

Assessment Affirms Quality Education

# PROVINCIAL ASSESSMENT REPORT 2005



**PUNJAB EDUCATION ASSESSMENT (PEAS)**

**EDUCATION DEPARTMENT,  
GOVT. OF THE PUNJAB**

In Collaboration With

NATIONAL EDUCATION ASSESSMENT SYSTEM (NEAS)  
MINISTRY OF EDUCATION,  
GOVERNMENT OF PAKISTAN



PEAS

## **Acknowledgment**

Punjab Education Assessment System (PEAS) would like to acknowledge with thanks the support it have received from many individuals and institutions in conducting the first round of provincial assessment in 2005. Thanks are particularly due to:

- Special Secretary Schools, Education Department.
- National Education Assessment System (NEAS), Ministry of Education, Islamabad.
- Additional Secretary Donor, Coordination Wing, Education Department.
- Director, Public Instruction Elementary Education.
- Executive District Officer (Education).
- Principals of Teacher Trainer Institutions.

In addition, thank Principals/Head of Schools are also due to the District Education Officers, Staff of Teacher Training Institutions, Test Administrators, Head Teachers, Teachers and specially Students for giving their valuable time to participate in assessing students learning achievements at National and Provincial Level.

## Table of Contents

List of Tables	v
List of Figures	vi
List of Abbreviations	vii
Executive Summary	ix
1 National Assessment	1
2 National Assessment Instruments	3
3 National Assessment Sample	5
4 Understanding NEAS Results	7
5 National Results	9
6 Background and Context Variables in Relation to Students' Achievement Scores	15
7 Constraints and Lessons Learned	29
<b>Annexes</b>	
1 National Test Sample	33
2 Instrument Development	39
3 Competencies and Sample Questions	43
4 Test Administration	53
5 Monitoring National Assessment	55
6 Marking and Coding Assessment Instruments	57
7 Item and Background Data Analyses	59

PRELIMINARY

## List of Tables

Table 1 National Sample for Grade-4 National Assessment 2005	6
Table 2 School Levels in the Sample	6
Table 3 Scaled Mean Score for Language and Maths	9
Table 4 Student Achievement by Location	11
Table 5 Student Achievement by Gender	12
Table 6 Mathematics Results for the Provinces/Areas ...	12
Table 7 Urdu Results for the Provinces/Areas	12
Table 8 Achievement of Sindhi and Urdu Medium Schools in Sindh ...	13
Table 9 School Level and Mathematics Achievement	13
Table 10 School Level and Urdu Achievement	13
Table 11 Student Age and Achievement	16
Table 12 Achievement in Relation to Home Language	16
Table 13 Mathematics Teaching at Home	17
Table 14 Urdu Teaching at Home	17
Table 15 Student Absenteeism	18
Table 16 Father's Education in Rural and Urban Areas	18
Table 17 Father's Occupation in Rural and Urban Areas	19
Table 18 Mother's Education in Rural and Urban Areas	19
Table 19 Teacher Gender and Student Achievement	20
Table 20 Teachers' Academic Qualification and Student Achievement	20
Table 21 Teachers' Highest Level of Professional Training and Student Achievement	20
Table 22 Achievement Scores in Relation to Teaching Experience	21
Table 23 Multi-Grade Teaching and Achievement	21
Table 24 Multi-Grade Teaching and Urdu Teachers	21
Table 25 Multi-Grade Teaching and Mathematics Teachers	22
Table 26 Teacher Assessment of Students through Homework	22
Table 27 Students Report Getting Homework	23
Table 28 Prizes, Incentives and Mathematics Achievement	23
Table 29 Punishment at School and Mathematics Achievement	23
Table 30 Participation in Co-Curricular Activities	23
Table 31 Questions Allowed in Class and Achievement	24
Table 32 Availability and condition of blackboards in schools	24
Table 33 Availability of School Library	25
Table 34 Availability of the Ministry Of Education Curriculum (2002)	25

Table 35 Presence of PTA/SMC in the Sample Schools	26
Table 36 PTA/SMC and Achievement	26
Table 37 Supervisor Visits	28
Table 38 Numbers of Schools (from NEMIS)	33
Table 39 Class 4 Pupils by Province/Area, Gender and R/U (from NEMIS)	34
Table 40 Proposed Sample: by Province/Area, Gender and Rural/Urban	34
Table 41 Achieved Sample by Province/Area	35
Table 42 Provinces/Areas with above-Average Participation of Students	35
Table 43 Provinces/Areas with below-average Participation of Students	36
Table 44 Sampling Outcome: Rural vs Urban	36
Table 45 Number of Master Trainers Trained by NEAS	53

PEAS

## List of Figures

Figure 1 Map of Pakistan Showing the Coverage of the 2005 Sample	6
Figure 2 Item Map - Urdu	10
Figure 3 Item Map - Maths	11
Figure 4 Availability and Condition of Blackboards in Schools	24
Figure 5 Availability of School Library	25
Figure 6 Availability of the Ministry Of Education Curriculum (2002)	26
Figure 7 Academic Qualification of Head Teachers	27
Figure 8 Professional Qualification of the Head Teacher	27
Figure 9 Percentages of Class 4 Pupils in Small Schools, by Province	37
Figure 10 Percentage of Students by Class Size: Balochistan	37
Figure 11 Percentage of Students by Class Size: Punjab	37

## List of Abbreviations

AIOU	Allama Iqbal Open University (Islamabad)
ACER	Australian Council for Educational Research
AEAC	Area Education Assessment Centre
ATC	Assessment Training Centre
AJK	Azad Jammu and Kashmir
B.Ed.	Bachelor of Education
CT	Certificate of Teaching
DfID	Department for International Development
DoE	Department of Education
ETS	Educational Testing Service
FCE	Federal College of Education (Islamabad)
FANA	Federally Administrated Northern Areas
FATA	Federally Administrated Tribal Areas
GCET	Government College for Elementary Teachers
IAEA	International Association for the Evaluation of Educational Achievement
ICT	Islamabad Capital Territory
IER	Institute of Education and Research (University of the Punjab)
IPI	International Partnering Institution
IRT	Item Response Theory
M.Ed.	Master of Education
MoE	Ministry of Education
MCQ	Multiple Choice Question
NAT	National Achievement Test
NEAS	National Education Assessment System
NIP	National Institute of Psychology (Quaid-e-Azam University)
NPCC	National Planning and Coordination Committee
NWFP	North West Frontier Province
PITE	Provincial Institute of Teacher Education
PTA	Parent Teacher Association
PTC	Primary Teaching Certificate
PPS	Probability Proportional to Size
PEACE	Provincial Education Assessment Centres
SAT	Scholastic Ability Test
SMC	School Management Committee
SD	Standard Deviation
TA	Technical Assistance
TOEFL	Test of English as Foreign Language
TIMSS	Trends in Mathematics and Science Study

## Executive Summary

The Government of Pakistan as well as Government of the Punjab is committed to improve the quality of education along with its efforts to increase enrolment and access. This commitment is reflected in the government's policy documents (National Education Policy 1998-2010, Education Sector Reforms 2001-05) and in its reaffirmation at international forums to include quality outcomes such as student achievement scores as performance monitoring indicators in the education sector. A specific priority within the overall plan is to:

*“build assessment capacity at the school, provincial and federal levels to better measure learning outcomes and improve the quality and effectiveness of program interventions.”* To achieve the targets Govt. of the Punjab has allocated funds more than ever in the history of Pakistan.

Most of the assessment activities in the past have been project driven in Pakistan. There is little institutional base either at the Federal or Provincial level to sustain assessment activities beyond the project timelines. Neither the traditional school examinations nor the external public examinations administered for certification purposes, nor the selection tests for admission to institutions of higher education, sufficiently fulfill the criteria of standardization and comparability required for a monitoring indicator.

National Education Assessment System (NEAS), Ministry of Education at Federal level and its associated centers in the Provinces and Areas (AJK, FATA, FANA), have been established as a priority programme to institutionalize students assessment as a permanent feature of education system. Punjab Education Assessment System (PEAS), Education Department, Govt. of the Punjab is responsible to conduct assessment tests for national and provincial purposes. Central to these efforts is the development of institutional capacity of the Departments of Education at the provincial and area levels to monitor standards of education .

Within Pakistan, NEAS/PEAS has established a partnership with the Institute of Education and Research (IER), University of Punjab, Lahore and with the Federal College of Education (FCE) Islamabad as its Assessment Training Centres (ATCs). The International Partnering Institution (IPI) of NEAS is the Australian Council for Educational Research (ACER), the University of Melbourne, Australia and the Educational Testing Service (ETS) in the United States.

NEAS/PEAS has planned as a sample-based national/provincial assessment at Grade 4 and at Grade 8, in four subjects:

- Language
- Mathematics
- Science
- Social Studies

The objectives of NEAS/PEAS are:

- **Informing Policy:** the extent to which geography and gender are linked to inequality in student performance.
- **Monitoring Standards:** how well the curricula are translated into knowledge and skills;
- **Identifying correlates of achievement:** the principle determinants of student performance and how resource allocation might be re-directed
- **Directing Teachers' Efforts and Raising Students' Achievements:** Assisting teachers to use data to improve student performance.

This report includes the details of curriculum based test development at NEAS and its associated centres, the basic issues and procedures for selection of national sample, particularly problems of low enrolment in some rural areas and application of sampling weights to rectify the imbalances of the population due to low enrolment in rural areas. Results of the first round of national assessment 2005, in terms of scaled achievement scores and the impact of background and context variables on students' achievement are also presented.

The achievement scores of IV Grade students in subject of Mathematics and Urdu are reported on a scale of 1-1000, with a mean set at 500 and a standard deviation of 100. This scale is the same as in TIMSS, SAT, TOEFL etc used in international studies. The scaled mean scores obtained in 2005 can be used as baseline for reporting trends over time as well as for comparing the average and relative performance of different groups of students. The 2005 baseline scaled mean scores can also be used to set targets for improvement of learning achievement to be reflected in future rounds of national/provincial assessment. On this scale, the Punjab Urdu scaled mean score is 382 and the Mathematics mean score is 430, both below the set mean of 500. The National Scaled mean in Urdu and Math is 369 and 421 .

The 2005 results show that achievement test scores are also well below the international average Mathematics score of 495 in TIMSS 2003 but there are four countries (Iran, Philippines, Morocco and Tunisia) whose Mathematics scores are lower than the Pakistan/Punjab Mathematics score. This comparison is only indicative and not entirely valid because of the differences in the NEAS/PEAS and TIMSS curriculum framework, sampling outcome, test administration conditions and students' learning environment.

The analysis of background and context variables in relation to achievement scores revealed that several variables related to student home background, teaching-learning processes and

teaching practices are associated with students' achievement positively or negatively.

Students who reported that they were allowed to ask questions in class scored significantly higher than those who were not allowed to ask questions. Similarly, students who reported getting homework were significantly higher than means scores of those students who did not get any homework in subjects of Mathematics and Urdu.

Supporting inputs from family were also associated with higher mean scores. Students who were taught by their father or brother at home scored significantly higher in Urdu and Mathematics compared to students who reported that their mother or sister taught them at home. In both the subjects, students who did not get help from any body scored the lowest.

The educational and occupational profile of fathers of students across urban and rural areas was similar, though the percentage of negative circumstances (fathers' unemployment, fathers not alive, and illiterate fathers) were slightly higher in rural areas. There was a significant percentage of illiterate mothers in both urban (35%) and rural (49%) areas. The majority of mothers in rural as well as in urban areas were reported to be housewives.

Students who remained absent from school due to illness performed significantly better than those who were absent due to babysitting or household and harvesting responsibilities.

Students of female teachers scored significantly higher in Urdu but not in Mathematics. Teachers academic qualifications ranged from secondary to masters level. However, teachers' qualification did not affect students' performance in either subject.

Reported availability of teaching resources (library, teaching kit, textbooks, teaching guides, curriculum document) did not make any significant difference to the mean scores of students in either Urdu or Mathematics.

There was a significant difference in mean scores of students whose teachers reported teaching two or more classes together in the same period as compared to those whose teachers taught a single class. The mean scores of the students taught by the teachers reporting teaching only one class were better than that of those who were taught by a teacher teaching more than one classes at a time.

