

**FORMATIVE EVALUATION STUDY OF PAKKI MATERIALS
DEVELOPED BY THE IMDC**

INTERIM REPORT

MARCH 1994

DIRECTORATE OF PRIMARY EDUCATION

NWFP

**INTERIM REPORT
FORMATIVE EVALUATION STUDY OF PAKKI MATERIALS
DEVELOPED BY THE IMDC**

DIRECTORATE OF PRIMARY EDUCATION

NWFP

EXECUTIVE SUMMARY

BACKGROUND

This evaluation of Pakki (Class One) materials is the second in a series of formative evaluation studies conducted by the Instructional Materials Development Cell (IMDC) of the Directorate of Primary Education. The purpose of the first study, an evaluation of Kachi (kindergarten) materials (see separate report), was to validate the effectiveness of the approaches and methods used in the new materials developed by the Cell for primary school children in NWFP. Specifically the purpose was to determine whether lessons patterned on "effective teaching practices"¹ would prove easy to use, engage the children and produce high achievement results. The results of the Kachi evaluation met or exceeded standards set by the developers and in most cases were higher. In addition, this phase pointed out the important point that if new materials were to be used effectively, a "slumbering" system of supervisors and teachers would have to be activated and held accountable for student learning. These findings are applicable to the Pakki materials which, in method and approach, are an extension of the Kachi materials.

The overall purpose of the present study has been to conduct a formative evaluation of three sets of instructional materials developed by the IMDC for Pakki classes. The materials include:

- An integrated textbook for Urdu language development that contains practice in reading, writing, and comprehension; the content includes science, social studies and subjects of national interest;
- An integrated textbook for Pashto language development, similar to the Urdu book, that contains practice in reading, writing, and comprehension; the content includes science, social studies and subjects of national interest

¹The "effective teaching practices" were identified from a sample of Pakistani teachers whose students had higher achievement results than students of teachers who did not use them, according to a BRIDGES study.

- A math book conforming with revised national objectives for Class One math
- Each of these sets of materials has an annotated teachers' guide that shows how to teach each lesson. These guides have also been evaluated in this study.
- The specific objectives of the evaluation study were:
 - to test whether the materials (textbooks and annotated teachers' guides) developed for the Pakki class in Urdu, Pashto, and math met the program objectives intended by the writers (see below);
 - to test whether the materials met specified standards in all the normal schooling environments of NWFP (urban/rural, male/female, large/small, multiclass/ singleclass) or with teachers having a variety of characteristics (differing academic or professional trainings, length of service, mother tongues, etc.)
 - to suggest actions, materials, or training that would support better use of the new materials
- The program objectives set for the three sets of materials were to produce materials that:
 - effectively teach the skills of literacy (reading, writing and comprehension) and numeracy;
 - are interesting to children;
 - help the teacher teach effectively and are easy to use

Indicators of success were set for all the program and study objectives. The materials' developers were to use the information collected during the study to take remedial actions either to revise their materials or improve training programs.

Though the main purpose of this evaluation was to test materials, a study of this kind cannot assess materials independent of the context in which they are lodged. In reality, 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.

EVALUATION STUDY

Pakki materials were distributed to classes which had completed the Kachi books. Because of delays in producing the second volume of Kachi language books, and later Pakki books, the districts ready to receive initial Pakki books constituted a larger group than the three experimental districts that started with Kachi books. Altogether Pakki materials were distributed to schools in 13 tehsils of 8 districts of NWFP in September 1994. Altogether 437 classes used math, 242 used Urdu and 133 used Pashto materials. All these classes were given Pakki Test 1 in the appropriate subject in November/December 1993, but only a few in each district were tested by the IMDC staff. For

reliability, the results below include only the scores from the IMDC-tested classes. The tests were designed to cover the main instructional objectives taught during the trial. They were also designed to drive teaching by covering concepts teachers might have ignored or found difficult to teach such as place value in math, or selecting a main idea from a paragraph of prose in language.

FINDINGS

Math materials. The IMDC staff tested more than 5000 students in 264 Pakki classes where new math books and teachers' guides had been used. The average score was 85 percent correct responses on the tested items. Over three-quarters (80 percent) of the classes met the IMDC standard of 75 percent or more correct on the tested items, well above the majority required. Classes taught by female teachers had a higher average (90 percent) than those taught by males (84 percent). There was virtually no difference between urban (85 percent) and rural classes (86 percent), single (85 percent) and multiclassrooms (86 percent), and between classes where teachers taught fewer than 40 students (86 percent) and classes where they taught over 40 students (84 percent).

Urdu materials. IMDC tested 3000 students in 138 Pakki classes in which the integrated Urdu text and teachers' guide had been used. The average score was 76 percent correct responses. Only half of the classes met the standard of 75 percent or more correct. Students of female teachers scored considerably higher (88 percent) compared with those of male teachers (72 percent). Urban classes obtained lower scores (71 percent) than rural classes (77 percent), and single level classes (72 percent) lower than multiclassrooms (77 percent). Achievement was the same in classes where teachers taught fewer than 40 students (75 percent) and where they taught more than 40 students (76 percent).

Pashto materials. The IMDC tested Pashto materials in 51 Pakki classes where the new integrated Pashto students' book and teachers' book had been used. Over 1000 students took the test. The average score obtained was 84 percent correct responses. More than three-quarters (82 percent) of the classes met the standard of 75 percent or more correct responses. Students of female teachers averaged a higher 89 percent compared to 82 percent for students of male teachers. Class scores were virtually the same in urban (85 percent) and rural schools (84 percent). They were slightly lower in single (83 percent) as compared with multiclassrooms (87 percent), but higher in classes where teachers taught fewer than 20 students (85 percent) than where they taught more than 60 students (74 percent). (Note: the IMDC did not test any of the middle size classes)

CONCLUSIONS

The results, with the exception of Urdu in boys' schools, were generally higher in the Pakki class than they had been in the Kachi classes during the first introduction of materials. This may result from teachers becoming used to the new methods or because the supervisory and training systems are working better at reinforcing what happens in the classroom.

