time for reading and writing practice after the radio lesson. The lessons were to be broadcast by Radio Pakistan on days when schools with varying holiday schedules were open simultaneously. Accompanying the radio program was a teachers' guide to translate the new vocabulary and word structures for each lesson into Urdu and to prepare the teacher for any actions she or he might need to take before or during the lesson. In addition, an alphabet and numbers workbook for Class 2 was developed to prepare children for reading and writing before they were formally introduced to the three levels of the radio program. Given the multi-class nature of most schools, children of other classes heard the lessons even when they were not required to know the material or complete the workbooks. Thus many school children would be exposed to the program from the earliest classes.

The development of the IRI program has been the responsibility of Curriculum Bureau Branch (CBB) staff detailed to the Instructional Materials Development Cell (IMDC). Under the supervision of a curriculum consultant provided by the Primary Education Development (PED) Program, one CBB staff member prepares the scripts based on a revised set of curriculum objectives while another prepares the teachers' guides and supervises the development of workbooks by IMDC staff. The lessons are prerecorded in the studios of Radio Pakistan with the help of two producers and an engineer especially assigned to the job. The staff of the broadcasting station have also been trained under the supervision of an expert provided by the PED program. The lessons are recorded on cassettes, sets of which are provided to schools in areas where reception is poor.

After the first test of the IRI program, the programs were improved and tried again in a small group of schools and again revised. Among the revisions recommended during this trial were a) even greater simplification of teachers' instructions in the radio lesson, b) an easier-to-understand teachers' guide, and c) an improved workbook. A set of ten 15 minute IRI training programs for teachers covering the first 30 student lessons were also developed and recorded.

In mid-September 1993 Radio Pakistan began broadcasting the lessons. The student lessons were timed to coincide with the end of the break so the teacher's program could be broadcast before the lesson and during the time when the teachers were free. Thus, the teacher program was broadcast at 10 45 for 15 minutes, there were 5 minutes of news followed by the IRI lesson from 11 05 until 11 25, after which there was a 10 minute workbook practice organized by the teacher.

3 STUDY PURPOSE

3 1 Purpose. The main purpose of the present formative evaluation is to test whether English interactive radio provides effective English instruction to primary school children. The following are the main evaluation questions

Do students develop English skills in the three areas of listening comprehension, reading and writing?¹

Does the IRI program produce satisfactory results in all kinds of NWFP schooling environments?

A positive response to these questions was to be taken as evidence that the IRI materials provided effective English instruction in the Pakistani context. A negative response to either question was to be taken as meaning that the program needed further review and possible modification.

3 2 The indicators of success The standards/indicators set by the developers to constitute evidence that the program objectives had been met were the following

3 2 1 English skills To indicate that the materials effectively produced the skills of listening, reading and writing, it was expected that more than half the sample classes in the province would receive a class average score of 75 percent or higher on a test of English prepared by the IMDC staff. The test was composed of items similar to those found in the radio lessons and workbooks at the mid-year mark.

3 2 2 Effective in all schooling environments To indicate that the program was effective in all schooling environments, achievement scores were expected to meet the standard of a majority of classes obtaining 75 percent correct answers or higher in each type of schooling environment (male and female, urban and rural, large and small classes, single and multi-classes) and with teachers having varying characteristics (academic and professional training, service experience, English fluency, etc.)

In reality, an evaluation of this kind cannot isolate and test radio lessons and their associated materials only, especially on so large a scale as the experimental sample used in this study. Rather it tests a comprehensive instructional system for introducing, monitoring and maintaining the IRI program--only one component of which is the lessons, workbooks and guides around which the system is organized.

¹ Speaking was not tested because of the difficulty of finding a uniform way to test so many children.

4 EVALUATION DESIGN

4 1 Phases The evaluation was conducted in three phases a) a preliminary trial phase to determine the main adaptations required to fit the program to the Pakistani context, b) a limited trial phase to test the modified program and c) an extensive experimental phase to see if children in a broad range of contexts learned English skills with the IRI program. The main findings of the first phase have already been noted above (details appear in ANNEX A). In the second phase, 30 draft lessons (on cassettes) with support training, and workbooks were introduced in 5 schools (2 boys' and 1 girls' urban school, and 2 boys' rural schools) between April of 1993 and the time for the summer break in June. The urban schools were visited once a week by developers and producers while the more distant rural teachers reported periodically to the IMDC for debriefing about their experiences with the program. Students in all schools were tested at the end of the trial period. The purpose of this phase was to finalize the radio lessons before disseminating them to a larger number of schools. Evaluation instruments and procedures for introducing the lessons were field-tested during this phase.

In the summer of 1993, the 30 lessons used in the limited trial, with their associated support materials and testing instruments, were modified and put into final form before IRI was broadcast to a larger sample of schools starting in September 1993. The developers continued producing the lessons for Level One, the 120 lessons of Level Two, and the 120 lessons of Level Three. It was assumed that the radio lessons, which were extensively modified according to recommendations of the first two evaluation phases, would require very little change in structure and approach in future development.² However, the teacher guides and workbooks were still given special scrutiny to ensure that they produced expected results in the weaker areas, for IRI, of reading and writing.

The present study is concerned with this third evaluation phase which involved the broadcast of Level One lessons to 300 schools starting in mid September 1993.

4 2 Sample An extensive sample of 300 schools was selected from the major tehsils (sub-districts) in NWFP. The schools included approximately 10 (6 male and 4 female) schools in each tehsil (see the attached tehsil list in ANNEX B). The IRI schools were ones where IMDC materials had already been introduced successfully, and where a system existed for introducing new programs. Schools were

²An important consideration in speeding up development of the three levels of English, and not waiting for further refinements from on-going evaluations, was the announcement that the PED program would discontinue funding in the summer of 1994.

selected in most major tehsils to give local officers experience with the program before it would be extended to all schools. Radios were given to the 300 schools in September, and soon thereafter approximately 30 extra radios were given to additional schools. Besides the radio, each school received two sets of batteries, the teachers' guide, student workbooks, and sometimes cassettes of the lessons when radio reception was poor.