Supporting inputs from community (PTA/SMC, supervisory visits) were not significant in determining the students' achievement. Achievement scores of the students from the schools

where non existence of SMC was reported were better as compared to those students where SMC was reported to be existing.

Finally, it would be pertinent to note that the development of a national/provincial assessment system is a complex and challenging task. The preparation and implementation of the 2005 assessment was carried out under tight deadlines. The grade 4 assessment instruments were piloted in April 2004 before the Technical Assistance (TA) input started. These were marked and coded by the NEAS/PEAS team; data entry was outsourced and item analysis was carried out when the TA input started in February 2005.

These constraints were confounded by insufficient staff in the NEAS/PEAS. To ensure that the assessment was conducted in an efficient and timely manner many of the staff were assigned multiple tasks. NEAS/PEAS should, therefore, try to fill staff vacancies on a permanent basis. The capacity building that takes place through hands-on work and training with support from Technical Assistance (TA) is dissipated when there is rapid turnover of staff, delay in administration decision and cooperation. This situation works against sustainability and institutionalization of the NEAS/PEAS network. Furthermore, NEAS/PEAS sample design in 2005 has posed some logistics problems and for 2006 NEAS has employed a modified design which will reduce some of the problems without compromising the representative or random features of the sample.

NEAS/PEAS has generated a large amount of data which can be used for research to identify specific variables that are associated with high student achievement. It is anticipated that when staff from NEAS and its associated centers proceed for higher studies they will be able to use this year's and future NEAS databases for carrying out such studies. Teacher training institutions and University departments of Education, Psychology and other Social Sciences in Pakistan can also use this database for research.

## NATIONAL / PROVINCIAL ASSESSMENT

Learning assessment is increasingly being used around the world to identify strengths and weaknesses of the education system. In a national assessment, measures of achievement in key curriculum areas are administered to students of various selected grade levels. Performance of students in selected sub domains can point to strengths and weaknesses in students' learning achievement within the curriculum areas and can show how the intended curricula are implemented in schools. Data on important student, teacher and school background factors affecting the learning outcomes are collected along with student achievement data to help identify areas needing improvement. This information could then help policy makers identify factors over which they can exert some control.

Like many developing countries, Pakistan is also faced with a problem of expanding enrollment while at the same time improving the quality of education remains a challenge. Little evidence however, is available in Pakistan on the assessment of quality of student learning. To fill-in this gap, a well planned and properly executed national/provincial assessment is needed. Considering this need, the Government of Pakistan and Govt. of the Punjab is committed to improve the quality of education along with its effort to increase enrolment and access. This commitment is reflected in the government's policy documents (National Education Policy 1998-2010, Education Sector Reforms 2001-05) and in its reaffirmation at international forums (Jomtien Declaration 1990 and EFA Assessment 2000, Dakar) to include quality outcomes such as student achievement scores as performance monitoring indicators in the education sector. A specific priority within the overall plan is to:

*"build assessment capacity at the school, provincial and federal levels to better measure learning outcomes and improve the quality and effectiveness of programme interventions."*

As a part of Government of Pakistan's commitment, the Federal and Provincial Governments have also taken some assessment initiatives under development projects in the past. Most of

these assessment activities have been project driven and there is little institutional base either at the Federal or Provincial level to sustain assessment activities beyond the project timelines. Consequently, there is also little evidence of the impact of project investments in training and human resource development in the field of student assessment. The initiatives taken so far are on a small scale and lack standardization and comparability. Neither the traditional school examinations administered for certification purposes, nor the selection tests for standardization and comparability required for a monitoring indicator.

In the backdrop of the outcomes of these initiatives, National Education Assessment System (NEAS) has been developed as a priority programme under the Ministry of Education's (MoE) Education Sector Reform Action Plan<sup>1</sup>.

It is anticipated that the national/provincial assessment will make available credible data to provide feedback to the education system to facilitate resource allocation, policy formulation, the improvement of the curriculum, textbook and materials development, and teaching, supervisory and management practices, all of which will contribute to the enhancement of the quality of education.

---

<sup>1</sup> Education Sector Reform Action Plan (2001-2005) Ministry of Education, Government of Pakistan

## NATIONAL/PROVINCIAL ASSESSMENT INSTRUMENTS

To measure students' learning achievement and to obtain information on factors that have been found to affect the quality of student learning, two different kinds of instruments were used for the assessment of grade 4 students in 2005. These were:

- Achievement tests in Language and Mathematics<sup>2</sup>, and,
- Student, Teacher and Head teacher Background Questionnaires<sup>3</sup>

Achievement tests were developed on the basis of the National Curriculum, 2002. The following processes were adopted to develop the achievement tests:

- Competencies<sup>4</sup> based on the first three levels of the cognitive domain of Bloom's Taxonomy (Knowledge, Understanding and Application) were developed for four subjects, keeping in view the content areas and objectives of the national curriculum;
- A table of specifications was designed for the tests which included content, learning outcomes, number of items and type of item;
- Multiple Choice and Constructed Response, Completion and Matching items were developed<sup>5</sup>. However it was decided to focus the tests on multiple choice and constructed response items to provide better standardization of test setting and marking;

Separate background questionnaires were developed for Head Teachers, Teachers and Students. These questionnaires contained questions to identify the association of various personal, home, school, teaching and community variables with student achievement. For example, the basic dimensions of this dataset included parental education and occupation; supporting inputs from home and community; students attitude towards school and teachers; the teachers' qualification and teaching practices, and multi-grade teaching etc.

---

<sup>2</sup> Instrument development Annex 2

<sup>3</sup> Instrument development Annex 2

<sup>4</sup> Test Competencies Annex 3

<sup>5</sup> Examples of test items are found in Annex 3

Both the mathematics and language achievement tests and through out the country the background questionnaires were piloted in 2004 on a sample of 4593 students<sup>6</sup>. To facilitate systematic conversion of test responses from booklets to computer, these tests were first marked, coded and scored manually followed by data entry in the computer. Classical item analyses were conducted (using ITEMAN software) to select items for the first round of national assessment in 2005. Items were selected on the basis of item statistics (difficulty, discrimination indices) and professional evaluation of item content by subject specialists. Some new items were also developed to improve the item pool and formats. These assessment instruments were administered in May 2005 to a national/provincial sample of 11977 students with the objective of establishing a baseline of achievement for grade 4 students in mathematics and language<sup>7</sup>.

---

<sup>6</sup> A report on the pilot testing 2004 is available from NEAS

<sup>7</sup> Test Administration Annex 4

## NATIONAL/PROVINCIAL ASSESSMENT SAMPLE<sup>8</sup>

In order to draw an adequately representative sample of the nation, the list of Government schools in the National Education Management Information System (NEMIS) database was used as the sampling frame. A stratified, random sampling procedure was used with three agreed explicit strata Province/Area, location (rural/urban) and gender (girls/boys). Data were also analysed by level of school (Mosque, Primary, Middle, Secondary and Higher Secondary) as an implicit stratum to explore if the achievement of grade 4 students was associated with the type of school they attended.

The Punjab sample design covered the entire province (Figure 1). Testing took place in all the 35 districts of Punjab a major step forward from previous sampling designs, which usually began by selecting a few districts, often for reasons of administrative convenience and financial constraints. The rate of participation was almost expected one. Instead of the desired total of 4800, the actual sample was 4016, or 85% of the target. Also to obtain an estimate of the actual population and to rectify the imbalances in population due to low enrolment in rural areas, sampling weights were applied.

Schools were selected in fixed proportions from the defined groups location (rural/urban) and gender (girls/boys).

**Table 1. Punjab Sample for Grade-4 National Assessment 2005**

	Targeted Sampled	Achieved Sampled
No. of Schools	240	237
No. of Students	4800	4016

<sup>8</sup> Item and Background Data Analyses Annex 7

## Punjab Map

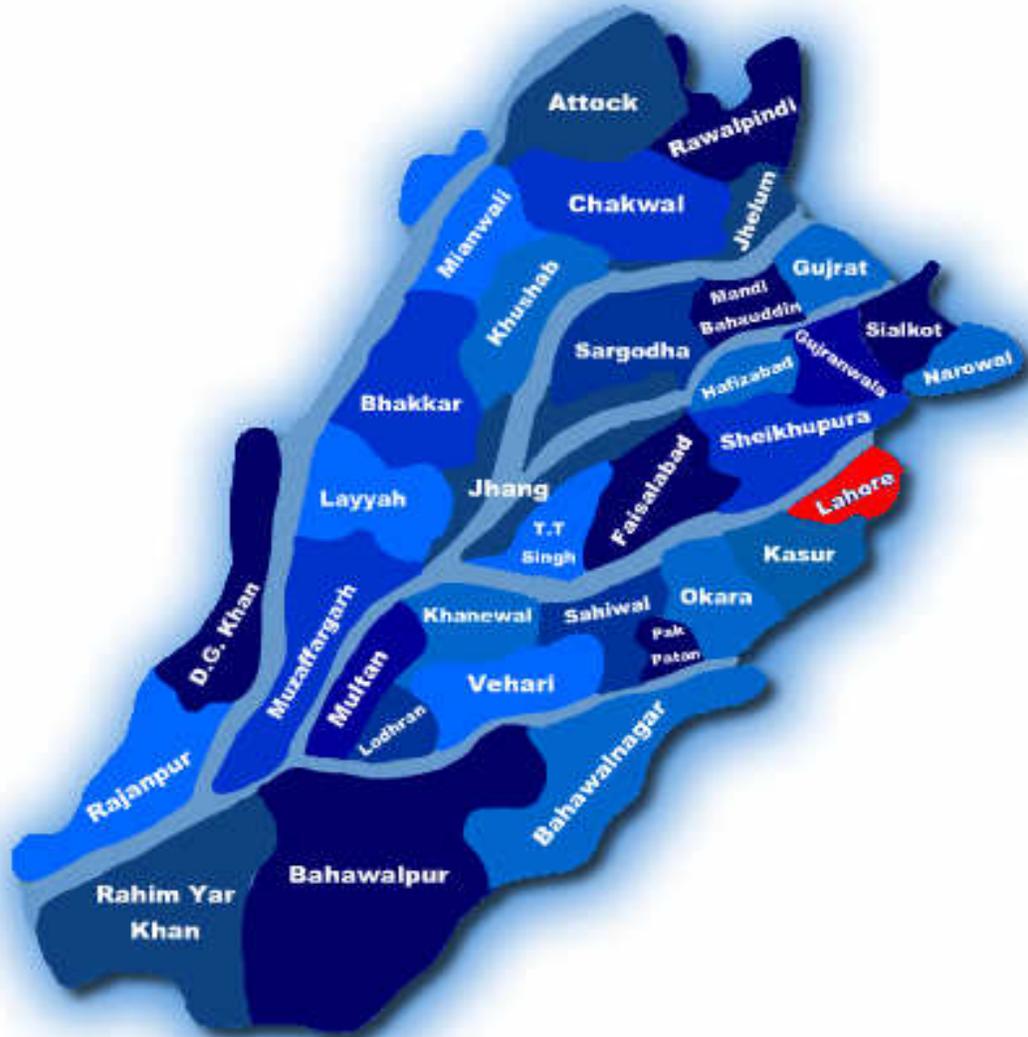


Figure 1. Map of Pakistan Showing the Coverage of the 2005 Sample

## UNDERSTANDING NEAS/PEAS RESULTS

Students' achievement is assessed in terms of their scores on the mathematics and Urdu tests. These scores provide a measure of how well the students answer items based on the National Curriculum. Applying IRT (one parameter RASCH model) calibrations, scores on each test were rescaled. The scales were centered on item difficulty in order to have an indicator of the adequacy of average performance and an indication of how well the test takers performed in relation to the desired level of achievement. Scaled scores could also be used to prevent item selection having an undue influence on the scores of students. The scaled score used in the NEAS/PEAS is the same as that used in many international assessments, for example, TIMSS, SAT, and TOEFL. The range of scaled scores is between 0 – 1000. The achievement scales are constructed so that a student achieving 50% of the marks in a test receives a scaled score of 500. The standard deviation, that is, the extent to which the score can vary from 500, is set at 100. This scale will be used to report scores each time the national assessment is conducted and will be used to report trends over time as well as the average and relative performance of different groups of students. The 2005 baseline scaled mean scores can also be used for improvement of learning achievement to be reflected in future rounds of national/provincial assessment.