Generally speaking the methods and approaches used in the new materials have been validated again, even on the Urdu side where girls' classes have produced very high results. To anyone observing

children taking the achievement tests it was clear that when teachers conscientiously taught according to the instructions in the materials, children performed well in all subjects. They could read and write simple sentences with understanding, and were able to extract a main idea from a short paragraph. In math they could perform, among other activities, simple addition and subtraction and represent place value on a picture of an abacus (see ANNEX A for test examples).

Success of program objectives. The materials substantially though not completely met the standards set by the developers. On the three program objectives they were assessed as follows:

Teaching literacy and numeracy skills: In Pashto and Math, achievement results were high, more than meeting IMDC standards. Urdu showed poorer results, falling just below the standard. Because three-quarters of the girls' classes surpassed the standard, however, compared with only 43 percent of boys, the problem seems not so much a problem with the materials as with something in the environment of boys' schools. Unfortunately the data do not provide a definitive answer for what the problem is, though several hypotheses are suggested in the body of the report.

Attractiveness of the materials. On all counts of being interesting to children, relevant to their experiences, suitable for their ages, etc. teachers gave the materials in all subject matters a high rate of approval. The responses exceeded the majority of positive answers required to meet the developers' standards.

Helping the teacher teach effectively. Most of the indicators that the materials helped the teacher teach effectively were met: a) The majority of the teachers responded positively to questions about the usefulness of the texts and annotated guides. b) Children in the main Pakki learning contexts (urban and rural, large classes--over 60 students, and small classes--under 20 students, and single and multiclassses) performed with differences but not ones that could be reliably tied to some aspect of the context. Pashto results were poorer in very large classes and Urdu was poorer in boys' schools and single classes. c) With the exception of male Urdu teachers, students of teachers with varying characteristics of training, experience, and languages also performed with no differences that could be tied reliably to their characteristics.

In the lowest subject Urdu, the proto-typically poor teacher in the results was a male teacher with high academic qualifications who was either inexperienced (up to 3 years in service) or very experienced (16 years or more service) teaching in an urban single-level class. The developers recognized this prototypical teacher from their experiences in DIKhan, Haripur, Mansehra, Abbottabad, Nowshera, Kohat and Peshawar where Urdu is the language of instruction. He was young, a political appointee with teaching as a last employment resort who spent more of his time looking after his personal business than attending class. Or he was older and coming close to retirement. Worn out from his years of teaching, he was unwilling to accept new ways of teaching that required his active participation.

RECOMMENDATIONS.

Support systems. What was true for the Kachi classes is also true for the Pakki classes,

materials of any kind require strong support systems to ensure that they are used effectively. Without continuous supervision and monitoring of results, and holding appropriate supervisors and educators responsible for student learning, the system will return to its previous state of indifference. If that happens no materials, however excellent, will have much impact. Observation from the field reinforces previous findings that considerable effort needs to be expended in building support systems.

Urdu. Borderline results in Urdu now for the second year of fieldtesting requires a closer look at why the results are not as high as one would like on the male side (see discussion in the main report). Some remedial actions that can be suggested for the short and long term are the following:

- Provide more training to male teachers in Urdu subject content as well as in teaching language to small children (training units are already available in the IMDC for this purpose)
- Provide more means of self-instruction in Urdu to students. IMDC supplementary reading materials (ready but not yet distributed to schools) which contain simple passages that children can read on their own may be helpful. However, they assume at least a minimum ability to read.
- The province should work toward staffing the lower classes of primary schools with female teachers. The results of the testing in all subjects show that the students of female teachers perform better at these grades. Some countries segregate classes up to Class Three in their own facilities. A less expensive option is to provide a shift for these children in schools that are already functioning.
- Interactive Radio Instruction in Urdu could be used to supplement current Urdu lessons or substitute altogether for them. By this means there would be greater uniformity to the instruction provided to children in this subject matter.
- Again, support systems that pay close attention to what is happening in classrooms are important. Teachers and supervisors need to be held accountable when students perform poorly, such as going again for retraining.

TABLE OF CONTENTS

EXECUTIVE SUMMARY.....

1 BACKGROUND.....

1.1 **The problem.....**

1.2 **Creation of the IMDC.....**

1.3 **Background studies.....**

2 EVALUATION STUDY.....

2.1 **Introduction.....**

2.2 **Purposes.....**

2.3 **Program objectives.....**

2.4 **Indicators of success.....**

2.5 **Sample.....**

2.6 **Proformas and procedures.....**

3 FINDINGS.....

3.1 **Extenuating circumstances in the field.....**

3.2 **Meeting achievement standards.....**

3.3 **Subject results.....**

3.4 **Class characteristics and results.....**

3.5 **Teacher characteristics and results.....**

3.6	Teacher ratings of the new materials.....
3.7	Teacher comparisons of new and old materials.....
3.8	Discussion of findings.....
4	CONCLUSIONS.....
4.1	General observations.....
4.2	Success of program objectives.....
5	RECOMMENDATIONS.....
	ANNEX A: Pakki Urdu, Math, Science tests and test item scores...