4.3 Introduction of the program Master trainers (ASDEOS) from the districts were brought in groups to the IMDC in Peshawar for training in how to introduce the program in their districts. They in turn trained the designated experimental teachers, the other ASDEOs and LCs in how to use IRI in the classroom. The ASDEOS and LCs were organized in a chain of command to supervise the use of the IRI program and to collect routine observation data.

4.4 Instruments and data collection In addition to an achievement test, two proformas were used in the evaluation (see forms in ANNEX C and achievement test in ANNEX E).

Form 1 An observation proforma was filled out each week in each class by an LC for the first two months of the program and thereafter once a month if the teachers were having no difficulty with the lessons. Part of the reason for this form was to provide some evidence that the supervisor had visited the classes.

Form 2 An interview/teacher characteristics form was to be filled out twice yearly, at the same time achievement tests were administered by the IMDC.

Achievement tests Achievement tests were administered twice yearly (before the summer break in June and before the winter break in mid-December). The first testing in mid-December 1993 was conducted in all schools by the IMDC/CBB staff. In May 1994 it will be conducted in 8 schools (5 boys and 3 girls schools) in each of 8 districts by CBB staff, and in the rest of the schools by local supervisors.

Each visit to the field yielded observational and descriptive information about how to improve procedures to support IRI materials. This information will be used when IRI English is disseminated to all schools in NWFP and as a base for preservice and inservice teacher training courses.

4.5 Baseline and summative evaluations The original design for the introduction of the IRI program in NWFP called for a pretest of English skill level before IRI was introduced and after a year of the program. A pretest was therefore conducted in February 1993 that yielded better-than-expected results given the observable weaknesses in the regular primary English program. Subsequent analyses revealed major flaws in both the test and the design, and

the effort to conduct the baseline study was subsequently abandoned (See ANNEX D for more details) At present, criterion-referenced tests are used to monitor the progress of English skill development

5 FINDINGS

5 1 Testing sample The first cycle of student testing was completed in December 1993 (see ANNEX B for the list of tested classes by tehsil) Of the approximately 330 schools participating in the program, IMDC/CBB staff tested 303 classes in 280 schools Chitral which had 13 schools following the program tested its own schools because of its inaccessibility (They were only able to send the results of four schools in time for this report) In Kohistan only 4 out of 10 schools could be tested because most were located too far away from the center In the rest of the schools, the program either was not implemented for various reasons or it was implemented late and therefore could not be tested

The findings are as follows

5 2 General findings The first results of the testing were very positive The overall provincial test result was 85 percent correct on the achievement test IMDC sets the standard for its programs that a majority (over 50 percent) of classes achieve an average 75 percent correct on an achievement test Altogether 84 percent or 254 classes had 75 percent or higher class scores and only 16 percent or 49 classes had less than 75 percent scores

Of the 303 classes tested, 184 were boys classes and 119 were girls classes The girls produced higher scores with 92 percent (109) of their classes meeting the standard, compared to the boys with 79 percent (145) meeting the standard

5 3 Specific findings

5 3 1 **Class characteristics and English achievement** One of the advantages of IRI is said to be the fact that it delivers a uniform quality educational program to all schooling environments regardless of the characteristics of that environment or the skill of the teaching staff that are found there Table 1 shows the IRI English class average scores in a variety of schooling environments in NWFP The class scores are virtually the same in all schooling environments tested, thus confirming for NWFP the uniform quality of the program

Table 1 Class characteristics and student achievement (Extensive IRI evaluation cycle)

Characteristics	Average class score of students (% correct)
Provincial average	85
Gender	
Male	83
Female	89
Location	
Urban	86
Rural	85
Class composition	
Single	87
Multi-class	85
Size of class	
1 to 20	86
21 to 40	87
41 to 60	85
60+	82

5 3 2 **Teacher characteristics and English achievement** IRI was designed for use where teachers do not know much English Since the radio substitutes almost completely for the teacher, one would not expect the characteristics of the teacher to have much effect on student achievement results Table 2 confirms that this is true The students of all kinds of teachers perform at essentially the same level The one exception was the lower performance of students having teachers with Kohistani as a mother tongue Only 4 out of 10 schools were tested in Kohistan and the lower result is more likely to represent such factors as the difficulty of distributing materials and radios to distant schools and the irregular attendance of students from migratory families

Table 2 Teacher characteristics and student achievement (Extensive IRI evaluation cycle)

Characteristics	Average class score of students (% correct)
Provincial average	85
Gender	
Male	83
Female	89
Academic	
Middle	85
Matric	86
FA	87
BA/BSc	85
MA/MSc	87
Division	
I	90
II	86
III	84
Profess Training	
Untrained	82
PTC	86
CT	87
BEd	85
MEd	89
Teaching experience	
Up to 3	86
4 to 10	86
11 to 15	89
16 or more	82
English fluency (self-rated)	
None	93
Poor	84
Fair	85
Good	89
Mother tongue	
Urdu	82
Pashto	86
Hindko	89
Kohistani	77

5 3 3 **Teacher ratings of IRI materials** A majority of teachers expressed the view that the IRI program was a good way to teach English to primary school children Table 3 shows that the majority expressed positive feelings about the program being interesting to children, helping them learn skills, and proving useful to the teacher The results in this table unfortunately are not completely accurate The translations of the English words "fair" and "good" in the rating scale were not accurate in Urdu Using the Urdu meanings, almost 100 percent ranked the program positively on all points