The public, policy makers, private sector and schools have an interest in international comparative assessments, using standardized procedures. Trends in Mathematics and Science Studies (TIMSS) are conducted periodically in over 60 countries under the auspices of the International Association for the Evaluation of Educational Achievement (IAEA). It uses national Science and Mathematics curricula for the assessment of learning achievement, and surveys of teachers and students to explore the context in which learning takes place. NEAS/PEAS is moving towards establishing a credible system of national assessment which can become a part of international assessments in the not too distant future.

Based on IRT analysis, Item Distribution Maps presented in Figures 2 and 3 provide interpretative information about a scaled score in terms of the skills and knowledge students with certain scores (ability) are likely to have. The item names (u1, u2, m1, m2 etc.) are shown located at their calibrations along with the person distribution. An "M" marker represents the location of the mean measure. "S" markers are placed one sample standard deviation away from the mean. "T" markers are placed two sample standard deviations away. Items placed along the scale in the item map demonstrate how skills correspond to levels of performance. For example, Figure 2 shows that in Urdu more students in the lower ability range (-1 to -3 on the person distribution side of the dotted line) found answering questions on a given text easier than writing a story themselves. Similarly in Mathematics (Figure 3) students in the lower ability range found recognition and simple computation items easier than problem solving and fraction items.

Scales are created for each subject independently so even though both Urdu and mathematics have the same numerical range (0 – 1000) average scores should not be compared across subjects.

### **Significance Levels**

To check whether differences in reported scores could have occurred by chance alone, significance tests are reported. A probability of  $p < 0.05$  means that the difference could occur by chance alone in only 5 out of 100 cases (usually accepted as a significant level of difference in the Social Sciences). Where  $p$  is  $< 0.01$ , the difference could occur by chance alone in only 1 out of 100 cases (significant difference); whereas if  $p$  is  $< 0.000$  it indicates that the probability of occurrence by chance is less than zero in a 1000 cases (highly significant difference).

### **Analysis of PEAS Results**

The PEAS used widely accepted statistical procedures in analyzing the data. After completion of the data entry process, data is converted into SPSS file format. SPSS is used for basic descriptive/summary statistics and correlational analysis the following soft wares are also used as the major data analysis tools:

- SPSS (Version 12.0(SPSS Inc.))
- ITEMAN (for classical Item analysis) by Assessment Systems (MicroCAT)
- WINSTEPS for IRT Analysis.

## PUNJAB RESULTS

This report presents the National and Punjab Province results of assessment of Urdu and mathematics achievement of a representative sample of grade 4 students in government and community schools.

Grade 4 students achieved a scaled mean score of 382 in Urdu language and scaled mean score of 430 in mathematics on a national/provincial scaled mean set at 500.

**Table 3. Scaled Mean Score for Language and Maths**

Subject	Scale Mean Score
Language (Urdu)	382
Mathematics	430

### Language Results

Grade 4 students in Punjab achieved a scaled mean score of 382 in the Urdu language test. Students' achievement in reading depended upon the cognitive demand of the task. Reading comprehension questions were found to be least difficult by students. Apparently, it seems to be an indication of the emphasis on reading from the textbook as the most common task students are asked to perform during lessons.

On the reading test lower ability students managed to answer correctly items requiring information or knowledge of facts from a given piece of text. However, students found language questions at the understanding level of Bloom's taxonomy most difficult to answer correctly. This may be seen as a reflection of the emphasis on rote learning in most government schools.

In writing, the tasks students found most difficult were different kinds of writing (writing stories, describing their own experiences) handwriting and items requiring knowledge about language (grammar and use of idioms, antonyms).

The item map under IRT summarises students' performance in Urdu by difficulty level and ability.

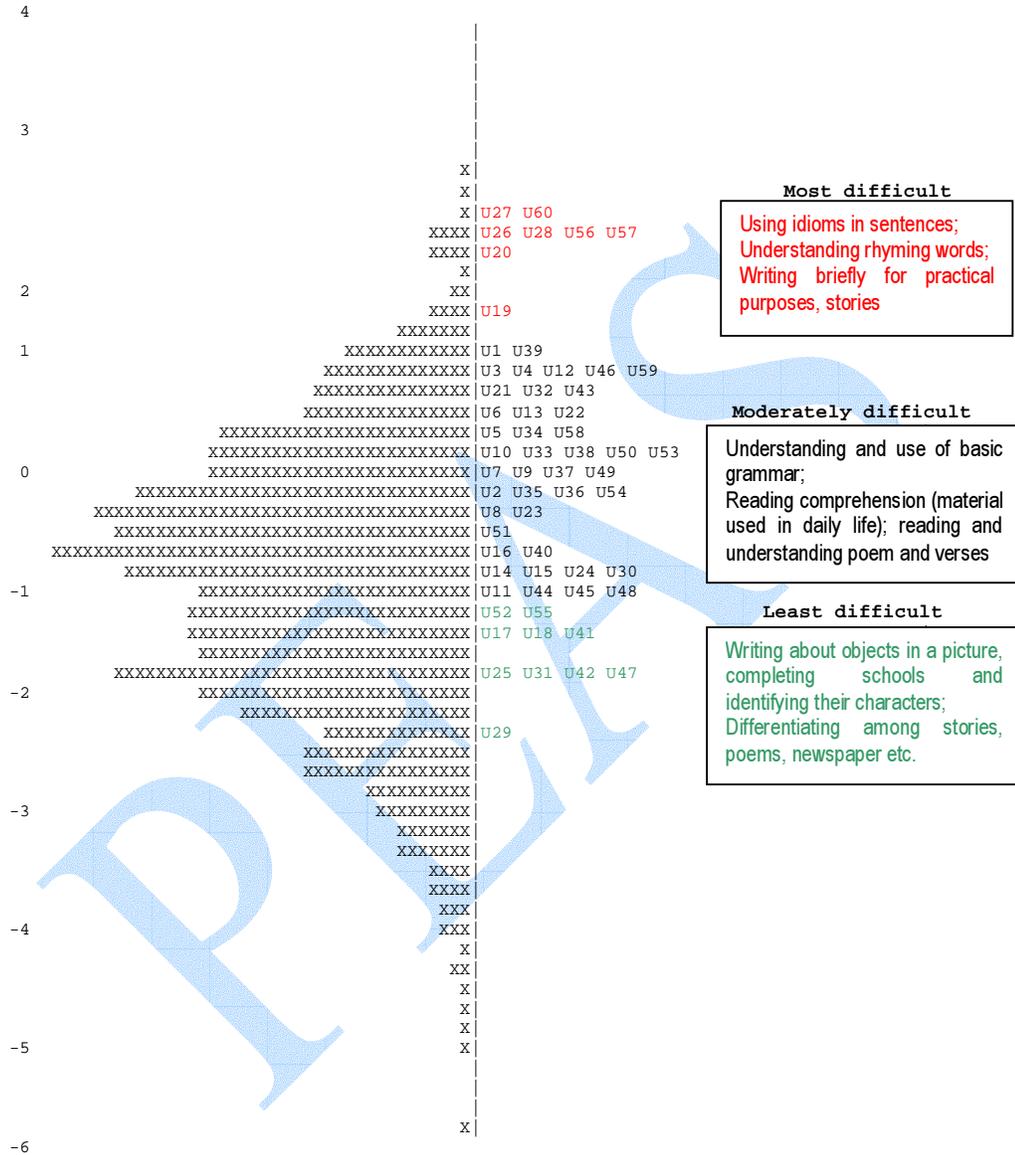


Figure 2. Item Map - Urdu

### Achievement in Urdu by Main Competency Areas

The Urdu achievement test assessed the reading and writing ability of the students as well as their competence in language and vocabulary. Students from Punjab achieved significantly better scores as compared to rest of the nation in reading and writing but their mean scores in language and vocabulary component were significantly lower form that of rest of the nation.

**Table 4. Mean Scores in Urdu by main competency areas**

Competency Areas	Mean Score		t-value	Level of Sig.
	Punjab	Rest of Nation		
Reading	40.38	34.62	160.85	.000
Writing	29.54	26.22	92.3	.000
Language & Vocabulary	22.86	24.42	50.02	.000

### Mathematics Results

Grade 4 students in Punjab achieved scaled mean score of 430 in the mathematics test. The students performed well in knowledge based items such as simple computation, questions about place value and estimation of length, capacity and weight. At the understanding level student achieved average results on items of computation, estimation of length and weight while students performed least well in the application tasks of fractions, odd and even numbers, problem solving, measurement of a line segment, information handling and line graphs.

### Achievement in Math by Level of Cognitive Domain

Mathematics test assessed students achievement in Mathematics in three levels of cognitive domain of Bloom Taxonomy i.e. knowledge, understanding and application. Students from Punjab showed better performance in all the three levels of cognitive domain as their mean scores out of hundred were significantly better than that of rest of the nation.

**Table 5. Mean Scores in Urdu by Levels of Cognitive Domain**

Achievement areas	Mean Score		t-value	Level of Sig.
	Punjab	Rest of Nation		
Knowledge	46.61	43.78	68.563	.000
Understanding	35.64	33.14	72.171	.000
Application	32.57	30.93	36.673	.000

The item map below summarizes students' performance in Mathematic by difficulty level and ability.

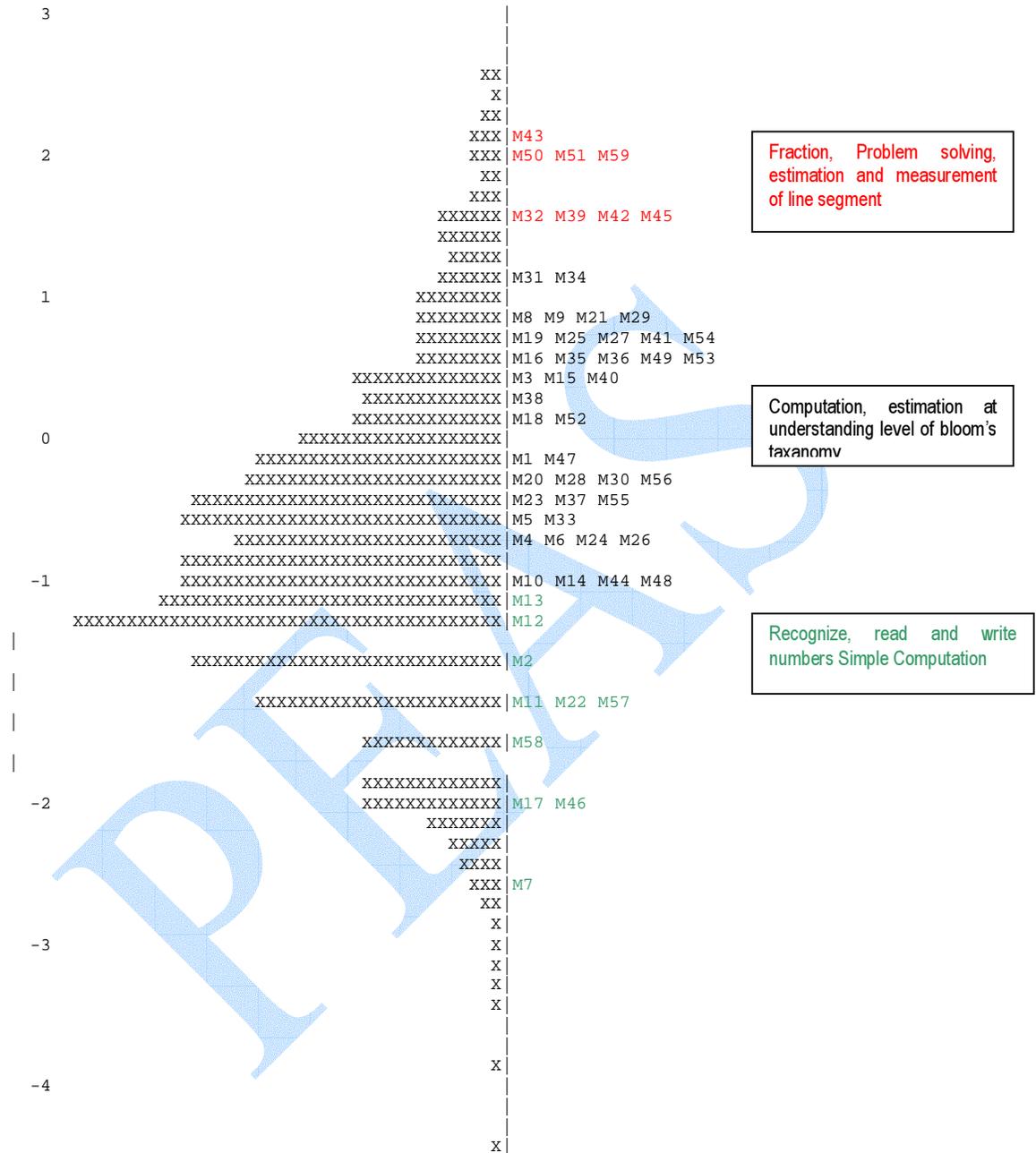


Figure 3. Item Map - Maths

**Student Achievement by Location**

There was a significant difference between the score of students in urban and rural areas in mathematics and Urdu. Scaled mean scores of students from urban area were significantly better in both subjects than the scores of students from rural areas. This difference may be an indication of some difference/inequality in learning environment of rural and urban schools. There may be difference of physical facilities or teaching staff.