INTERIM REPORT
FORMATIVE EVALUATION STUDY OF PAKKI MATERIALS
DEVELOPED BY THE IMDC

DIRECTORATE OF PRIMARY EDUCATION

NWFP

1 BACKGROUND

1.1 The problem. Many educators, teachers, students and parents have complained that existing instructional materials in Pakistani schools are not effective means of teaching literacy and numeracy to primary school children. They complain that the content of the books is sometimes too difficult for the teacher to teach or for students to understand at their levels of cognitive development. Subject content matter is often irrelevant to the children's lives and experiences, and is therefore uninteresting to both teachers and students. The materials may be unclear, or even confusing to young children. In most cases, they are so dependent on a teacher's direction that they cannot be used alone by students when a teacher is called away to teach another class, as is the case in the multi-classes found in most rural schools. The teachers, who are poorly trained, may be weak in the content matter or may not know techniques for teaching effectively. The training they receive assumes instruction in classrooms with a single class level and a modest number (40) of students. The materials themselves are not organized systematically according to accepted standards for curriculum development. There is no transition between one book and another, as for example between the Pakki (Class One) alphabet books and reading books. Teachers have not been trained specifically in how to use the books. Because of this and inadequacies in the books themselves, teachers fall back on rote memory of the subject matter to simplify the task of instruction.

Because of these and other difficulties, a decision was taken in the Directorate of Primary Education to produce new instructional materials for the primary level that addressed the major problems discussed above.

1.2 Creation of the IMDC. In October 1991, an Instructional Materials Development Cell (IMDC) was established in the Directorate of Primary Education in Peshawar. The staff were recruited from promising young primary school teachers with high (BA and BSc) academic qualifications and at least 5 years teaching experience who had applied to participate in a preliminary workshop on instructional materials development. Out of 50 teacher candidates in the workshop, the 16 (later 4 dropped out) who showed most aptitude were chosen to become subject specialists and researchers for the Cell. Primary school teachers were chosen because of their experience working with primary

age children in conditions that exist in the schools of NWFP.

The aim of the Cell is to develop the instructional materials and associated annotated teachers' guides required for the main academic subjects in Kachi (kindergarten) through Class Five. The staff of the IMDC write the materials, which are then checked for accuracy by experts in the subject matter, informally tested in local classrooms by the writers, and when perfected to the degree possible by the writers themselves, a formative evaluation such as the present study is conducted in experimental schools, and the materials are again modified based on the recommendations of the study. The materials are prepared in yearly installments, one class at a time. At the time of this report Kachi, Pakki and Class Two materials had been developed. Kachi materials had been field-tested and this report summarizes the results of the first months of field-testing of the Pakki materials.

1.3 Background studies. To serve as a basis for developing primary programs, studies were conducted in 1991/92 in NWFP schools. Among them were studies of the range of conditions in early primary (Kachi) and multi-grade classes which detailed the types of schooling environments where instructional materials would be used. Both types of classes showed large variations in the numbers and ages of students as well as deficiencies in instructional aids, materials and classroom facilities. Particularly significant in terms of the poor quality of instruction were the heavily teacher-dependent materials which made it difficult for a teacher to assign independent work to students, and causing them to fall back on memorization as an easy way to solve the problem. Teachers also seemed to lack academic qualifications as evidenced in their low graduation (division) scores and the numerous mistakes observers noted them making in the classroom. It was concluded that the education program could be improved in both types of classes with instructional materials that addressed the most important needs of these classrooms, specifically, improving student skills, and making the materials both user-friendly and less teacher-dependent.

The studies made clear that improving student learning in this and other early primary classes would increase the learning of the bulk of primary children. Indeed if learning were significant in these early classes, a full school year could be added to the academic program at an early age before students, especially girls, were likely to drop out without having consolidated basic skills. Moreover if the materials were interesting to children, it was possible that schooling might be attractive enough so children would continue on to higher classes, rather than dropping out in large numbers between Kachi and Pakki Classes as statistics of the time seemed to indicate they were doing.

As a result of helping to conduct these studies many district staff became directly acquainted with problems in primary classes while at the same time learning basic skills of data collection and classroom observation that would be useful later on when the new programs were introduced.

Based on the findings of the studies, the IMDC staff developed materials first for Kachi and later for Pakki Class. The materials were intended to replace existing Pakki materials which had informally been used for both Kachi and Pakki classes. The new materials were based on a phonetic approach that allowed children to decode words and simple sentences by the end of the Kachi year. thus providing the transition that had previously been missing in the books, between alphabet letters and reading passages. Multi-class conditions were used as the basis for developing the materials on the

assumption that the majority of NWFP schools contained multi-classes and that it would be easier to adapt materials suited to multi-class environments to single class environments rather than the other way around.

2 EVALUATION STUDY

2.1 Introduction. The evaluation of Pakki materials is the second in a series of formative evaluation studies conducted by the IMDC. The first, an evaluation of Kachi materials (see separate report) was conducted in 3 phases. In the first phase, Kachi materials were introduced into a small sample of schools. A variety of data--from observations, interviews, and achievement tests were collected and intensively analyzed. The purpose of this phase was to validate the basic structure and approaches used in the materials, that is, to determine whether lessons patterned on "effective teaching practices"² would prove easy to use, engage the children and produce high achievement results.

At the end of this phase of the Kachi evaluation, the results met or exceeded standards³ set by the developers. Thus the basic approaches and structures used in the materials seemed to be validated. This phase pointed out the important difficulty that if the new materials were to be used effectively, a "slumbering" system of supervisors and teachers would have to be activated and held accountable for student learning. The IMDC staff could motivate teachers to use materials during the trial, but they could not serve in this capacity forever. Therefore the next phases of the Kachi evaluation paid close attention to making the supervisory staff accountable for student performance.