Table 3 Teacher Ratings of IRI English program

Characteristic	% Rating		
	Poor	Fair*	Good*
Interesting to child	1	36	61
Helped child learn skills	0	38	62
Useful to the teacher	0	39	61

* These rankings were mistranslated into Urdu as better ("fair") and best ("good")

Teachers were also asked to compare the IRI English Program with the previous English program they used Almost all responded that it was better than the previous program they had used

5 3 4 **Questions about difficulties with the program** Because radio lessons depend for their impact on listeners understanding the words that are spoken and responding appropriately to cues, teachers were asked about their own and students' difficulties with these aspects of the program They were also asked about any difficulties they had had in completing the workbook exercises Most teachers responded that neither they nor their students had any problems with these activities Table 4 shows their responses to these questions

Table 4. Teacher-reported difficulties with English program

Difficulty	% Answering		
	Yes	No	Sometimes
Teachers			
Understanding words	2	84	14
Using teacher instructions	4	92	4
Cuing students correctly	3	93	4
Helping with workbooks	3	92	5
Students			
Understanding words	3	81	15
Responding to cues	4	86	10
Completing workbook exercises	4	87	8

5 3 5 Test items The test (see ANNEX E for the test and an item break down) was composed of three parts listening comprehension, reading and writing. The children scored well in all three areas listening 87 percent, reading 84 percent and writing 82 percent. Observers, however, felt that the children were not responding as quickly to questions requiring reading and writing skills. The few test items in these two skill areas may not have picked up weaknesses that in fact exist.

6 CONCLUSIONS AND RECOMMENDATIONS

The IRI program has been well received in the schools as well as among education officials both at the provincial and federal levels. The Federal Curriculum Wing, after observing the program's success, has organized a workshop to write a new English curriculum based on the IRI lessons. It is tentatively planning to extend the program to all of Pakistan. If Pakistan Broadcasting Corporation in Islamabad broadcasts the programs to the whole of the country, this would help resolve the problem of poor reception in some areas of NWFP because of the stronger capacities of the federal radio facilities.

6 1 Answers to the study questions The main study questions have been answered positively in the first testing cycle.

Do students develop skills in the three areas of listening comprehension, reading and writing?

The overall scores on the achievement tests were listening 87 percent, reading 84 percent and writing 82 percent. 84 percent of the sample classes obtained scores of 75 percent or more, much higher than the 51 percent majority required to meet the standard.

Does the IRI program produce satisfactory results in all kinds of NWFP schooling environments?

Classes in all schooling environments (male and female, urban and rural, large and small, single and multi-classes) and having teachers with a variety of characteristics (academic and professional training, service experience, English fluency, etc.) all obtained achievement score of 75 percent correct answers or higher.

As a consequence, the positive response to the two evaluation questions means that the IRI program in the first extensive testing trial meets the expectation of the developers. By the criteria defined in the achievement test, and by ratings of the teachers, the program is capable of providing primary school children in NWFP with effective English instruction.

6.2 Recommendations The following are recommendations that come from observation in the field.

- o **Speaking** Speaking is an essential skill that was not tested in the evaluation cycle. Informal tests of this skill show the children respond very well to simple questions like "What is this?" and other questions to which they have been exposed in the radio lesson. However, no systematic test was conducted of this skill. Though difficult to conduct with large numbers of children, some sort of test needs to be devised to test this skill in the IRI classrooms.

- o **Reading and writing** Although children scored high in the two skill areas of reading and writing, IRI by its nature is weakest in developing these skills. These skills depend more heavily on the capabilities of teachers who in the primary system of NWFP are weak in English. Careful monitoring of these skill areas should be conducted routinely with special attention reserved for workbooks and teachers guides to ensure that they adequately build these skills. Also teacher training for IRI needs to stress these two areas.

- o **Field support** As with all programs placed in the field, the essential requirement in effectiveness is to have strong systems to support their use. The number of schools, while not large, which could not be tested because something went awry in the implementation of the IRI program, suggests the need for continuing emphasis on these support systems. The support systems include initial training in use of the radio, the workbooks and the teachers' guides, routine supervisory

visits, and periodic assessment and feedback to the teachers. In this way, it may be possible to prevent a teacher, while there is still time to do something about it, from ignoring the program altogether.

o **Next phases.** In the next phase, when the program is extended to all schools in September 1994, the Education Department will have to make decisions about how to handle the "radio question." Will teachers or communities be willing to provide their own radios or will the government provide a radio to each school? Is there a place in the recurring budget for replacing batteries? This issue needs to be discussed and resolved before the program starts. Also the recurring budget will need to allocate resources for radio air time, which over the next years will increase as more English levels are introduced.

o **Other subjects** The success of the IRI program in developing language skills where teacher capability is weak, suggests that similar programs may be the answer for poor skill development in other subjects. For example, existing school programs are not now able to produce high levels of Urdu competency in most primary school children. The problem is compounded in Pashto areas where the mother tongue is very different from Urdu. Many teachers themselves have difficulty with Urdu and fall back on their mother tongues when teaching in the classroom. Other subjects like math and science where teachers score very low on tests of Class Five competency are also areas where IRI could supplement existing programs or substitute for them altogether. IRI programs could also provide quality instruction to nonformal classes of over-school-age children, to out-of-school girls and to adults, or to preservice or inservice teachers to improve their training.

ANNEX A
FINDINGS
FROM THE FIRST TRIAL OF
INTERACTIVE RADIO INSTRUCTIONAL (IRI) MATERIALS
IN NWFP SCHOOLS--NOVEMBER 1992

SOURCES FOR THE FINDINGS

The findings and recommendations below come from 4 sources a) systematic observations of 5 IRI lessons in 11 classrooms by English speaking Pakistani private school teachers and government GCET trainers b) a two-hour debriefing session of the same observers c) descriptive reports by a number of specialist observers including those with a wide variety of specialist backgrounds relevant to IRI implementation and d) my own observations in all but one of the classes

GENERAL FINDINGS

Children and teachers both seemed to enjoy the test lessons very much. When asked by the observers, they invariably responded that they liked to learn/teach English in this way. All but one teacher felt the lessons were better than those they had used before. Teachers reported that they felt the radio lessons were a good way for children to learn English. Overall, children showed a high level of engagement in the lessons as demonstrated in periodic counts of engagement by observers.