**Table 4. Student Achievement by Location**

Subjects	Location	Weighted Number	Scaled Mean Score	t-value	Level of Sig.
<b>Urdu</b>	Urban	211617	417.2758	128.042	.000
	Rural	624426	372.1392		
<b>Math</b>	Urban	211617	434.1920	13.760	.000
	Rural	624426	430.4610		

#### **Student Achievement by Gender**

There was no significant difference between the scaled mean score of girls and boys in mathematics. However, there was a significant difference between boys and girls' scores in Urdu. Girls achieved a higher scaled mean score than boys.

**Table 5. Student Achievement by Gender**

Subject	Gender	Weighted Number	Scaled Mean Score	t-value	Level of Sig.
<b>Urdu</b>	Boys	514931	352.4091	254.029	.000
	Girls	330795	431.7060		
<b>Math</b>	Boys	514931	424.2318	65.997	.000
	Girls	330795	441.9273		

### Mathematics Results for the Provinces/Areas

Students from Balochistan, Punjab, FANA, FATA and ICT scored higher in mathematics as compared to the rest of the nation. Students from NWFP, Sindh and AJK had significantly lower scores in mathematics than the other Provinces and Areas.

**Table 6. Mathematics Results for the Provinces/Areas**

Province / Area	Provincial/Area Mean	Rest of Nation	Probability
Balochistan	443(*)	420	.000
NWFP	415	422(*)	.000
Punjab	430(*)	412	.000
Sindh	402	426(*)	.000
AJK	395	422(*)	.000
FANA	418(*)	412	.000
FATA	451(*)	420	.000
ICT	439(*)	421	.000

### Urdu Results for the Provinces/Areas

Balochistan, Punjab, FANA and ICT Urdu scores were significantly higher as compared to the rest of the nation. NWFP, Sindh, AJK and FATA scaled scores were significantly lower as compared to the rest of the nation.

**Table 7. Urdu Results for the Provinces/Areas**

Province / Area	Provincial/Area Mean	Rest of Nation	Probability
Balochistan	403(*)	368	.000
NWFP	327	378(*)	.000
Punjab	382(*)	356	.000
Sindh	367	369(*)	.000
AJK	359	369(*)	.000
FANA	375(*)	369	.000
FATA	345	369(*)	.000
ICT	433(*)	368	.000

### School Level and Mathematics Achievement

All levels of schools in Punjab are performing better as compared to the rest of nation as scaled mean scores in Mathematics are better as compared to the rest of the nation. Performance of the students from middle school in Punjab as well as in the rest of the nation is lowest one. Mean Mathematics scores of the students from Higher Secondary Schools are the highest and of the students from middle school are the lowest in Punjab. Students from

mosque schools and primary schools achieved better means scores in mathematics as compared with middle and high schools.

**Table 9. School Level and Mathematics Achievement**

School Level	Punjab Mean by Level	Mean: Rest of The nation
Higher Secondary	464	350
Secondary/High	421	418
Middle	412	408
Primary	440	414
Mosque	458	387

### **School Level and Urdu Achievement**

Scaled mean Urdu scores of the students from all levels of the schools in Punjab are better than that of rest of the nation except middle schools. Scaled mean scores of mosque schools students were found to be highest and of middle schools students were lowest. Scores from primary schools and high schools were same but better than students from middle schools.

**Table 10. School Level and Urdu Achievement**

School Level	Punjab Mean by Level	Mean: Rest of The nation
Higher Secondary	440	338
Secondary/High	387	361
Middle	366	375
Primary	387	355
Mosque	427	326

## BACKGROUND AND CONTEXT VARIABLES IN RELATION TO STUDENTS' ACHIEVEMENT SCORES

Findings of previous research in the field of achievement testing and educational research indicate that students' achievement is determined by:

- Student characteristics such as the home background and attitudes towards education;
- The teaching and learning process and teaching practices such as teacher attitude, order and discipline, the variety of teaching strategies, the assignment of homework and providing feedback to students on their work;
- School conditions and climate such as effective leadership, the general facilities of the school, the organized curriculum, flexibility and autonomy;
- Supporting inputs such as parent and community support and effective support from the education system.

Besides testing the students in mathematics and Urdu throughout Punjab, students were administered questionnaires to collect information about their background, socio economic status of the family, family support for studies and many factors which may effect their learning. Background data were collected from the students to investigate the impact of background on achievement of students in different subjects. Data were analyzed to explore the effect of various factors on students' achievement. The details of the achievement of students in the subjects of Urdu and mathematics with respect to different aspects of their background are discussed in the following sections.

### **Student Age and Achievement**

The Punjab sample included 10.4% 9 year old students and 4% were 8 years old while 2.7% of the students were aged 14 years and older. Students aged 9 (the desired age level for grade 4) got the highest mean scaled scores on both the tests. Student aged less than 9 years scored the lowest on both the tests. Similarly, students aged 14 and above (5.2%) got low mean scores on both the tests. This finding may have implications for teachers training to develop teaching skills which supports student learning in a wide age range in the same class.

### Students' age and achievement in Urdu

Age of the students included in the study ranged from 8 to 17 years. Students from age 14 to 17 were included in the single group and univariate analysis of variance was run to see the significance of difference between the mean scores of different age groups of students. It was found that scaled means scores of students of 9 years old were highest and significantly better than all other groups. It is worth mentioning that students aged 9 (the desired age level for grade 4) got the highest mean scaled scores in Urdu.

**Table 11. Students' Age and Urdu Achievement**

Age of the student	Mean	Comparison Groups	Mean Difference	Level of Sig.
age 8 years	377	age 9 years vs. age 8 years	24.58	.000
age 9 years	402	age 10 years	7.69	.000
age 10 years	394	age 11 years	26.58	.000
age 11 years	375	age 12 years	35.60	.000
age 12 years	366	age 13 years	.83	.000
age 13 years	401	age 14 and 14+	25.37	.264
age 14 and 14+	377			

### Students' age and achievement in Mathematics

Univariate Analysis of Variance was run to see the significance of difference between the mean scores of different age groups of students. It was found that scaled means math scores of students of 9 years old were highest and significantly better than all other groups. It is worth mentioning that students aged 9 (the desired age level for grade 4) got the highest mean scaled scores in Mathematics also.

**Table 11. Student Age and Math Achievement**

Age of the student	Mean	Comparison Groups	Mean Difference	Level of Sig.
age 8 years	397	age 9 years vs. age 8 years	47.5	.000
age 9 years	445	age 10 years	5.5	.000
age 10 years	439	age 11 years	27	.000
age 11 years	417	age 12 years	16.6	.000
age 12 years	428	age 13 years	7.7	.000
age 13 years	437	age 14 and 14+	10.8	.000
age 14 and 14+	434			

### Teaching Mathematics at Home

Mathematics scores appear to be affected by family support. Students whose fathers helped them in Mathematics scored higher than those who did not get any support in studies from fathers. Similarly students getting help from brothers and sisters scored higher as compared to students getting help from mothers.

Table 13. Mathematics Teaching at Home

Family member who supports	Mean Score of the students	Comparison Groups	Mean Difference	Level of Sig.
Father	457	Father vs. Mother	50	.000
Mother	407	Brother	28	.000
Brother	429	Sister	31	.000
Sister	426			

### Urdu Teaching at Home

Urdu language achievement is also affected by family support. For language achievement however, sisters support seems to be more effective compared to support from brothers. Fathers also seem to play an important role in improving the learning of Urdu language as compared to mothers. Students getting help from fathers scored better than students getting support from mothers.

Table 14. Urdu Teaching at Home

Family member who supports	Mean Score of the student	Comparison Groups	Mean Difference	Level of Sig.
Father	402	Sister vs. Father	6	.000
Mother	367	Mother	35	.000
Brother	380	Brother	22	.000
Sister	408			

It will be seen in the section on mothers' education that about half the mothers of urban students and two-thirds mothers of rural students are reported to be illiterate, therefore it is not surprising that they cannot help their children much with their studies.

### Study at Home

It should also be noted that students' scores are higher if someone helps them with studies at home. Student who have reported getting *no help* with studies at home have lowest scores in Urdu and Mathematics.

**Table 14. Study at Home**

Subjects	Do you study at home?	Scaled		Level of Sig.
		Mean Score	t-value	
Urdu	Yes	388	276.28	.000
	No	321		
Math	Yes	433	50	.000
	No	399		

### Father's Education in Rural and Urban Areas

Twenty eight percent of fathers in rural areas are illiterate while 18% of fathers in rural areas were reported to be illiterate. In urban areas 33% fathers were reported to be educated to matriculation level while 25% of fathers were reported to be educated to matriculation level.

**Table 16. Father's Education in Rural and Urban Areas**

Education Level	Rural	Urban
Illiterate	28 %	18 %
Primary	35 %	29 %
Matriculate	25 %	33 %
Intermediate	7 %	9 %
Bachelor	3 %	5 %
Master	2 %	5 %

### Fathers Education Level and Students Achievement Level in Mathematics

Scaled math mean scores of the students whose fathers had master degree was highest and those of having primary level of education was lowest. There was no difference between the mean scores of the children whose fathers were illiterate or had primary level of education. Children whose fathers had intermediate level of education scored better then those who had bachelor, matriculation and primary level of education.

**Table 16. Father's Education Level and Mathematics Achievement**

Fathers' Education Level	Mean Score of the student	Comparison Groups		Mean Difference	Level of Sig.
Illiterate	443.9299	Master's	Illiterate	32.4204	.000
Primary	433.3085		Primary	43.0418	.000
Matriculation	435.2884		Matriculation	41.0619	.000
Intermediate	441.0108		Intermediate	35.3395	.000
Bachelor's	407.3023		Bachelor's	69.0480	.000
Master's	476.3503				

### Fathers Education Level and Students Achievement Level in Urdu

Scaled Urdu mean scores of the students whose fathers had master degree was highest and those of having primary level of education was lowest. There was no difference between the mean scores of the children whose fathers were illiterate or had primary level of education. Children whose fathers had intermediate level of education scored better then those who had bachelor, matriculation and primary level of education

**Table 16. Father's Education Level and Urdu Achievement**

Fathers' Education Level	Mean Score of the student	Comparison Groups	Mean Difference	Level of Sig.
Illiterate	391.3736	Master's	Illiterate	.000
Primary	390.9122		Primary	.000
Matriculation	395.8356		Matriculation	.000
Intermediate	415.1039		Intermediate	.000
Bachelor's	395.5444		Bachelor's	.000
Master's	475.7848			

## Father's Occupation in Rural and Urban Areas

Table 17. Father's Occupation in Rural and Urban Areas

Occupation	Rural	Urban
Labourer	22%	19%
Government Job	16%	20%
Private Job	19%	18%
Unemployed	9%	9%
Other	30%	30%
Deceased (not alive)	3%	2%

Only lower socio-economic status fathers' appear to send their children to government schools in urban areas. Similarity the difference between father's education and occupation in rural and urban areas should be noted.

## Mother's Education in Rural and Urban Areas

There was a significant percentage of illiterate mothers in both urban (49%) and rural (64%) areas.