The Pakki materials in method and approach are an extension of the Kachi materials. Therefore, because the methods and approaches have been validated in field-tests of the Kachi materials, the Pakki evaluation follows a simplified design. The evaluation continues to check student learning results coming from use of the new materials, and to assess which schooling environments and personnel, if any, have more difficulty using the materials. Close contact with the field during assessments, makes IMDC staff aware of how well or poorly their materials are working and where they need revision. They also observe in the field where other types of teacher support are needed. A number of training units and supplementary materials have been developed out of this experience.

2.2 Purposes. The overall purpose of the present study has been to conduct a formative evaluation of three sets of instructional materials developed by the IMDC for Pakki classes. The materials include:

²The "effective teaching practices" were ones used by a sample of Pakistani teachers whose students had higher achievement results than students of teachers who did not use them, according to a BRIDGES study. They consisted of a series of steps stressing practice: review of previous material, focusing on a topic or objective, presentation of new material with clear examples, guided practice, independent practice and homework. IMDC lessons in texts and teachers' guides are organized around these steps.

³ Because the aim was to develop a teacher performance score, the original standard set by IMDC developers which related to the individual scores of children was changed to a class score. The new standard was that the majority of the classes in a district would have 75 percent or higher correct answers.

- An integrated textbook for Urdu language development that contains practice in reading, writing, and comprehension; the content includes science, social studies and subjects of national interests
- An integrated textbook for Pashto language development, similar to the Urdu book, that contains practice in reading, writing, and comprehension; the content includes science, social studies and subjects of national interest
- A math book conforming with revised national objectives for Class One math

Each of these sets of materials has an annotated teachers' guide that shows how to teach each lesson. These guides have also been evaluated in this study.

The specific purposes of the evaluation study were:

- to test whether the materials (textbooks and annotated teachers' guides) developed for the Pakki class in Urdu, Pashto, and math met the program objectives intended by the writers (see below);
- to test whether the materials met specified standards in all the normal schooling environments of NWFP (urban/rural, male/female, large/small, multiclass/ singleclass) or with teachers having a variety of characteristics (differing academic or professional trainings, length of service, mother tongues, etc.)
- to suggest actions, materials, or training that would support better use of the new materials

2.3 Program objectives. The main program objectives set for the three sets of materials were to produce materials that:

- effectively teach the skills of literacy (reading, writing and comprehension) and numeracy;
- are interesting to children;
- help the teacher teach effectively and are easy to use

The materials' developers were to use the information collected during the study to take remedial actions either through revision of their materials or improved training programs.

Though the main purpose of this evaluation was to test materials, a study of this kind cannot assess materials independent of the context in which they are lodged. In reality, the study tests the 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.

2.4 Indicators of success. The standards/indicators set by the writers to constitute evidence that the program objectives had been met are found below. Anything less than the standard would be cause for concern.

Skills: To indicate that the materials effectively produce the skills of literacy (reading, writing, and comprehension) and numeracy, it was expected that at least half of the classes tested by IMDC staff⁴ would score 75 percent correct answers or higher on a test of the content covered during the trial. The test used in the evaluation was similar to that found in the major review section at the end of the first half year of the materials.

Attractiveness: To indicate that materials were attractive to children, a majority of teachers would rate the materials relevant to the experiences and age levels of children, and that the materials seemed interesting to the children.

Teacher effectiveness: To indicate that the materials helped the teacher teach effectively and were easy to use: a) a majority of teachers were expected to respond positively to a set of questions asking about the usefulness of the student texts and guides, and b) over half the classes would score 75 percent correct answers or higher on an achievement test of the content covered during the trial, and c) the class scores in the Pakki learning contexts (male and female, urban and rural, large and small classes, single and multi-classes) and with varying types of teachers would reach 75 percent or more on an achievement test.

2.5 Sample. Pakki materials were provided to classes which had completed the Kachi books. Because of delays in producing the second volume of Kachi language books, and later Pakki books, the districts ready to receive initial Pakki books constituted a larger group than the three experimental districts that started with the Kachi books. These earlier districts had been chosen on the basis of several criteria: the medium of their instruction, the enthusiasm and willingness of their staff to participate in the testing, and the absence of programs such as PEP II, and the Pak-German project that might interfere with the introduction of the new materials. By the time of the Pakki evaluations, PEP II had been disbanded and the LCs that were part of that project had been accommodated within the district supervisory staff. The districts and sub-districts that received the new Pakki materials were thus a larger and less homogeneous group than those which received the initial introductions of Kachi materials. The schools also had more experience with new materials and the supervisors were becoming used to the system of supervision that held everyone accountable for their part in student learning results.

⁴The number of schools which by the time of the Pakki evaluation were receiving new materials had increased to over 775, beyond the capacity of IMDC to test each one. The bulk of the schools were, therefore, tested by district supervisors, with a more limited sample spot-checked by IMDC staff.

Table 1 shows the full sample of schools in 13 tehsils of 8 districts that started using Pakki materials in September 1994. Altogether 437 used math, 242 used Urdu and 133 used Pashto materials. All these classes were given Test 1 in the appropriate subject(s) in November/December 1993, but only a few in each district were tested by the IMDC staff. The rest were tested by the supervisory staff. A sample of 20 children (unless the enrollment was smaller than 20) were selected from each class for testing.