There was also concrete evidence that children learned with this method. Most children answered informal questions posed by observers correctly showing they understood the lessons and the vocabulary. They could recall English words and attach mother-tongue equivalents to the words.

The lessons required that the children think rather than simply parrot responses. Several observers noted that children comprehended the lessons better than their teachers, responding appropriately to instructions much of the time without looking to them for cues, and sometimes even disregarding wrong cues. The lessons were therefore a positive educational experience over and above the language skills they taught.

Modifications will be required in every lesson, since much the same difficulties are found in all of them. Overall these modifications require making the lessons more systematic and teacher-independent and improving on the technical points described below and in a memo prepared by Pamela Brooke. The main problem that needs to be

overcome is the teachers' lack of English competence. The teachers selected for the test unintentionally had much higher schooling qualifications than are generally found among average primary teachers, yet still had difficulties with the language. The general form of the lessons and mode of presentation appear to work well in Pakistani classrooms and should be retained.

The responses of both teachers and students improved over the five lessons of the test.

SPECIFIC FINDINGS

The following observations were made during the test period.

Reception Because the lessons were on tape, the radio could be turned to the right volume so all children could hear the lessons easily. Sometimes the observers rather than the teachers adjusted the radio, which may mean that the teachers need to be made conscious of the importance of tuning the radio so all children can hear it properly. Observers were not able to test whether reception was a problem, and therefore this may be a potential problem that needs to be kept in mind. Similarly the questions of electrical supply, the cost of batteries, and who should pay for them was not addressed.

Language The diction and pronunciation of the radio voices was poor. This raises the question of whether any recognizably local accent will produce English of an acceptable international standard. It will be important to find speakers who speak clearly and avoid characteristic mistakes such as "vun" for "one," and "datz" for "that's." Animal sounds should be accepted English ones, such as "moo" for cow.

The arguments about which accent to use included the following.

Pakistani. For Teachers understand a Pakistani-accented English best and all the teachers in the test stated a preference for that accent if the programs continue. A Pakistani accent does not raise issues of colonialism or "cultural imperialism" and therefore is less vulnerable to criticism than British or American accents. **Against** People who learn this accent have difficulty understanding "international" brands of English. IRI is an opportunity to give the up-coming generations "good" English. Pakistani-accented English has regional variations that will make it more difficult to expand IRI to other provinces. The accent also has class characteristics that complicate its use.

British. For: It is a native accent, better understood by Pakistanis than American accents. **Against:** It raises political sensibilities. It has less universality in terms of technology and media.

American. For. It is a native accent that is becoming more important internationally in technology, print and visual media "It has an easier grammar than British " **Against** Pakistani are not accustomed to this accent and it raises political sensibilities

There is a strong educational argument for teaching a native accent, as would be the case if French or any other language were taught Arab countries have spent large sums of money to import native speakers A compromise might be to take a fully competent Pakistani English speaker who had the opportunity to become bilingual at an early stage

Character names: No difficulty was raised with the female name "Bano " The male name Ali, however, was objected to on two grounds, 1) that the pronunciation becomes incorrect when sung (two "Ls" become necessary when sung to two drawn-out beats), and 2) that Ali is so associated with one Islamic sect that it may be objected to in some districts The name Aslam was offered as an alternative

Accompaniment The music was lively and considered generally appropriate by the observers (later ASDEOs objected to the music and said they feared complaints from conservative teachers and parents) but at certain points it was so loud it obscured the voices, and during pauses where teachers were asked to speak in mother-tongue, it was so loud children could not hear the teacher's instructions This apparently was a technical problem that can be corrected in the final product For reasons of local sensibilities about music, efforts should be made to keep the tunes simple and geared mainly to conveying vocabulary (no rhythmic nonsense sounds or actions that respond to the music beyond imparting vocabulary meaning)

Gestures The gestures which we recommended to teachers for cuing children to respond to radio instructions caused confusion until teachers ignored them and resorted to their own normal gestures Some of the gestures already had established meanings (raising the hands meant for children to stand, not respond verbally) that contradicted the meanings intended for the cuing If teachers are to be used extensively for cuing children in future programs, teachers should be allowed to develop their own gestures for sing, silence, continue singing, stand, etc

Class level The lessons worked better with Class Three children than with Class One or Kachi children The latter became restless after a short time, were highly dependent on teacher instructions, and frequently made wrong responses even when given the proper cues Class Three children, on the other hand, comprehended the lessons well enough in some cases to compensate for poor teacher direction The lessons, perhaps unintentionally, seemed aimed at the level of Class Three children There was too much listening, too many new words and too many instructions for younger children,

they enjoyed the lessons but had difficulty following them

Written instructions for teachers The printed instructions for the teachers contained a number of mistakes which made it difficult for the observers to determine how well they might have prepared the teachers for cuing the students. Several observers felt that teachers, in any case, would not actually have read the instructions to prepare for lessons. An observer with experience in the field suggested that teachers would read the instructions if they helped him or her "save face" by assisting with the English.

Written instructions seem to be a necessary way to convey to a teacher what he or she must know about the lesson. Everyone agreed that instructions should be clear, concise and simply written in Urdu, and should appear in a consistent format that teachers can become familiar with. All new vocabulary should be translated into Urdu so teachers fully understand the English. Teachers need instructions in what they must do before, during and after the class.