Table 18. Mother's Education in Rural and Urban Areas

Education	Rural	Urban
No Response	< 1%	< 1%
Illiterate	63.7%	48.8%
Primary	24%	26.6%
Matriculation	8.7%	16.1%
Intermediate	1.4%	4.5%
Bachelor's	< 1%	< 1%
Masters	< 1%	< 1%
Other	< 1%	< 1%

## Mother's Occupation in Rural and Urban Areas

The vast majority of mothers take care of household duties whether they live in rural or urban areas.

## Teachers, Teaching Practices and Students' Achievement

### Teacher Gender and Student Achievement

There was a significant difference between teacher gender and student achievement in Urdu tests. Students of female teachers had higher scores in Urdu. However, the majority of girls (80%) and boys (90%) were being taught by teachers of their own gender. More rigorous statistical analysis is needed to determine if the better performance of girls in Urdu is a reflection of the generally reported research finding that girls are better than boys in language skills or the higher scores in this case reflect better teaching by female teachers who may also have better language skills.

**Table 19. Teacher Gender and Student Achievement**

Gender	Urdu Probability Mean	Probability	Maths Mean	
Male	351.42		421.90	
Female	401.36	0.000	414.61	n. s

No significant difference was observed between teacher gender and students' achievement in mathematics. The generally reported finding of better mathematics achievement of boys in some studies is not supported by these NEAS results.

### Teachers' Academic Qualification and Student Achievement

While none of the differences were statistically significant it should be noted that the majority of mathematics teachers were matriculates or graduates. Also 15% of mathematics teachers and 13% of Urdu teachers were master's degree holders. Students of matriculate teachers performed equally well in comparison to students of teachers with master's degrees.

**Table 20. Teachers' Academic Qualification and Student Achievement**

Teachers Academic Qualification	% Urdu	Urdu Mean	% Maths	Maths Mean
Secondary	36	368.56	35	430.02
Intermediate	19	387.27	20	427.87
Bachelors	30	355.67	30	402.75
Masters	13	369.80	15	420.70

### Teachers' Highest Level of Professional Training and Student Achievement

Training received by teachers did not significantly affect students' performance in any of the subjects. None of the mean score differences were significant at 0.05 level.

**Table 21. Teachers' Highest Level of Professional Training and Student Achievement**

Teachers' Training Maths	% Urdu	Urdu Mean	% Maths

		Mean		
PTC	416	66	367	64
CT	446	16	377	18
B.Ed.	386	15	358	16
M.Ed.	449	02	402	02

PEAS

### Achievement Scores in Relation to Teaching Experience

Students whose teachers had 16 – 20 years’ teaching experience achieved the highest scores in both Urdu and Mathematics while Urdu scores of students were the lowest when taught by teachers whose teaching experience was 5 years or less. However, none of these other differences were statistically significant.

**Table 22. Achievement Scores in Relation to Teaching Experience**

Teaching Experience (in years)	% Urdu	Urdu Mean	% Maths	Maths Mean
0-5	15	357.08	15	411.82
6-10	14	354.54	16	414.54
11-15	27	374.36	26	421.92
16-20	24	390.07	24	433.95
21-25	11	338.41	9	428.35
26 and above	8	372.91	8	427.73

### Multi-Grade Teaching and Achievement

Students whose teachers took more than one class at the same time scored lower than those whose teachers taught single classes. These differences however, were not significant statistically for either the mathematics or Urdu tests.

**Table 23. Multi-Grade Teaching and Achievement**

Teachers who Probability Teach more than One class together	Urdu Mean	Probability	Maths Mean	
No	371.19		420.41	
Yes	360.50	n.s.	408.48	n.s.

### Multi-Grade Teaching and Urdu Teachers

Thirty percent of all the Urdu teachers surveyed reported that they had to teach more than one class at the same time.. In Punjab 22.6% of teachers taught multiple classes.

**Table 24. Multi-Grade Teaching and Urdu Teachers**

Province/Area	Scaled Mean	Probability
Yes	74.3%	25.7%
No	69.7%	30.3%
	77.4%	22.6%
	61.8%	38.2%
	60.0%	40.0%
	44.9%	55.1%
	33.3%	66.7%
	87.8%	12.2%
Total	69.6%	30.4%

### Multi-Grade Teaching and Mathematics Teachers

In Punjab 23% of teachers taught multiple classes in the same period. From the information obtained by the questionnaires it can be seen that FATA and FANA have more multi-grade teaching in both subjects taking place than any of the other Provinces and Areas.

**Table 25. Multi-Grade Teaching and Mathematics Teachers**

Province/Area	Scaled Mean	Probability
No	74.1%	25.9%
Yes	66.3%	33.7%
	77.0%	23.0%
	66.7%	33.3%
	55.8%	44.2%
	47.5%	52.5%
	35.4%	64.6%
	85.8%	14.2%
Total	69.9%	30.1%

### Students' Assessment by Teachers

The survey of the sample schools used in the 2005 Punjab Assessment found that classroom observation techniques to assess students' performance were used by only 40% of the teachers. The most common form of assessment was written tests - 81% of teachers in the sample preferred to use this method. Monthly tests were arranged by 87% of teachers while 71% of the teachers prepared regular progress reports for their students. 89% of the teachers stated that they informed parents of their children's achievement through the progress reports. The question arises as to why student scores in the Punjab assessment are below average in Urdu and Mathematics in spite of this fairly good assessment and reporting practice mentioned by teachers. A more detailed study of how these assessments are carried out and why they do not help in improving student performance is needed. Teachers' evaluation of students by homework did not affect their performance. There was no significant difference in achievement between students who were evaluated by homework and those who were not evaluated by homework.

**Table 26. Teacher Assessment of Students through Homework**

	Urdu Mean	Probability	Maths Mean	Probability
No	366.53		422.75	
Yes	375.99	n.s.	423.50	n.s.

### Homework and Achievement

Students who reported getting homework scored significantly higher on both Urdu and Mathematics tests.

**Table 27. Students Report Getting Homework**

	<b>% Urdu</b>	<b>Urdu Mean</b>	<b>Probability</b>	<b>% Math</b>	<b>Math Mean</b>	<b>Probability</b>
No	4.7	297.94		4.7	386.93	
Yes	84	376.52	0.000	84.1	424.95	0.003

### Patterns of Reward and Punishment in Schools

Students who reported getting prizes and incentives scored significantly higher in Mathematics than those who did not get incentives. Similarly, students who reported not getting punishment scored significantly higher in Mathematics than those who got punished. However, neither prizes nor punishment affected performance in Urdu.

**Table 28. Prizes, Incentives and Mathematics Achievement**

	<b>Mean</b>	<b>Probability</b>
No (33%)	401.004	
Yes (67%)	433.883	0.01

**Table 29. Punishment at School and Mathematics Achievement**

	<b>Mean</b>	<b>Probability</b>
No (60%)	432.07	
Yes (40%)	405.58	0.05

### Participation in Co-Curricular Activities

Students participating in co-curricular activities performed higher on both tests as compared to those who did not participate in these activities. The raw frequencies on the co-curricular activities reveal that majority of students participate in the recitation of the Holy Quran/Na'at and games. Fewer students participated in debates and drama.

**Table 30. Participation in Co-Curricular Activities**

	<b>% Urdu</b>	<b>Urdu Mean</b>	<b>Probability</b>	<b>% Math</b>	<b>Math Mean</b>	<b>Probability</b>
No	22.2	346.81		20.2	402.45	
Yes	68.3	379.17	0.05	68.3	428.46	0.04

### Questions Allowed in Class and Achievement

The scores of the students who were not allowed to ask questions scored significantly lower than those who were allowed as questions.

**Table 31. Questions Allowed in Class and Achievement**

Classroom Participation	Urdu Mean	Probability	Maths Mean	Probability
No	287.80		382.73	
Yes	379.65	0.03	435.59	0.02

## School Conditions and Climate

### Availability and condition of blackboards in schools

From the table below it can be seen that while the majority of the assessment sample schools had blackboards, 55% of the urban sample schools and 52% of the rural sample schools stated that their blackboards needed repair.

**Table 32. Availability and condition of blackboards in schools**

Availability of Blackboard	Urdu Mean	Probability	Maths Mean	Probability
No	373.49		398.17	
Yes	368.85	n.s	424.06	0.03*

### Rural Urban

0  
500  
1,000  
1,500  
2,000  
2,500  
3,000  
N

Facilities - Blackboards

No

Satisfactory Condition

Needs Repair

13%

35%

52%

11%

34%

55%

Material for Chart

**Figure 4. Availability and Condition of Blackboards in Schools**

The availability of the blackboard did not have a significant effect on students' Urdu achievement but it had a significant effect on students' mathematics achievement.

### Availability of School Library

The presence of a library did not affect student achievement. There was no significant difference in the performance of students on either the Urdu or Mathematics tests in schools which had a library compared to the achievement of students where schools did not have a library.

**Table 33. Availability of School Library**

Library	Urdu Mean	Probability	Maths Mean	Probability
No	367.45		418.65	
Yes	380.63	n. s	427.98	n. s

More probing questions need to be asked about the use of blackboards and libraries in terms of frequency of use and purpose for which they are used, to determine why such a useful learning resource as a library does not have an impact on student achievement.

Rural Urban  
L100  
L100  
L100  
N  
Is there a school library?  
No  
Yes  
27%  
73%  
32%  
68%

Material for Chart

**Figure 5. Availability of School Library**

#### **Availability of the Ministry Of Education Curriculum (2002)**

The reported availability of the ‘curriculum’ in schools did not contribute toward students’ achievement scores in either Urdu or mathematics. However, it is important to note that the ‘curriculum’ (nisaab) was probably interpreted by most of the schools as the textbooks, because MoE curriculum documents are neither distributed to schools nor they are available in market. The detail is:

**Table 34. Availability of the Ministry Of Education Curriculum (2002)**

Availability of Curriculum		Urdu Mean	Probability	Maths Mean	Probability
No	14%	362.46		411.52	
Yes	86%	371.48	n.s	423.13	n.s

**Rural Urban**

1,000  
2,000  
3,000  
4,000  
5,000  
N  
Curriculum present in school?  
**No**  
**Yes**  
**91.8%**  
**92.2%**  
**7.8%**  
**8.2%**

**Material for Chart**

**Figure 6. Availability of the Ministry Of Education Curriculum (2002)**

As noted in the case of fathers' education and occupation the similarity between facilities or lack of them in rural *and urban* government schools is evident from the above tables.

**Supporting Inputs for Schools**

**Presence of PTA/SMC in the Sample Schools**

It can be seen that the majority of the sample schools had PTA/SMC and in the majority of cases the funding (**over 90%**) came from the Government.

**Table 35. Presence of PTA/SMC in the Sample Schools**

Location	No %	Yes %	Funds from Government %	
Rural	10	90	94	6
Urban	8	92	92	8
Total	7	93	93	7

**PTA/SMC and Achievement**

The presence of the PTA/SMC in a school did not significantly affect the performance of students in either Urdu or Mathematics.

**Table 36. PTA/SMC and Achievement**

PTA/SMC Score	Urdu Mean Score	Probability Score	Maths Mean Score	Probability
Yes	369		418	
No	404	n.s.	436	n.s.

**Academic Qualification of Head teachers**

**Twenty-eight percent** of Head teachers in the sample schools in rural areas reported having a Bachelors' Degree as compared to **33%** Head teacher in urban schools. **Thirtyeight** and **31%** rural and urban head teachers had postgraduate degrees, respectively.

15% of head teachers in the sample schools in rural areas and 18% of the head teachers in urban sample schools have only a secondary qualification.

0 Rural Urban  
**Location Code**  
 0  
 500  
 1,000  
 1,500  
**N**  
 Educational  
 Qualification  
 Secondary level  
 Intermediate  
 Bachelor's  
 Post Bachelor's  
 Other

Material for Chart

**Figure 7. Academic Qualification of Head Teachers**

**Professional Qualification of the Head Teacher**

There was no significant difference in the professional qualifications of Headteachers in the urban and rural sample schools.

Only 43% of the rural and 41% of the urban Head teachers had qualifications of B.Ed. and M.Ed. 57% of head teachers in the rural school sample and 57% of head teachers in the urban school sample had PTC or CT Qualifications.