Table 1: Schools using Pakki materials as of September 1994

Tehsil	Schools	Math/Sc.	Urdu	Pashto
Bannu	Boys	42	42	--
	Girls	20	20	--
Battagram	Boys	9	9	--
	Girls	2	2	--
Chitral*	Boys	21	21	--
	Girls	8	8	--
Chitral* (Mastooj)	Boys	12	12	--
	Girls	--	--	--
Dir (Timargara)	Boys	48	--	48
	Girls	8	--	8
Kohat	Boys	24	24	--
	Girls	8	8	--
Kohat (Hangu)	Boys	6	6	--
	Girls	4	4	--
Lakki	Boys	36	36	--
	Girls	4	4	--
Mansehra	Boys	38	38	--
	Girls	8	8	--
Mardan	Boys	36	--	--
	Girls	14	--	--
Mardan (Takht Bai)	Boys	8	--	--
	Girls	4	--	--
Swat (Saidu Sh.)	Boys	43	--	43
	Girls	14	--	14
Swat (Alpuri)	Boys	12	--	12
	Girls	8	--	8

Totals	437	242	133
--------	-----	-----	-----

*Chitral results do not appear in later tables because, being inaccessible most of the year, they test their own materials.

2.6 Proformas and procedures. There were 3 instruments used to evaluate the introduction of the Pakki materials: achievement tests, an interview/teacher characteristics form and an observation form.

Achievement tests (Forms 25, 26 and 27 for Math, Urdu and Pashto) were administered in late November or early December after teachers had been using the material for approximately two and a half months (the time varied depending on when the classes received the materials).⁵ The test was based on review tests found in the materials but IMDC staff rearranged pictures/figures and made other changes to avoid possible memorization of the tests. The tests covered the main instructional points during the trial period. Part of the purpose of the test was to motivate teachers to teach difficult items that they might otherwise decide to ignore. In language, for example, children were asked to identify the main idea of a short passage which they read themselves, and in math several questions were concerned with the difficult concept of place value. A maximum of 20 children were tested in each class.

A teachers' characteristics and interview form (Form 1) was filled in at the same time as the test was administered. The information from this form was used to determine if there were any consistent differences in the way teachers with a variety of characteristics used the materials and whether the materials were interesting to children and useful for the teacher.

Classroom **observation proformas (Form 3)** were left with supervisors to be filled in during one lesson of each unit in each school. The information in these proformas was used exhaustively in the initial Kachi phase, but in later phases was not entered into the computer when the sample became so large as to make the data unwieldy. At that point Form 3 was simplified to become a routine monitoring instrument that could be used in any class for observing teacher performance.⁶

3 FINDINGS

The findings presented here are from the first cycle of Pakki testing, using Test 1 for each of three subject materials: Math, Urdu and Pashto, at approximately the mid-point of the materials.

3.1 Extenuating circumstances in the field. As in the Kachi evaluation, a number of extenuating

⁵ Test 2 is given in late May before the summer holidays.

⁶ Modifications in the proforma for this purpose were a consequence of the recommendation, shown to be essential in the Kachi evaluation, that a routine monitoring system be developed to support teachers in all schools.

circumstances in the field may have affected the results of the study. They include:

- **Late starts** to the distribution of materials in all districts and special problems within localized areas caused some teachers to review Kachi materials again and again, and others to revert to older Textbook Board books which contradict some of the methods used in the IMDC materials.
- **Frequent transfer** of teachers and supervisory upset the process of introducing the materials in some places since incoming staff were not usually trained in the new materials.
- **Frequent casual and other leaves of teachers** meant sometimes that the children had not been taught as much of the materials as expected. When this occurred, the IMDC staff only tested the children on the parts of the tests that had been covered.

One of the priorities identified in the Kachi study was the need to develop support systems for new materials. Part of this process required locating problems as quickly as possible and remedying them on the spot so acceptable student learning could be produced. This approach, not appropriate in a summative design, is consistent with formative evaluation which gathers information so developers can improve their materials and programs. After the earlier Kachi phases validated the method and approach, the Pakki phase has focused more attention on building the support system for the materials, by involving district personnel in training teachers and assessing student learning. To remove as much bias as possible from this report, however, scores below are only included when IMDC staff have administered the test.⁷

3.2 Meeting achievement standards. Table 2 shows the results of the Pakki achievement tests in Math, Urdu, and Pashto.⁸ The table shows the number of classes meeting or not meeting the standard of 75 percent or higher correct responses on the achievement tests.

Table 2: Number of class scores meeting IMDC standards (only IMDC tested classes reported) --December 1994**

Tehsil	Classes	Math		Urdu		Pashto	
		>75%	<75%	>75%	<75%	>75%	<75%
Bannu	Boys	28	9	15	28	--	--
	Girls	19	1	10	7	--	--

⁷This is not any more so much because supervisors inflate the scores, but rather to give a higher level of confidence in the results. Comparison between the scores of supervisor-administered and IMDC-administered tests now show fewer marked differences than in earlier efforts to involve the supervisors in testing.

⁸Note that some girls' classes did not receive their books on time. This was because the books were delivered to the male SDEO office and were not transferred to the female office for distribution to schools.

Battagram	Boys	1	0	1	0	--	--
	Girls	--*	---	--*	---	--	--
Dir (Timargara)	Boys	9	0	--	--	7	2
	Girls	6	0	--	--	4	0
Kohat	Boys	10	10	11	14	--	--
	Girls	8	1	6	2	--	--
Kohat (Hangu)	Boys	4	2	0	6	--	--
	Girls	2	2	4	0	--	--
Lakki	Boys	22	9	17	12	--	--
	Girls	3	0	4	0	--	--
Mansehra	Boys	11	3	1	0	--	--
	Girls	--*	---	--*	---	--	--
Mardan	Boys	28	5	--	--	--	--
	Girls	19	0	--	--	--	--
Mardan (Takht Bai)	Boys	5	4	--	--	--	--
	Girls	4	0	--	--	--	--
Swat (Saidu Sh.)	Boys	14	3	--	--	13	3
	Girls	16	1	--	--	14	3
Swat (Alpuri)	Boys	4	1	--	--	4	1
	Girls	---	---	--	--	---	---
Totals		213	51	69	69	42	9
		81%	19%	50%	50%	82%	18%

*The classes in this district/tehsil received their books late.