Workbook activities if designed with mostly self-instruction can help reinforce lesson content without depending overly much on teachers.

Teacher difficulties During the lessons, teachers were observed having difficulty with the following:

Cuing children to respond with the proper rhythm so they answered in the pauses meant for their responses. Without this, they drowned out the word that was being modelled.

Identifying specific word commands. Instead of singing, "Stand up, sit down," for example, teachers cued the children to do the action of standing up and sitting down. This meant the children did not learn the word "sing."

Understanding oral commands. Teachers had difficulty with the radio instructions "Say good morning teacher," "In your mother tongue, tell the children," "put a line on" (some observers did not know what was required with this last instruction), and sound stories.

Teacher direction Most teachers did not comprehend English enough to follow any but the simplest instructions. Almost all did not respond appropriately to requests to give instruction in the mother tongue. If they are to continue to be given instructions in English, the English must be extremely simple and accompanied by some other reinforcement like a tone that tells them when they should react.

Children's responses Children sang songs enthusiastically even though (probably because of poor diction of the voices) they

sometimes missed words like "it's time to get up and go " They may have to go through these songs more slowly the first few times

They usually looked for teacher direction, but in a number of cases they responded in ways contradictory to teacher instruction, often in more correct form than asked for by the teacher

They caught on quickly to the idea that they needed to listen and then respond As the lessons progressed they became better and better at doing this

Children and observers sometimes had difficulty distinguishing goat sounds from sheep, etc

Student engagement In general the children were orderly and attentive during the lessons However, there were variations in their engagement in the lesson Where patterns could be discerned in the observers' proformas, student engagement was usually highest in the middle of the lesson and lowest at the beginning and ends The beginning might be explained by the fact that a number of teachers made no effort to focus student attention before they turned on the radio and it took awhile to get children fully engaged in the lesson The lower student engagement at the end may be a result of simple "attention fatigue" after a period of intense concentration

Children also tended to be most highly engaged during the songs and when they were asked to do some activity such as stand up and sit down, clap hands and snap fingers, point at, and verbal responses They became restless during dialogues of Bano and Ali, especially when these continued on without giving them a chance to respond There tended to be more of these long dialogues in the beginning of the lessons which may also have contributed to lower attention at this time Shorter modelling should help to solve this problem

Workbook assignments. Class Three students could write numbers easily because of their previous experience with numbers and letters If radio lessons are to be the only English course children receive, then there needs to be a more careful plan devised for teaching writing Writing should proceed more slowly than in Lesson 5 where children were suddenly asked to write numbers from 1 to 10 Attention will have to be given to forming letters and numbers as well as to the direction of writing from left to right One option might be to continue teaching younger children English alphabet and number writing while they listen to radio lessons for the proper pronunciation Workbooks (notebooks) also seemed to work better with Class Three than One

Vocabulary Certain vocabulary words gave difficulty for a mix of reasons They need to be reviewed for their appropriateness, especially when it is expected that certain objects will be available in the classroom for demonstration Pencils and

notebooks, for example, are not necessarily present in the classroom. Objects may be known by other English terms ("copies" instead of notebooks) that conflict with the radio usages. Some objects such as pencils are difficult to point at so that the whole class can see. When they are held up, the meaning of "point at" is lost. Scripts reviewed with these points in mind should work better.

Observer presence. Someone raised the question of whether the teachers would have worked so hard at trying to understand the English and guide the class if observers had not been present. How would the lessons work without the presence or attention of a teacher? Can cues be made so simple that the lessons are virtually teacher-independent?

Teacher training. The training program which was given to the teachers before the test was not adequate. Teachers need actual practice in cuing "students" to follow lessons. Simply showing them a video of Lesson One and giving them a verbal presentation of points to follow is not enough. During the test, several teachers asked questions that showed they had not fully absorbed even the little training they received or were unclear about what they were to do. Questions such as the following were asked:

Can I push "pause" and translate the new words in Pashto?

Can I teach the new words before (or after) the class?

Can I play the tapes again for students to practice?

What do I do with the students' notebooks after they have copied from the blackboard?

A follow-up training for the teachers after a period of using the lessons would be ideal so teachers could raise any difficulties they faced. An alternative would be to anticipate questions in the teachers' guide or on a training program broadcast and answer them.

The teachers seemed to feel insecure facilitating a program they were hearing for the first time. Provision for screening the program before the lesson, as in a radio broadcast the night before would help.

It was suggested that training programs for teachers take two forms: a product-specific training that helped teachers become better facilitators, and a capacity-building training that helped teachers improve their English language capability. Both could be broadcast in the evening and some arrangement might be made for testing participating teachers and giving those who meet a specified standard, a certificate of English teaching competency, perhaps through AIIOU.

CONCLUSIONS

Overall, although there were numerous problems in the details of the program lessons, teacher notes, and training, none of these appear insurmountable if modifications are made that address the problems noted above

Answers to evaluation questions:

Do teachers behave according to directions from the radio voices and the notes in teachers' guides?

By the end of the test period most teachers were responding correctly most of the time to the radio cues. The written instructions contained mistakes that made it difficult to determine whether they could be used effectively. None of the teachers reacted correctly to all the radio instructions, and there was a great deal of variation in teacher responses, from no reaction to instructions to incorrect reactions to appropriate responses. In the teachers' defense the cues and written instructions were not always clear, and some required a level of English that the teachers did not nor could not be expected to possess. Gestures recommended in training turned out to be gestures that teachers normally use with different meanings which caused some confusion at first.

Do students respond accurately to the prompts of the radio program and the cues of their teachers?

Students generally responded accurately to radio and teacher cues, with some exceptions. Most of their confusions had improved during the test period. Sometimes children responded correctly when teachers gestured incorrectly. Improvements can be made in the diction and pronunciation of words that children either heard incorrectly or imitated improperly from inadequate modelling. With better radio voices and teacher understanding of lesson instructions, most of the problems students experienced would be solved.