Again, the educational and professional qualifications of the rural and urban Head teacher are not very different in the government schools sampled here

Rural Urban  
**Location Code**  
 0  
 500  
 1,000  
 1,500  
 2,000  
**N**  
 Teaching Qualification  
 of Principal  
 PTC  
 CT  
 B.Ed.  
 M.Ed.  
 Other

Material for Chart

**Figure 8. Professional Qualification of the Head Teacher**

### Supervisor Visits

Fewer visits of supervisors (less than three-monthly) are linked with high mean scores on the mathematics achievement test. There is no significant difference in the Urdu scores associated with the frequency of supervisor's visits to schools. These reports need to be corroborated from other sources and follow up questions need to be asked about what do supervisors do when they visit schools. Frequency of Supervisor visits is mostly similar for rural and urban schools except for more weekly visits in urban areas.

Location	Weekly Visits %	Maths Mean Score	Twice a month visits	Maths Mean Score	Monthly %	Maths Mean Score	Three Monthly %	Maths Mean Score	Other (Less often)	Maths Mean Score
Rural	12		13		46		19	415.385	9	
Urban	20	410.793	12	420.618	40		18		10	
Total	17		13		43	414.709	18		9	451.897

## CONSTRAINTS AND LESSONS LEARNED

### Constraints

The development of a national/provincial assessment system is a complex and challenging task. It involves the training of staff, the development of test frameworks and items, the organization of the assessment instrument booklets as well as the printing, collating, distribution and collection of the assessment instruments throughout Pakistan country or province under tight deadlines. The grade 4 assessment instruments were piloted in 12 districts of country in April 2004 before the TA input started. These were marked and coded by the NEAS/PEAS team; data entry was outsourced and item analysis was carried out when the TA input started in February 2005.

Between February 2005 and the 4<sup>th</sup> of May 2005 with collaboration of provincial centers the NEAS team with TA assistance reviewed and developed, collated and distributed the assessment instruments and undertook the training of the test administrators. A cycle of activities that takes at least two years in well establishes assessment centres around the world was completed by NEAS and its associated centres in less than a year from instrument development to the administration of large scale assessments.

The above constraints were confounded by insufficient staff in the NEAS, PEACEs and AEACs to ensure that the assessment was conducted in an efficient and timely manner. Many of the staff were assigned multiple tasks due to this constraint;

Problems of low student enrolment were found especially in rural/hilly areas specifically in girls schools. Some schools had less than 5 students at grade 4 level.

The transportation of the assessment materials to enable training and testing to take place was difficult. The return of the assessment instruments after testing was also difficult particularly from remote areas. The approved flat and uniform rates of transportation allowance did not take into account the availability or non-availability of public transport facilities.

## **Lessons Learned**

NEAS/PEAS need to fill staff vacancies on a permanent basis. The staff working in the assessment centres should be retained. The capacity building that takes place through hands on work and training with support from Technical Assistance (TA) is dissipated when there is rapid turnover of staff. This situation works against sustainability and institutionalization of the NEAS network..

Rapid staff turnover also mitigates against the timely implementation of national/provincial assessment activities. More grades and subjects are scheduled to be added to the NEAS implementation plan. Without stability in the technical and office support staff, delivering yearly national assessments at the required level of quality assurance will become increasingly difficult for NEAS/PEAS. In addition, there is an ambitious training plan for staff in the NEAS network. If some staff proceed on training there is a vital need for back up staff to be available to continue the implementation of the planned assessment cycles.

NEAS sample design in 2005 has posed some logistics problems and for 2006 NEAS has employed a modified design which will reduce some of the problems without compromising the representative or random features of the sample. Even so, low enrolments in rural and girls' schools are so evident in the NEMIS database that they cannot be excluded for reasons of logistic convenience, if defensible statements are to be made about the achievement of students.

## **Conclusion**

The findings presented in this report are an attempt by NEAS/PEAS to establish a baseline of student achievement at grade 4 level in Urdu and Mathematics as a starting point for monitoring National and Provincial level trends in the learning achievement of students. The background variables found to be positively or negatively associated with students' Urdu and Mathematics scores are not necessarily causative and need to be considered with caution.

This first round of assessments has generated a pool of data that can be analysed in greater detail and supplemented with rigorous quantitative and qualitative studies to identify more specifically the correlates of achievement in the students' learning environment which can be supported through teacher training and learning materials development. Currently NEAS/PEAS do not have the human and financial resources to undertake such studies. NEAS/PEAS is working under a time bound Implementation plan to deliver several rounds of national/provincial assessments at grade 4 and grade 8 levels by June 2008. It is anticipated that when staff from NEAS and its associated centre proceed for higher studies they will be able to use this year's and future NEAS/PEAS databases for carrying out such studies.

Teacher training institutions and University departments of education/psychology can also use this database for research studies.

PEAS

ANNEXES

PEAS

## NATIONAL/PROVINCIAL TEST SAMPLE

### The Proposed Sample Design

Three sample design options were presented to the National Planning and Coordination Committee, and after extended debate, the Committee agreed on the following basic parameters:

- 32 explicit strata:** Punjab with largest number; other Provinces/Areas approximately pro rata.
- Stratum size; Punjab:** 1537 per stratum (total 6148); Sindh 784 per stratum (total 3136); other provinces have 267 per stratum (1068 per province); areas have a stratum size of 97 (388 per area).
- Total sample size:** 12,972.

This decision was taken as the basis for the final sample design. However, when the full sampling frame became available, the facts it revealed forced the NEAS/pPEAS team to adapt the basic design in several ways.

Table 38 below outlines the numbers of schools in the various provinces/areas, by “gender” and “rural-urban”. It should be noted that there are in fact mixed as well as single-sex schools in Sindh and AJK, as well as ICT. It should also be noted that the “urban” sub-sector is unusually small in AJK and FANA.

**Table 38. Numbers of Schools (from NEMIS)**

Prov./Area	Rural			Urban			Total
	Female	Male	Mixed	Female	Male	Mixed	
Balochistan	2609	7431	0	400	977	0	11417
NWFP	7802	15960	0	703	1115	0	25580
Punjab	25058	30878	0	3535	3849	0	63320
Sindh	5633	17324	15176	1909	2317	1772	44149
<b>Total Provs.</b>	<b>41102</b>	<b>71611</b>	<b>151176</b>	<b>6547</b>	<b>8258</b>	<b>1772</b>	<b>1444466</b>
AJK	2597	3134	0	115	102	0	5948
FANA	253	584	622	35	42	54	1590
FATA	1859	3023	0	0	0	0	4882
ICT	116	119	20	45	34	63	397
<b>Total Areas</b>	<b>4825</b>	<b>6860</b>	<b>642</b>	<b>195</b>	<b>178</b>	<b>117</b>	<b>12817</b>
Grand Total	45927	78471	15818	6742	8436	1889	157283

Table 39 below outlines the numbers of students in the various provinces/areas. Comparing these two shows that schools in some areas, notably Balochistan, are very small.

**Table 39. Class 4 Pupils by Province/Area, Gender and R/U (from NEMIS)**

Prov./Area	Rural			Urban		Total
	Boys	Girls	Boys	Girls		
Balochistan	24941	15281	14625	11241	66088	
NWFP	175479	89132	24403	18016	307030	
Punjab	441103	281374	93380	100797	916654	
Sindh	153477	60767	71143	64255	349642	
<b>Total Provs.</b>	<b>795000</b>	<b>446554</b>	<b>203551</b>	<b>194309</b>	<b>1639414</b>	
AJK	28020	23772	1798	2077	55667	
FANA	7429	4781	925	1183	14318	
FATA	28522	11490	0	0	40012	
ICT	3466	2876	3941	4283	14566	
<b>Total Areas</b>	<b>67437</b>	<b>42919</b>	<b>6664</b>	<b>7543</b>	<b>124563</b>	
Grand Total	862437	489473	210215	201852	1763977	

On this basis, the following principles were defined for sample selection:

- The number of explicit strata would be 38 rather than 32, to allow for the “mixed” schools in Sindh, AJK and ICT.
- Every explicit stratum was to be represented in the sample.
- The smallest number to be sampled from any explicit stratum was 100.
- Where possible, reporting would be based on a minimum of 900 cases in each explicit stratum – which would give a precision of 16% of SD at 95% confidence.
- Because of the large sample size it was necessary to collapse strata – for example, to report for Rural vs Urban and Boys vs Girls in a Province, rather than Rural Boys vs Urban Boys or Rural Girls vs Urban girls.
- Where strata are collapsed in this way, the target was to be 900 cases in each stratum.
- Schools were to be selected with Probability Proportional to Size (PPS).
- The maximum sample size should not exceed 15,000, if at all possible.

On this basis, a revised sample design was agreed. This is given in Table 40 below:

**Table 40. Proposed Sample: by Province/Area, Gender and Rural/Urban**

Prov./Area	Rural			Urban			Total
	Female	Male	Mixed	Female	Male	Mixed	
Balochistan	500	500	0	500	500	0	2000
NWFP	500	500	0	500	500	0	2000
Punjab	1200	1200	0	1200	1200	0	4800
Sindh	500	500	500	500	500	500	3000
AJK	500	500	0	100	100	0	1200
FANA	300	300	300	100	100	100	1200
FATA	600	600	0	0	0	0	1200
ICT	100	100	100	100	100	100	600
<b>Total</b>	<b>4200</b>	<b>4200</b>	<b>900</b>	<b>3000</b>	<b>3000</b>	<b>700</b>	<b>16000</b>

Even this large design inevitably involved some compromises. For example, no attempt was made to include enough cases in ICT to report separately for explicit strata; no attempt was made to report for urban strata in AJK and FATA.

### The Achieved Sample

The total NEAS sample in 2005 was 11,977 grade 4 students. Table 41 provides information regarding the actual numbers of schools and students involved in the 2005 grade 4 national assessment.

**Table 41. Achieved Sample by Province/Area**

Province/Area	Number of Schools	Number of Students
Balochistan	90	1305
NWFP	99	1630
Punjab	237	4061
Sindh	152	2128
AJK	57	745
FANA	59	719
FATA	60	847
ICT	30	542
Total	784	11977

The sample design achieved covered the entire nation well. Testing took place in 117 out of 126 districts and federal regions. The rate of participation was lower than expected (74.86% of the planned sample). In case of province of Punjab testing took place in – districts our 35 districts and rate of participation was 85%.

This in itself is not a major problem; 12,000 and 4061 at national and Punjab level.

**Table 42. Provinces/Areas with above-Average Participation of Students**

Province/Area	Planned	Achieved	%
ICT	600	542	90.33
Punjab	4800	4061	84.60
NWFP	2000	1630	81.50
Grand Total	7400	6233	84.23

High rate of participation from relatively densely-populated provinces and areas. Provinces and areas with below-average participation rates, tended to be those with scattered populations, at least in some areas. It is therefore not at all surprising to discover that the main reason for the shortfall was lower participation in rural areas.

PEAS

**Table 43. Provinces/Areas with below-average Participation of Students  
Planned Achieved**

**(Number)**

**Achieved**

**%**

FANA 1200 719 59.92

AJK 1200 745 62.08

Balochistan 2000 1305 65.25

FATA 1200 847 70.58

Sindh 3000 2128 70.93

Grand Total 8600 5744 65.79

Participation in rural areas in general was much lower than in urban areas, as summarized in Table 44, below.

**Table 44. Sampling Outcome: Rural vs Urban**

**All Pakistan Rural Urban**

Intended 9300 6700

Achieved 5684 6293

Achd. As % of Intdd. 61.12 93.93

Punjab has less than 5% in schools with five or fewer class 4 pupils, and less than 15% in schools with ten or fewer pupils in class 4. Sampling, therefore, is inevitably a far greater challenge in Balochistan than Punjab, because the proportion of small schools is so much higher.

Similarly, the problem is not equally great in rural and urban areas. As chart 2 below shows, the number of class 4 students in Balochistan who are in the very smallest classes (5 or less) is around 20% for rural boys, and 30% for rural girls. By contrast, very few urban boys or urban girls are in such small schools. Therefore, gathering a sample in rural areas poses a far greater challenge than in urban areas, at least in Balochistan. The same is not true in the same way for Punjab, as chart 3 shows. Around 5% or less of Punjab students are in very small schools, even in rural areas.