**The columns should read "Equal to or more than 75%" and "Less than 75%"

Note: Chitral test their own materials so do not appear here.

The table shows that results in math and Pashto surpassed the standard of a majority having 75 percent or more correct answers, but were just below the standard in Urdu. In the Kachi trials, the Urdu materials also showed comparatively poorer results than the other two subjects. While these results still surpass by a great deal the standards (33 percent) set for subjects taught through Textbook Board materials, the lower scores of Urdu are still cause for concern (see discussion below).

3.3 Subject results. Results below are disaggregated by subject.

3.3.1 Math materials. The IMDC staff tested students in 264 Pakki classes where the new students' math books and teachers' guides had been used. More than 5000 students took the IMDC

test.

- The average score was 85% correct responses on the test.
- Over three-quarters (80 percent) of the class scores met the standard of 75 percent or more correct responses.
- Classes taught by female teachers had higher scores (94 percent met the standard) than those of male teachers (80 percent). Female students averaged 90 percent to the males' average of 84 percent.
- The score was virtually the same in urban (85 percent) and rural schools (86 percent).
- Achievement was almost the same in single (85 percent) and multiclassrooms (86 percent).
- Achievement was the same in classes where teachers taught fewer than 40 students (86 percent) as in classes of over 40 students (84 percent).

3.3.2 Urdu materials. The IMDC tested 138 Pakki classes in which the integrated Urdu text and teachers' guide had been used. Almost 3000 students took the achievement test.

- The average score was 76 percent correct responses.
- Only half of the classes met the standard of 75 percent or more correct.
- Students of female teachers scored considerably higher (73 percent of classes met the standard) than students of male teachers (43 percent). The female average score was 88 percent compared with 72 percent for the males.
- The average was lower in urban (71 percent) than in rural schools (77 percent).
- Achievement was lower in single (72 percent) as compared to multiclassrooms (77 percent).
- Achievement was the same in classes where teachers taught fewer than 40 students (75 percent) and where they taught more than 40 students (76 percent).

Students had particular problems with test items where sentence structures in Urdu differ from what for most is their mother tongue, Pashto. In Urdu, gender (masculine and feminine) requires changes in verb endings, in adjectives and in prepositions, and these tend to be different from Pashto. No effort has been made before the new materials to emphasize practice of items where these differences exist.

3.3.3 **Pashto materials.** The IMDC tested Pashto materials in 51 Pakki classes where the new integrated Pashto students' book and teachers' book had been used. Over 1000 students took the test.

- The average score was 84 percent correct responses.
- More than three-quarters (82 percent) of the classes met the standard of 75 percent or more correct responses.
- Students of female teachers scored slightly higher (86 percent met the standard) than students of male teachers (80 percent). Female students averaged 89 percent to 82 percent for males.
- The score was virtually the same in urban (85 percent) and rural schools (84 percent).
- Achievement was slightly lower in single (83 percent) as compared with multiclassses (87 percent).
- Achievement was higher in classes where teachers taught fewer than 20 students (85 percent) than where they taught more than 60 students (74 percent). (Note: the IMDC did not test any of the middle size classes)

3.4 Class characteristics and results. Table 3 shows the results by subject and by various class characteristics. Again math and Pashto had high scores irrespective of the class characteristic. The one exception was a lower average class score for Pashto in classes of more than 60 students. Urdu had three groups falling below 75 percent correct: male, urban and single classes.

Table 6: Class characteristics and student achievement (first Pakki evaluation cycle)

Characteristics	Average score of students (percent correct)		
	U	M	P
Provincial average	76	85	84
Gender			
Male	72	84	82
Female	88	90	89
Location			
Urban	71	85	85
Rural	77	86	84
Class composition			
Single	72	85	83
Multi-class	77	86	87
Size of class			
1 to 20	75	86	85

21 to 40	--	96*	--
41 to 60	--	--	--
60+	76	84	74

*There were too few cases to be reliable.

3.5 Teacher characteristics and results. One of the study questions asked whether the materials could be used effectively with teachers having different characteristics and in different environments. Already we have seen that in all subjects, classes taught by female teachers have higher average scores than those taught by male teachers. Sometimes, as in Urdu the difference is dramatic.

Table 7 shows the average score for students whose teachers have different characteristics. Again, math and Pashto have high scores in all categories of teacher (when single cases are removed) except for the lower score in math for middle pass teachers. Because of its poor overall scores, Urdu has lower than standard scores in more than half the categories of teachers. It is difficult on the surface, however, to see a pattern in most of the lower scores for this subject that relates systematically to the teacher characteristic. All the results seem to be consistently affected by teacher gender, and to lesser degree, a teacher's years of experience. Teachers with many years of experience tend to have students with lower scores, a factor observed in the field as greater resistance to new ways of teaching.

Table 7: Teacher characteristics and student achievement

Characteristics	Average class score (% correct)		
	U	M	P
Provincial average	76	85	84
Gender			
Male	72	84	82
Female	88	90	89
Academic			
Middle	84	68	28*
Matric	75	87	80
FA	75	85	89
BA/BSc	73	84	88
MA/MSc	48*	94	93
Division**			
I	65	81	83*
II	76	87	84
III	72	82	87
Profess. Training			
Untrained	88*	93	89
PTC	77	86	83
CT	71	82	85
BEd	66	92	91*
MEd	82*	96*	96*
Teaching experience			

Up to 3	73	88	88
4 to 10	78	85	85
11 to 15	67	89	90
16 and more	74	83	77
Mother tongue			
Urdu	63*	89	95*
Pashto	72	86	85
Hindko	86	81	--
Kohistani	90*	90*	--
Saraiki	31*	51*	--

*These cases are too few (3 or less) to be reliable.