Does the adapted IRI program appear to work equally well in varying schooling contexts?

The IRI lessons worked better with the Class 3 children than with Class 1 children. Vocabulary was presented too quickly for Class 1 children, and the continuous rather rapid pace of the lessons made it difficult for small children to follow correctly even though they participated enthusiastically. Multi-classes will all invariably listen to programs because of their spatial proximity, but each group responds according to its own capacity to listen and respond. Again classes from two to five seemed to follow the lessons better than younger children. Large classes sometimes had difficulty hearing cues, when the noise of singing or clapping

overpowered the radio voices. Some observers felt the girls listened and responded more accurately than the boys. Urban/rural differences were not marked except when they were associated with differences in class size.

RECOMMENDATIONS

Recommendation One: If IRI lessons and teacher notes are modified as detailed above, it seems likely that they can be successfully adopted in Pakistani primary schools. They should be introduced first in a small number of "laboratory" classes to confirm that their deficiencies have been corrected. Then they can be extended to an experimental group of 300 schools and finally to all schools.

Recommendation Two: If IRI lessons are to continue in much the same pattern as exists now, then they should be offered from Class Three on. The three IRI levels can be introduced to Classes Three, Four and Five consecutively after a period of adjustment. In multi-classes (which constitute the vast majority of NWFP schools) all the children might listen to the radio, but workbooks would only be offered to Classes Three through Five. By the time younger children reached Class 3, they should be very good at the listening/responding parts. A special program to teach basic reading and writing skills might be introduced in the early years. Head teachers would be given the discretion to use the arrangement that best fits their school, including possibly finding more radios and covering more classes. The only limitation would be in the classes for which workbooks were designed.

To ease into a regular schedule, the first three years might be as follows:

In the first year of the program introduce Level One workbooks in Classes 3, 4, and 5.

In the second year introduce Level One workbooks in Class 3, and Level Two workbooks in Class 4, and 5.

In the third year introduce Level One workbooks in Class 3, Level Two workbooks in Class 4, and Level Three workbooks in Class 5, and in each following year follow this same procedure.

Other recommendations

- o Lessons should be reduced to 18-20 minutes plus a 10 minute workbook exercise. This corresponds with the class lesson period.

- o All music should be removed. Tones and other uses of rhythm which can not be construed as singing or melody can be used.

- o All cultural problems marked by the script writer should be removed or modified, including the name change for one of the radio voices
- o Dialogue between actors should be cut to a minimum so there is more practice by children. Make sequences shorter when they are introduced so children understand all words. Example "It's time to get up and go "
- o Consistent cues such as tones when students or teachers must respond should be provided. Cues that help children respond in the proper interval rather than on top of voice model would also be useful
- o Animal sounds were sometimes not recognizable and carried on for too long
- o Minimize instructions for teachers during the radio lesson. Prepare teachers for anything they must do by writing instructions in Urdu in teachers' guides and giving Urdu equivalents of new English vocabulary words
- o Teachers' guides need to be written in very simple Urdu because teachers are often weak in that language and are not used to reading to prepare themselves for class. Include instructions for what teachers should do before, during and after the radio broadcast
- o Workbooks need to be as self-evident as possible to compensate for teachers' low level English skills
- o Provide lots of review because of interruptions in the school year
- o Radio voices need to be either native or bi-lingual speakers of English with little accent and clear diction
- o Workbooks should have a picture of a boy and a girl on the cover so teachers/students have something to point at when learning gender structures in single-sex schools
- o Care should be taken not to expect certain objects in classrooms such as pencils. The script writer has marked most of the ones that are not available

ANNEX B

LIST OF TEHSILS AND IRI CLASSES TESTED
IN NOVEMBER AND DECEMBER 1993

District	Tehsil	Boys' classes	Girls' classes	Total
ABBOTTABAD	ABBOTTABAD	7	4	11
BANNU	BANNU	7	4	11
CHITRAL	CHITRAL	4	-	4
	MASTUJ	-	-	3
CHARSADDA	CHARSADDA	7	6	13
	TANGI	8	6	14
D I KHAN	D I KHAN	4	4	8
	KULACHI	6	3	9
DIR	TIMARGARA	6	4	10
KARAK	KARAK	6	4	10
	B D SHAH	5	-	5
KOHAT	KOHAT	6	7	10
	HANGU	3	4	7
KOHISTAN	DASSU	1	1	2
	PATTAN	1	1	2
MANSEHRA	MANSEHRA	5	4	9
	BALAKOT	4	2	6
MARDAN	MARDAN	7	6	13
	TAKHAT BAHI	6	4	10
MALAKAND	SAMA RANIZAI	8	6	14
	SWAT RANIZAI	7	4	11
NOWSHERA	NOWSHERA	6	5	11
PESHAWAR	PESHAWAR	14	6	20
SWAT	SAIDU SHARIF	8	10	18
	ALPURI	5	3	8
SWABI	SWABI	8	3	11
	LAHORE	6	2	8
BUNNER	DAGGAR	6	2	8