**Figure 9. Percentages of Class 4 Pupils in Small Schools, by Province**

**Figure 10. Percentage of Students by Class Size: Balochistan**

**Figure 11. Percentage of Students by Class Size: Punjab**

**Balochistan Students**

0  
20  
40  
60  
80  
100  
1-5 6-10 11-15 16-20 > 20

Urban Boys

Urban Girls

Rural Boys

Rural Girls

**Punjab Students**

0  
10  
20  
30  
40  
50  
60  
70  
80  
90  
1-5 6-10 11-15 16-20 > 20

Urban Boys

Urban Girls

Rural Boys

Rural Girls

**Pakistan Province s**

0  
10  
20  
30  
40  
50  
60  
70  
1-5 6-10 11-15 16-20 > 20

Balochis tan

Punjab

Sindh

NWFP

**Data for Charts**

PRELIMS

**Options for Improving Coverage**

Data-gathering problems are essentially rural, and are concentrated in the particularly sparsely-populated provinces and areas. It is however critical that students in sparsely populated

rural areas should be represented in the NEAS sample, because in some provinces and areas they are such a large proportion of the population that the results would be significantly distorted if they are not included. Whatever measures are taken to refine the NEAS sample design, it should respond to this reality. Two specific measures are:

- i. eliminate the very smallest schools from the frame; and
- ii. increase the number of schools where the average size is small.

PEAS

# INSTRUMENT DEVELOPMENT

Item writing for the 2005 National Assessment was originally conducted in 2002 when a Competency Development and Item Writing Workshop took place.

The workshop objectives were as follows:

- To develop technical capacity and sustainability for test development that did not exist in 2002
- To develop grade 4 Competencies identifying an hierarchy of abilities;
- To develop Tests Specifications according to the weightage identified in the curriculum documents;
- To develop test items for four subjects, Urdu, Mathematics, Science and Social Studies and Sindhi language and Sindhi translation\*.

The participants in the Workshop were members of the Curriculum Wing and Subject Specialists from the Provincial and Area Bureaus of Curriculum/Assessment Centres. The workshop was activity based with all participants being involved according to their subject specialism. The outcome of the workshop was the development of **competencies, test specifications and items** (100 for each subject). These competencies and test items were used for the pilot testing of mathematics and Urdu in 2004 and for large scale testing of mathematics and Urdu in 2005.

The assessment framework developed was the foundation for the national assessment and was the basis for all item development. The assessment framework consisted of two organising dimensions – the content dimension and the cognitive dimension. The content domains define the specific subject matter covered by the assessment, and the cognitive domains define the sets of behaviours expected of students as they engage with the subject content (e.g. knowing facts and procedures, using concepts, solving problems, reasoning). Each content domain has several topic areas (e.g., number is further categorized by whole numbers, fractions and decimals, integers, and ratio, proportion, and percent; reading is further categorised by reading for information, reading to find the main idea, identifying genres, vocabulary, punctuation etc.).

On the basis of these domains and the competencies developed, test specifications for all four subjects (Urdu, mathematics, Science, Social Studies) were prepared and test items developed. The specification table provided a guideline to the development of a comprehensive reliable, valid and practical test for the pilot testing in 2004. The same tests were later reviewed and finalised in national level workshops.

\* Statement about Sindhi language and Sindhi medium tests.

The development of test specifications and framework are essential if the testing activity is going to measure the elements for which it is being constructed. A test has to have a clearly stated purpose and should clearly describe the content areas and the grade level for which it has been developed. Also the length of time required for the test should be determined as this would have a direct effect on the number of items in the test and also the breadth of the curriculum to be tested.

The development of a test specification ensured that the test measured a representative sample of the curriculum content and its objectives. It ensured that the curriculum content was more likely to be assessed in a balanced way. Writing a test specification required:

- A list of all the instructional objectives and hierarchy. The framework is developed on the basis of three levels of achievement – knowledge, comprehension, application according to Bloom’s Taxonomy. Knowledge consists of the simple recall of specific information; comprehension enables the student to demonstrate their understanding of the knowledge obtained and application is the student’s ability to use their knowledge and understanding in different contexts and situations.
- The content of the subject areas.
- The weightage to be given to each of the instructional objectives.

From this framework, competency and test specification development **test items** were written. The test items consisted mainly of multiple choice items with a few constructed response items. The main reason for this was that multiple choice items were easily marked and coded while constructed response items required a large amount of training and exemplification of the acceptable answer for a national standard to be achieved.

A workshop was held in February 2003 to develop the **pilot test** in language and mathematics from the pool of items. Two parallel tests were designed. Items which tested the key competencies and those competencies that were able to be tested in a pencil and paper test were to be included. The weightage given to the specific content areas was according to the weightage given in the 1995 National Curriculum.

The tests were further reviewed and printed for pilot testing in 2004. Pilot testing was required for the NEAS to ensure that the demands of the tests were appropriate and also to identify items which were reliable, valid and discriminated appropriately.

A sample of 4953 students from across the whole of Pakistan took part in the pilot testing in 2004 in mathematics and Urdu language.

From the result of the **statistical analysis** of the pilot items (using ITEMAN software), items were selected and additional items developed to “fill in the gaps”. After item selection and writing additional items, formats for large scale testing were developed and administered in May 2005 to a national sample, to establish a baseline for grade 4 languages and mathematics achievement.

The items were then organized into three test booklets. Three test booklets were developed so that students would not be able to copy from each other. The items were ordered in a spiral form.

The **Background Questionnaires** for Head Teachers, teachers and students were also developed. These questionnaires looked at such things as school conditions and climate; teachers and teaching practices; Supporting Inputs for Schools; and, students' home backgrounds. Difficulties were found in constructing some of the questions, as well as in ensuring sufficient coverage of each background and context variable in relation to the length of the questionnaires and the time it would take for the personnel in the sample schools to complete them.

The **distribution** of assessment materials to the Provincial Education Assessment Centres (PEACEs) and the Area Education Assessment Centres (AEACs), which were managing the further distribution of materials and training of Test Administrators under tight deadlines, was a challenging task. All the materials arrived for the national assessment but there were some delays as the deadlines for delivery were very narrow. Some of the **difficulties** experienced in instrument development are listed below:

- Lack of continuity of trained personnel;
- Lack of TA support after 2002;
- Delay in analyzing the 2004 pilot test statistics;
- Ensuring that the items address the 2002 National Curriculum. The original items had been based on the 1995 national Curriculum;
- Ensuring that there were sufficient items to cover the major part of the curriculum;
- Difficulties in spiraling items so that students in the different booklets were given similar items at the same stages in the booklets;
- The time taken to answer the background questionnaires;
- The need to review the appropriateness of some of the questions in the light of the 2005 assessment.

## COMPETENCIES AND SAMPLE QUESTIONS

### Mathematics Class 4 Competencies and Hierarchy

#### No. Curriculum Content Competency

The student will be able to:

1

#### Number and Algebra

##### 1.1 Numbers

Recognise, read and write numbers up to one crore

##### 1.2 Urdu numerals

Read and write Urdu numerals

##### 1.3 Lac, million, crore

Understand the conversion of lac, million and crore

##### 1.4 Odd and even numbers

Identify even and odd numbers

##### 1.5 Addition and subtraction Add and subtract numbers – mentally up to 50

add two 3 digit and 4 digit numbers

Subtract any 4 digit number from a 5 digit number

##### 1.6 Multiplication Multiply and divide 3 and 4 digit numbers by two digit number

1.7 Distributive property Understand and apply the distributive property of multiplication over addition and subtraction

##### 1.8 Pakistani currency Use Pakistani currency notes in simple calculations and problems

1.9 Fractions Understand different types of fraction (equivalent, proper, improper, mixed)

1.9 Smaller and greater fractions Identify smaller and greater fractions with the same denominator

##### 1.10 Addition, subtraction and multiplication of fractions

Add, subtract and multiply fractions with the same denominator and with different denominators not greater than 10

##### 1.11 Addition and subtraction of decimal fractions

Add, subtract decimal fractions up to 3 decimal places

2

#### Measurement and Geometry

2.1 Estimation Show an understanding by estimating length, capacity and weight

2.2 Computation and problem solving Compute and solve problems involving similar and compound units of measurement

2.3 Time Demonstrate knowledge of time using clocks, resources such as timetables, calendars

2.4 Time Problems Compute and solve problems involving time (simple and compound units) in daily life

##### 2.5 Geometrical shapes

Know and differentiate between square, circle, rectangle

2.6 Line measurements Estimate, measure a line segment

##### 2.7 Boundaries, interior and exterior

regions of common plane figures

Show an understanding of boundaries, interior and exterior

regions of common plane figures

National Assessment Report 2005

44

#### No. Curriculum Content Competency

The student will be able to:

##### 2.8 Perimeter of rectangle and square

Compute the perimeter of rectangle and square

#### 3 Information Handling

3.1 Line graphs Identify and interpret line graphs and to show an

understanding of uses in daily life

**No. Competency Knowledge Understanding Application**

**1 Number and Algebra**

1.1 Recognise, read and write numbers up to 1 crore

Recognise the numbers

Find the place value Use real life examples

e.g., differentiate

between the

population of different cities, to

demonstrate

knowledge and

understanding

1.2 Read and write Urdu numerals

Recognise the Urdu numerals

Read and write Urdu numerals

1.3 Understand the conversion of lac, million, crore

Know that;

1 lac = 1 million

10 million = 1 crore

Convert crore to lac to millions and vice

versa

1.4 Identify even and odd numbers

Define even and odd numbers

Categorise even and odd numbers

Demonstrate that things/objects can be paired and odd

things/objects cannot

1.5 Add and subtract numbers

- mentally up to 50

add two 3 digit and 4 digit numbers

- subtract any 4 digit number from a 5 digit number

Add and subtract mentally up to 50

Add two 3 digit and 4 digit numbers both

horizontally and vertically

Subtract any 4 digit number from a 5 digit

number both horizontally and

vertically

Recognise and use appropriate

mathematical

applications for

simple calculations

Use appropriate

mathematical

applications to solve simple problems in number, money, measurement

1.6 Multiply and divide 3 and 4 digit numbers

by a two digit number  
Multiply 3 and 4 digit numbers by two digit numbers both vertically and horizontally

Divide 3 and 4 digit numbers by two digit numbers both vertically and horizontally

Recognise and use appropriate mathematical applications for simple calculations

Use appropriate mathematical applications to solve simple problems in number, money, measurement

1.7 Understand and apply the distributive property of multiplication and subtraction

Demonstrate knowledge and understanding of distributive property of 'x' over '+' and '-' through completing simple calculations  
Solve problems related to daily life

**45**  
**No. Competency Knowledge Understanding Application**

1.8 Use Pakistani currency notes in simple calculations and problems

Add, subtract using Pakistani currency notes in daily life

Add, subtract, multiply and divide using Pakistani currency notes in daily life

Solve problems related to daily life

1.9 Understand different types of fractions

(equivalent, proper, improper, mixed)

Demonstrate knowledge and define common fractions, kinds and decimal fractions

Distinguish between  
different fractions  
Write and convert  
compound fractions  
into improper and  
vice versa

1.10 Identify smaller and  
greater fraction with  
the same denominator  
Recognise the terms  
smaller and greater  
Understand the  
difference between  
greater and smaller  
fractions

Compute daily life  
problems relating to  
fractions

1.11 Add, subtract and  
multiply fractions  
with the same  
denominator and with  
different  
denominators not  
greater than 10

Recognise the  
different applications  
in adding, subtraction  
and multiplying  
fractions with the  
same denominator  
Use appropriate  
applications in simple  
calculations

Use appropriate  
applications to  
compute daily life  
problems relating to  
fractions

1.12 Add, subtract decimal  
fractions up to 3  
decimal places.

Recognise and define  
decimal fractions  
Convert decimal to  
common fractions and  
vice versa

Use appropriate  
applications to  
compute daily life  
problems relating to  
decimal fractions

## **2 Measurement and Geometry**

2.1 Show an  
understanding and  
estimating length,  
capacity and weight  
Recognise the need  
for formal units of  
measurement

Know that:

1km = 1000m

1m = 100 cm

1kg = 1000gm

100kg = 1 quintal

1 litre = 1000 ml

Demonstrate that

objects are of different lengths, capacities and weight

Convert and

interchange units

Measure the length, capacities and weights

Use formal units of measurement in daily life

2.2 Compute and solve problems involving similar and

compound units of measurement

Compute simple examples involving similar and

compound units of measurement

Solve problems involving formal units of measurement using examples from daily life.