**Division refers to the graduating score with "I" the highest.

3.6 Teacher ratings of the materials. Teachers were asked to rate specified characteristics of the materials, on a scale from one to three, where "3" was equal to "good". Table 8 shows these ratings with regard to the student books. Urdu books were rated lower than the math and Pashto books.

Table 8: Teacher ratings* of Urdu, Math and Pashto student books

Characteristic	% Rating good		
	U	M	P
Suitable for child's age	79	92	98
Relevant to child's experience	78	93	98
Interesting to children	79	92	98
Help children learn	79	92	98

* There were 3 ratings: 1=poor 2=fair 3=good

Table 9 shows the percent of teachers rating the teachers' guides as "good". The overall ratings were high, but again the Urdu is rated lower than the other two guides. In another question asking about the use of the guides, 97 percent of Urdu teachers said they only used the guides some days, while 99 percent of the math teachers and 100 percent of the Pashto teachers said they used the guides every day. This is a startling difference that may suggest one of the reasons why the Urdu scores are comparatively lower.

Table 9: Teacher ratings* of Urdu, Math and Pashto teacher guides

Characteristic	% Rating good		
	U	M	P
Directions are clear	80	93	99

Relevant to child's experience	78	92	99
Easy to teach with this method	78	92	99
Useful to the teacher	78	93	99

* There were 3 ratings: 1=poor 2=fair 3=good

3.7 Teacher comparisons of new and old materials. Teachers were asked to compare specified characteristics of the new materials with materials they had used before for Pakki classes. They overwhelmingly reported that the new materials were better on all dimensions compared to the materials they had been using before. Table 10 shows these comparisons.

Table 10: Teacher comparisons* of new Urdu, Math and Pashto student texts compared with old texts

Issue	% Rating "better than old"*		
	U	M	P
Systematic organization	82	94	99
Interesting to the child	82	93	99
Related to child's experience	82	94	99
Helping child learn	81	93	99
Easy to teach	82	94	99

* There were 3 comparisons: 1=not as good as old books 2=the same as old books 3=better than old books

3.8 Discussion of the findings. With minor exceptions, results for math and Pashto were high among all Pakki groups and in all schooling environments. The only results that did not meet standards were the Pakki Urdu materials, and there mainly among the students of male teachers. Since Urdu results have been borderline during both Kachi and Pakki trials, it would be helpful to understand why. The answer undoubtedly lies in a combination of factors that are difficult to give causal priority to. A number of explanations have been suggested, but most have some negative evidence that tends to disconfirm them as main factors. The following are some of the proposed explanations:

- "Urdu is a second language and therefore it is more difficult to learn." However, girls achieved at approximately the same high rate in Urdu as they did in Pashto.

- "The new Urdu books are at fault." However, the method and approach in the Urdu books is the same as in the Pashto book, and girls had no trouble achieving at a high level with the same materials.
- "Male teachers in NWFP are more likely to speak Pashto as a mother tongue than female teachers. Because Paktuns are conservative, they are reluctant to allow their women to work, and therefore female teachers are more likely to be Hinko or Saraiki speakers. Both these languages are closer to Urdu and therefore female teachers would be more likely to be better teachers of Urdu." In this study, students of teachers who were Hindko speakers scored higher (86 percent) in Urdu than students of teachers who were Pashto speakers (72 percent). However, more Urdu teachers in the female sample claimed Pashto as a mother tongue (74 percent) than on the male side (70 percent), and more male teachers (26 percent) claimed Hindko as a mother tongue than female teachers (16 percent).
- Female teachers are better with languages than male teachers. However, female PTC students do less well than male PTC students in a Class Five test of Urdu (see the Teacher Content Knowledge Study and subsequent tests of PTC students). Moreover, male teachers tend to be academically more highly qualified than female teachers, a factor that usually produces higher achievement in students.
- Female teachers are more likely to communicate in Urdu with one another at school because of their differences in mother tongue and therefore they are more fluent and thus better teachers. Female schools are more likely to be located in or close to urban areas where it would be more likely to find a heterogenous set of teachers communicating in Urdu or Urdu-like languages. BRIDGES' studies showed that female teachers were more likely to speak Urdu in their classes than male teachers.
- Females are more nurturing and therefore better teachers in the lower grades. Better results were produced in all subjects by females, even though the differences were not as dramatic in Pashto or math as in Urdu. Female teachers may constitute a more motivated group because it is more difficult for them to overcome the social and cultural barriers to becoming teachers.

The explanation for poor male performance in Urdu probably lies in a combination of factors: the fact that men may have a number of distracting economic interests, that they may have less patience with the difficult task of teaching small children a second language, that their style of instruction is more likely to be formal and directive rather than responsive and interactive, and that more of their schools lie in rural areas that are heavily Pashto speaking and therefore where Urdu is more difficult to learn. Another clue may be found in the intriguing piece of information that though almost all teachers said they used the teachers' guide daily for Math and Pashto, almost all the teachers said they only used the teachers' guide sometimes for teaching Urdu. This does not explain poor male performance, but it does suggest a basic difference between Urdu and the other two subjects. Perhaps teachers find it difficult to read Urdu.

4 CONCLUSIONS

The conclusions of the study are the following:

4.1 General observation. The results, with the exception of Urdu in boys' schools, were generally higher in the Pakki class than they were in the Kachi classes in the first introduction of the materials. This may be because teachers are becoming used to the new methods and because the supervisory and training systems are better able to reinforce what happens in the classroom.