23

HARIPUR	HARIPUR	4	5	9
LAKKI	LAKKI	8	3	11
TANK	TANK	6	3	9
BATTAGRAM	BATTAGRAM	5	3	8
<hr/>				
	TOTAL	184	119	303

ANNEX C

EVALUATION FORMS

Date _____

FORM-1

EVALUATION STUDY OF IMDC MATERIALS
TEACHER INFORMATION SHEET

This questionnaire should be completed at the time of testing for each subject

1. District of school _____
2. Tense _____
3. Village/town _____
4. School name _____
5. Official type of school boys _____ girls _____
6. Location _____
7. Class level of new materials kachi _____ One _____ Two _____ Three _____
8. Subject of test Urdu _____ Math _____ Pashto _____
9. Number of test Test 1 _____ Test 2 _____ Test 3 _____
10. Teacher name _____
11. Teacher sex male _____ female _____
12. Experience _____ (Number of years teaching primary)
13. Teacher mother tongue _____
14. Schooling (tick one)
 - _____ Middle pass _____ Matric _____ FA, FSc _____ BA, BSc _____
 - _____ MA, MSc _____
15. Division level (tick one)
 - _____ First division _____ Second division _____ Third division _____
16. Teacher training (tick one)
 - _____ Untrained _____ PTC _____ CT _____ BEd _____ MEd _____ Other (please explain) _____
17. How many classes are you responsible for teaching altogether _____ (Number of classes, include kachi as a separate class)
18. How many students are you responsible for teaching altogether _____ (Number of students, including kachi, One, Two, etc)

19 During the trial, how often did you use the new student books?

Not at all _____ Some days _____ Daily _____

If you did not use the student materials, why not? _____

20 Do you have any comments about the student textbook _____

21 During the trial, how often did you use the method in the teacher guides?

Not at all _____ Some days _____ Daily _____

If you did not use it daily, explain why not? _____

22 Do you have any comments on the teachers' guides?

23 How would you rank student books on the following points

	Poor	Fair	Good
23a Suitable to the age of children a			
23b Objects in book known to children b			
23c Interesting to children c			
23d Help children learn subject d			

24 How would you rank teachers' guides on the following points

		Poor	Fair	Good
24a	Directions are clear	a		
24b	Events/people in stories are similar to those in local area	b		
24c	Easy to teach with this method	c		
24d	Useful to teacher in teaching this subject	d		

25 How would you compare the new books with what you used before for this class?

		Not as good	The same	Better
25a	Systematic organization of lesson	a		
25b	Interesting to the child	b		
25c	Related to children's experience	c		
25d	Helping child learn	d		
25e	Easy to teach	e		

CLASS OBSERVATION FORM

A separate form should be filled in for each teacher Do not fill a form when the class teacher is not present

- 1 Supervisor name _____ 2 Designation _____
 3 Date _____ 4 Arrival time _____ 5 Departure time _____
 6 Teacher name _____ 7 Teacher number _____
 8 School name _____

- 9 How many students are taught by this teacher? _____ (No present)
 10 How many classes are taught by this teacher? _____

TICK THE CORRECT BOX

	NO	YES	No opport *
11 Children are seated in the class so all can see the blackboard			
12 Children are seated so teacher can walk among them and monitor their work			
13 Children are grouped in ways suitable for instruction (by class, ability, etc)			
14 Teacher maintains an orderly classroom where learning is not disturbed			
15 Teacher maintains discipline without harsh punishment, physical or verbal			
16 Teacher revises previous work to ensure children have learned it			
17 Teacher explains work in clear and simple ways			
18 Teacher guides student practice with tasks/questions that help them learn			
19 Teacher gives independent work so children practice by themselves			
20 Concrete examples/visual aids are used to help children understand lesson			
21 Teacher gives feedback to children in a way that helps them understand lessons			

*No opportunity to observe

22 The teacher gives homework (Supervisor check homework)

Not at all____ Some days____ Most days____ Every school day____

23 The teacher corrects homework (Supervisor check homework)

Not at all____ Some days____ Most days____ Every school day____

24 List the strong points of the teacher in teaching this lesson

25 What improvements can the teacher make in his/her teaching?

26 Did the teacher make improvements requested during the supervisor's last visit? (Tick one)

- No improvement was requested_____
- None of requested improvements were made_____
- Some requested improvements were made_____
- All requested improvements were made _____

I have discussed all the above points with the teacher so he/she understands what I have observed. I, the undersigned will visit this teacher during the month of _____ to follow up

Signature of supervisor

The supervisor has discussed with me the strong points in my teaching and ways I can improve. I will try to make the improvements requested by the time the supervisor visits again

Signature of teacher

ANNEX D

BASELINE/SUMMATIVE EVALUATIONS

The original design for implementing the IRI program in Pakistan called for providing a baseline test of student English knowledge and comparing the results with post-test results after a year of IRI instruction. For a number of reasons it was not possible to do this kind of evaluation. This annex chronicles the efforts to perform the task and the reasons why it failed.

In February 1993, at the end of the 1992/93 school year, a sample of 30 classes from Bannu, Swat and Peshawar was selected. In each of the two districts of Bannu and Swat, 3 girls' and 7 boys' schools were drawn randomly from areas near enough to the district capital to make it possible to supervise them regularly. The Bannu and Swat schools had been among those selected originally for the introduction of IMDC Kachi materials. The Peshawar schools were selected from the approximately 40 schools that make up the Cantonment area (other areas of the district were slated for other experimental programs). These schools were all schools where IRI materials were to be introduced in September 1993. Students from Classes Three, Four and Five were tested using the same test that was to be given students of the same teachers after the first level of IRI English was completed in the following year. The test was recorded on cassette for uniform administration. If students of Classes Four and Five showed high performance on the test, then a similar procedure would take place in February 1994 and 1995, i.e. tests for Levels 2 and 3 would be administered in consecutive years before each level of IRI was introduced.

This lapped design presumably allowed for comparison of IRI and normal English programs under similar conditions of instruction. The "control" consisted of the students of teachers at the end of the year preceding the introduction of IRI, while the "experimental" group consisted of the students of the same teachers a year later after the treatment. The theory was that pre-intervention scores could be compared with post-intervention scores more reliably if there were no contamination by the new course. A main assumption was that the same teachers would be teaching similar children under the same conditions each year, making the "control" and experimental groups more comparable.

The problems registered about this evaluation were the following:

1. **Assumption about the stability of staffing.** The design assumed that teachers would stay the same from year to year. Our experience in the introduction of other primary materials was that teachers transfer continuously and that it would be even more likely for them to change when the subject is English. During the duration of our week-long test of IRI in November 1993 teachers in 3 out of 8 of the experimental classes changed. The likelihood of

transfer was even greater as the time duration increased, so that one could expect very few of the same teachers in the experimental classes to remain a year later

2 Assumption that there will be little contamination of the "control" group The contamination in fact is of the IRI lessons. Practically speaking, the comparison at the time of the end of the first IRI year will be between the performances of children having 3 years of the normal course and those having 2 years of the normal course plus one year of the IRI course. After IRI is fully implemented at 3 levels, conditions will change and there will only be one year of IRI workbook preparation before Class 3. This means that the test of the first 3 (or 1) IRI levels will not be tests of the program that will eventually be put into effect. In other words, what is tested at the end of year one will not be the program that will be institutionalized after a few years, and the test will not provide an uncontaminated comparison of old and new programs.

3 Test It is difficult to find a Level 1 test that will be fair for pre-IRI and post-IRI students. General English curriculum objectives are very limited for Level 1, and they are generally covered by both programs. Children of both programs are likely to do well in such a limited test. Should the test include items that go beyond the normal program and show IRI to advantage? If so, then to be fair, vocabulary items from the normal program that may not have appeared in IRI need also to be included. In other words there is no available set of independent measures that can be drawn upon to prepare a fair test.

4 Broadcasting times The IRI programs start in September to coincide with the time when all NWFP schools are open, but it is the middle of everyone's year. Thus testing in February will not give the results of an entire IRI year to compare with a full year of the existing program. IRI will therefore not show up as well as it should.

5 Morality of administering difficult tests. If the tests are geared to IRI and way over the heads of normal program children, the exercise of testing them is discouraging for them and their teachers. Is this ethical to do?

Despite these reservations, the IMDC/CBB conducted the baseline tests. A visiting IRI consultant worked with the staff to develop an appropriate test, with vocabulary items that were common to both IRI and the normal English program. In February 1993 this test was administered to students in Classes Three, Four and Five³ in the 30 selected sample schools. Students obtained surprisingly high

³Class Five in all the Bannu schools could not be tested because they were already on holiday.

marks considering how poor their English course appeared in actual classroom use. The average scores were 58 percent correct for Class Three, 65 percent correct for Class Four and 71 percent correct for Class Five.

Already at the time of testing in February, it became clear that the test had a number of flaws that may have accounted for the high scores. First, some of the vocabulary words such as bus, ball, cycle, glass, and baby, are words that are commonly used in Urdu and Pashto. Then others had ambiguous pictures such as an item where the child must identify the person pointing and in three of the four pictures a person appears to be pointing. Finally, there were some questions where the child should identify a verb structure, but could answer the question correctly by knowing instead the noun in the sentence--one of the English words used in Urdu or Pashto.

In September supervisors in all the districts were trained, and IRI began in the 300 selected experimental schools. When the staff returned for testing in late November, they found that the local supervisors, for arbitrary reasons,⁴ had switched the IRI program away from twenty-one of the pretest schools, and therefore, only nine of the test schools were left having the IRI treatment. This virtually eliminated the usefulness of the pretest.

Given the basic flaws in the design, the difficulty in finding an appropriate test, the unreliable pretest results, and the changes in the sample, it seemed better to drop the attempt to do this kind of evaluation. The staff continued with periodic criterion-referenced testing to monitor the results of the program.

⁴The IRI program was highly desirable and the supervisors may have switched the program to favored schools, or to schools where they thought a teacher's better level of English would improve the program's chances of success.

ANNEX E

ENGLISH TEST I AND ITEM ANALYSIS

ITEM ANALYSIS ON ENGLISH TEST I

ITEMS	%CORRECT	
Q1	92	
Q2	89	
Q3	92	
Q4	94	Listening 87%
Q5	87	
Q6	86	
Q7	84	
Q8	69	
Q9	96	
Q10	85	
Q11	90	
Q12	80	
Q13	82	Reading 84%
Q14	81	
Q15	86	
Q16	91	
Q17	78	
Q18	63	Writing 82%
Q19	88	
Q20	91	
Total	85	

ENGLISH IN ACTION

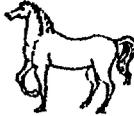
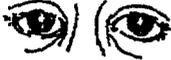
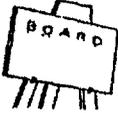
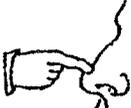
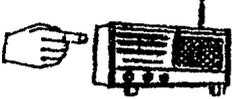
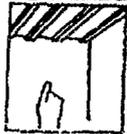
TEST 318

SERIAL NO. _____ TEACHER'S CODE _____

SCHOOL'S NAME _____

TEACHER'S NAME _____ MARKS OBTAINED _____

Circle the picture that matches what the radio says

1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Read, and match with the picture

11.

I am standing

I am sitting



12.

Touch your nose

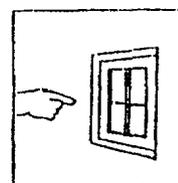
Touch your head



13.

Point to the window

Point to the floor



14.

He is a boy

She is a girl



15.

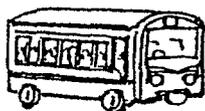
That is a hen

That is a duck



Write the beginning letters in the blanks

16



us :

17



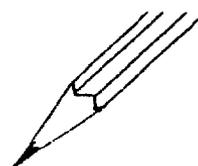
and

18



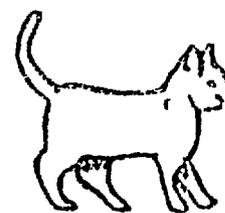
ater

19



encil

20



at