Generally speaking the methods and approaches used in the new materials have again been validated, even on the Urdu side where girls' classes have produced very high results. To anyone observing children taking the achievement tests it becomes immediately obvious that when teachers conscientiously teach according to the instructions in the materials, children perform well in all subjects. They read and write simple sentences with understanding, and are able to extract a main idea from a short paragraph. In math they perform, among other activities, simple addition and subtraction and represent place value on a picture of an abacus (see ANNEX A for test examples).

4.2 Success of program objectives. The materials substantially though not completely met the standards set by the developers. On the three program objectives they were assessed as follows:

4.2.1 Teaching literacy and numeracy skills: In Pashto and Math, achievement results were high, more than meeting IMDC standards. Urdu showed poorer results, falling just below the standard. Because three-quarters of the girls' classes surpassed the standard, however, compared with only 43 percent of boys, the problem seems not so much a problem with the materials as with something in the environment of boys' schools. Unfortunately the data do not provide a definitive explanation for the poor performance, though several hypotheses have been suggested above.

4.2.2 Attractiveness of the materials. On all counts of being interesting to children, relevant to their experiences, suitable for their ages, etc. teachers gave the materials in all subject matters a high rate of approval. The responses exceeded the majority of positive answers required to meet the developers' standards.

4.2.3 Helping the teacher teach effectively. Most of the indicators that the materials helped the teacher teach effectively were met: a) The majority of the teachers responded positively to questions about the usefulness of the texts and annotated guides. b) Children in the main Pakki learning contexts (urban and rural, large classes--over 60 students, and small classes--under 20 students, and single and multiclassses) performed with few exceptions to the standard. Pashto results were poorer in very large classes and Urdu was poorer in boys', urban, and single classes. c) With the exception of gender and years of teacher experience (long experience produced lower results), students of teachers with varying characteristics of training, division, and languages also performed with no differences that could be tied reliably to teacher characteristics.

In Urdu, the proto-typically poor teacher in the results was a male teacher with high academic

qualifications who was either inexperienced (up to 3 years in service) or very experienced (16 years or more service) teaching in an urban single-grade class. We have all seen this teacher in DIKhan, Haripur, Mansehra, Abbottabad, Nowshera, Kohat and Peshawar where Urdu is the language of instruction. He is either young, with teaching as a last employment resort, who spends more of his time looking after his personal interests than attending class, or he is older and coming close to retirement. Worn out from his years of teaching, he is unwilling to accept new ways of teaching that require his active participation.

5 RECOMMENDATIONS

5.1 Support systems. What was true for the Kachi classes is also true for the Pakki classes, materials of any kind require strong support systems to ensure that they are used effectively. Without continuous supervision and monitoring of results, and holding appropriate supervisors and educators responsible for student learning, the system will revert to its earlier somnolence. When that happens, no materials however excellent will have much impact. Observation from the field reinforces what was identified as a necessity in the Kachi evaluation, the importance of spending considerable effort in building support systems.

5.2 Urdu. Borderline results in Urdu now for the second year of fieldtesting requires a closer look at why the results are not as high as one would like on the male side. Some remedial actions that can be suggested for the short and long term are the following:

- Provide more training to male teachers in Urdu subject content as well as in teaching language to small children (training units are available in the IMDC for this purpose)
- The province should work toward staffing the lower classes of primary schools with female teachers. Some countries segregate classes up to Class Three in their own facilities. A less expensive option is to provide a shift for these children in schools that are already functioning.
- Interactive Radio Instruction in Urdu could be used to supplement the current lessons or substitute altogether for them. By this means there would be greater uniformity to the instruction provided to children in this subject matter.
- Again, support systems that pay close attention to what is happening in classrooms are a must. There must also be consequences of poor performance such as requiring teachers whose children are not producing adequate achievement results to go again for retraining.

ANNEX A

PAKKI URDU, MATH AND SCIENCE TESTS AND ITEM TEST SCORES

ITEM TEST SCORES

Below are the percent of correct answers out of the possible correct answers on each item of English Test I administered by the IMDC in late November/early December 1993.

URDU ITEM SCORES

ITEM	%
Q1	63
Q2A1	84
Q2A2	80
Q2B1	77
Q2B2	71
Q2C1	62
Q2C2	64
Q2D1	75
Q2D2	72
Q2E1	64
Q2E2	63
Q3A1	83
Q3A2	80
Q3B1	77
Q3B2	74
Q3C1	70
Q3C2	68

Q4A1	94
Q4A2	92
Q4A3	90
Q5A1	82
Q5A2	83
Q5A3	78

Total 76

MATH ITEM SCORES

Q1	90
Q2	89
Q3A1	97
Q3A2	96
Q3A3	94
Q4A1	95
Q4A2	95
Q4A3	82
Q5	90
Q6	86
Q7	95
Q8	94
Q9	96
Q10A1	97
Q10A2	97
Q11A1	77
Q11A2	68

Q11A3	69
Q12	75
Q13A1	72
Q13A2	60
Q13A3	60
Q14A1	92
Q14A2	90
Q14A3	71
Total	85

PASHTO ITEM SCORES

Q1	88
Q2A1	79
Q2A2	82
Q2B1	84
Q2B2	84
Q2C1	82
Q2C2	79
Q2D1	76
Q2D2	79
Q3A1	90
Q3A2	85
Q3B1	93
Q3B2	91
Q3C1	73
Q3C2	74

Q3D1	81
Q3D2	80
Q3E1	72
Q3E2	78
Q4A1	95
Q4A2	95
Q4A3	95
Q5A1	87
Q5A2	88
Q5A3	87
Total	84