2.3/2.4 Demonstrate knowledge of time using clocks, resources such as timetables, calendars

Compute and solve problems involving time (simple and compound units) in daily life

Read the time from clocks

Know the number of days in a week and in a month (solar and lunar calendar)

Compute simple examples.

Place events in time sequences

Work with timetables, calendars to solve simple problems

2.5 Know and differentiate between square, circle,

Identify and name shapes

Identify the different properties of shapes

(shapes which roll, Classify shapes by

simple properties (shapes which roll,

National Assessment Report 2005

46

**No. Competency Knowledge Understanding Application**

rectangle stack, slide, have straight/curved sides)

stack, slide, have straight/curved sides)

2.6 Measure and estimate

line segments  
Demonstrate the difference between a line and line segment  
Measure the given line segment by using a ruler  
Solve problems regarding the use of line segments using examples from daily life

2.7 Show an understanding of boundaries, interior and exterior regions of common plane figures.  
Show an understanding of the meaning of the terms boundaries, exterior, interior regions of common plane figures  
Identify geometrical figures

Demonstrate an understanding by identifying examples found in daily life.  
Demonstrate knowledge of boundaries exterior and interior of common plane figures in daily life

2.8 Compute the perimeter of rectangle and square  
Demonstrate knowledge of the meaning of the term perimeter  
Calculate the perimeter using appropriate formulae  
Calculate the perimeter of everyday objects e.g. book, table, school, classroom

### **3 Information Handling**

3.1 Identify and interpret line graphs and show an understanding of uses in daily life  
Demonstrate knowledge of the difference between pictorial, bar and line graphs  
Explain the horizontal and vertical scales  
Understand the number line  
Interpret the information given in a

simple line graph

## Urdu Language Class 4 Competencies and Hierarchy Levels

### No. Curriculum Content

**Competency: The student will be able to:**

#### 1 Reading

1.1 Read silently, simple writing; understand and explain meaning in own words

1.2 Read simple texts aloud

1.3 Read material used in daily life e.g., children's books, handwritten material, newspaper, advertisements and receipts

1.4 Read poems/verses from poems for enjoyment.

Memorise favourite verses and show understanding of the meaning

1.5 Differentiate between the different types of writing (story, poem, newspaper, letter)

#### 2 Writing and Handwriting

2.1 Copy simple sentences and texts

2.2 Take correct dictation

2.3 Write about things/objects in pictures/incomplete story/picture and identify their important characteristics

2.4 Write, in own words, a read story

2.5 Write briefly for a simple practical purpose, e.g., simple letter, incident, story etc.

2.6 Write briefly about personal experience e.g. diary

#### 3 Knowledge of Language and Vocabulary

3.1 Understand and use basic grammar (noun, verb, adjective, subject, object, pronoun, plural)

3.2 Understand and use antonyms, rhyming words, words beginning with the same letter

3.3 Arrange according to alphabetical order

3.4 Understand and use idioms appropriately in sentences

47

### No. Competency Level 1 Level 2 Level 3

#### 1 Reading

1.1 Read silently, simple writing; understand and explain meaning in own words.

Read simple sentences silently.

Understand meaning of simple sentences through, e.g., rearranging jumbled sentences; identifying important points in a text; suggesting a title.

Follow instructions given in a text

1.2 Read simple texts aloud

Read simple sentences aloud using correct pronunciation and intonation.

Use punctuation to demonstrate an understanding of meaning

Read with fluency to convey meaning of the text.

Able to answer simple questions on the read text.

1.3 Read material used in daily life e.g., children's books, handwritten material, newspaper, advertisements and receipts.

Read simple material  
silently, with  
enjoyment.

Understand meaning  
of simple material  
through, e.g.,  
rearranging jumbled  
sentences, identifying  
important points in a  
text. .

Follow instructions  
given in the material  
1.4 Read poems/verses  
from poems for  
enjoyment.

Memorise favourite  
verses and show  
understanding of the  
meaning

Read simple  
poems/verses from  
poems aloud and  
silently, with  
enjoyment.

Memorise favourite  
verses

Understand the  
meaning of a simple  
poem

1.5 Differentiate between  
the different types of  
writing (story, poem,  
newspaper, letter)

Know the difference  
between prose and  
poetry

Differentiate between  
the different types of  
writing (story, poem,  
newspaper, letter)

Understand  
implications of, e.g.,  
the environment, over  
population, and make  
predictions

### **No. Competency**

3 Knowledge of Language and Vocabulary

3.1 Understand and use basic grammar (noun, verb, adjective, subject, object, pronoun, plural)

3.2 Understand and use antonyms, rhyming words, words beginning with the same letter

3.3 Arrange according to alphabetical order

3.4 Understand and use idioms appropriately in sentences

National Assessment Report 2005

48

### **No. Competency Level 1 Level 2 Level 3**

#### **2 Writing**

2.1 Copy simple  
sentences and texts.

Copy simple  
sentences and texts  
correctly for the most  
part

Form letters and space  
words legibly for the  
most part

Copy simple  
sentences and texts  
correctly in good

handwriting

2.2 Take correct dictation Write one sentence correctly

Write at least three sentences correctly with correct punctuation and spelling

Write at least five sentences correctly with correct punctuation and spelling

2.3 Write about things/objects in pictures/ incomplete story/picture, and identify their important characteristics

With the aid of printed words and phrases will write a sentences about things/objects in pictures

With the aid of printed words and phrases will write at least three sentences about things/objects in pictures, identifying the important characteristics.

Write at least five sentences regarding things/objects in pictures.

Write an ending to a story.

Write a title or heading for a picture story

2.4 Write in own words, a read story.

Write one sentence, using own words, about a read story.

Write, at least three sentences, using own words, about a read story.

Write at least five sentences, in own words, about a read story.

2.5 Write briefly for a simple practical purpose, e.g., simple letter, incident, receipt, invitation etc.

Complete, e.g., a letter, by filling in the gaps.

With the aid of printed words and

phrases will write a few sentences.  
Write several sentences, in own words  
2.6 Write briefly about personal experience, e.g. diary  
Write at least one sentence, using own words, about a personal experience, e.g., Eid, visiting a friend.  
In writing a diary is able to write date etc. correctly  
Write, at least three sentences, using own words, about a personal experience, e.g., Eid, visiting a friend.  
In writing a diary is able to write date etc. correctly  
Write at least five sentences, in own words, about a personal experience, e.g., Eid, visiting a friend.  
In writing a diary is able to write date etc. correctly

**No. Competency**

3 Knowledge of Language and Grammar

3.1 Understand and use basic grammar (noun, verb, adjective, subject, object, pronoun, plural)

3.2 Understand and use antonyms, rhyming words, words beginning with the same letter

3.3 Arrange according to alphabetical order

3.4 Understand and use idioms appropriately in sentences

49

## **Mathematics Grade 4 Test Examples**

**Example 1:**

Competency 1.4 Number and Algebra

Odd and even Numbers (knowledge)

**Example 2:**

Competency 1.1 Number and Algebra

Recognise, read and write numbers up to one crore

- Place value (understanding)

**Example 3:**

Competency 2.2 Measurement and Geometry

Compute and solve problems involving similar and compounds of measurement (understanding)

National Assessment Report 2005

50

## **Urdu Grade 4 Test Examples**

**Example 1:**

Competency 3.1 Knowledge about language

Understand and use basic grammar

(Noun, Verb, Adjective, Subject, Object, Pronoun, Plural)

**Example 2:**

Competency 3.1 Knowledge about language  
Understand and use basic grammar  
(Noun, Verb, Adjective, Subject, Object, Pronoun, Plural)  
51

**Example 3:**

Competency 2.3 Write about a picture and identify the important characteristics  
(level 1)

**Example 4:**

Competency 1.3 Read and understand material used in daily life (level 2)  
National Assessment Report 2005  
52

**Example 5:**

Competency 1.4 Read poems/verses and understand their meaning (level 3)

PEAS

## TEST ADMINISTRATION

The tests were administered in May 2005, by teachers in the Provinces and Areas according to the instructions in the Test Administration Guideline Booklet<sup>9</sup> developed. The teachers were trained by the Master Trainers from the PEACEs and AEACs, FCE, IER, AIOU, NIP who had received training in NEAS. The Lead Trainers as well as the NEAS staff monitored the test administration.

For uniformity of test administration, NEAS trained Master Trainers at Federal level. The Master Trainers then trained the Test Administrators (Teachers, Subject Specialists, GCET and PITE faculty) who were to administer the assessment instruments in the Provinces/Areas. The master trainers trained by NEAS were as follows:

**Table 45. Number of Master Trainers Trained by NEAS**

<b>PEACEs/AEACs</b>	<b>Number of Master Trainers</b>
Balochistan	5
NWFP	4
Punjab	10
Sindh	7
AJK	2
FANA	7
FATA	2
ICT	8
FCE	1
IER	1
AIOU	1
NIP	1
<b>Grand Total</b>	<b>49</b>

<sup>9</sup>The Test Administrator's Guidelines are available from NEAS

National Assessment Report 2005

These master trainers trained test administrators throughout Pakistan. The number of test administrators trained was as follows:

<b>PEACEs/AEACs</b>	<b>No of Master Trainers</b>
Balochistan	200
NWFP	200
Punjab	480
Sindh	300
AJK	120
FANA	120
FATA	120
ICT	60
<b>Grand Total</b>	<b>1600</b>

Some of the **difficulties** identified in the test administration were as follows:

- There was a delay in the delivery of the Test Administration Manual to the Provinces/Areas, resulting in a delay in the training/non-training of the test administrators;
- Some of the Test Administrators did not always appreciate the need for the assessment to be conducted in a rigorous manner;
- Test Administrators did not always use the examples in the test booklets to familiarize the students with the test methodology;
- Test Administrators did not always follow the guidance given in the guidance booklet.

# MONITORING NATIONAL ASSESSMENT

Monitoring was undertaken in two areas:

- Monitoring the Administration of National Assessment, and,
- Monitoring the marking and coding of the assessment instruments

### **Monitoring the Administration of National Assessment**

The main objective of monitoring the national assessment was to ensure the validity of the National Assessment data. It is important that all aspects of the national Assessment is standardized, including the administration of the assessment instruments.

The monitors consisted of the Master trainers from NEAS, PEACEs and AEACs, FCE, IER, AIOU and NIP. The monitors monitored the assessment activity and reported back to the NEAS on how the test administrators followed the guidelines given during the test administration training.

### **Monitoring the Marking and Coding of the Assessment Instruments**

Marking and coding of all the scripts in the 2005 National Assessment was undertaken in NEAS. Instructions regarding the marking and coding were given to the personnel involved.

The responsibility of the monitors was to:

- Ensure that the marking and coding was conducted in an efficient and fair manner;
- Provide support to those markers and coders who were having some difficulty;
- Identify marking and coding discrepancies and correct them where possible;
- Take a 10% sample of the marking and coding sheets to identify the validity of the marking and coding;
- Ensure that the correct code was being used on the coding sheets;
- Ensure that no marker malpractice is taking place such as the deliberate alteration of a mark to inflate or deflate a student's original mark
- Ensure that the assessment instruments were complete and returned appropriately when the marking and coding was completed.

Some of the difficulties identified in these activities are found below:

- There was some duplication of assessment instrument monitors in the schools – a need for improved communication between the monitors;
- Test Administration Monitors need to be given specific instructions as to how they should behave/react if they observe that the assessment instruments are not being administered appropriately;

National Assessment Report 2005

- There was a need for the training of the markers and coders to be more rigorous;
- There was a need for more extensive monitoring of both the assessment booklets and coding sheets – 10% of the number of students on a coding sheets was not sufficient to ensure that the marking was being conducted efficiently and fairly;
- Little information was available of any evidence of irregularities in test administration such as identical but unlikely patterns of response in multiple choice answers in more than one booklet; identical mistakes and peculiarities in the scripts of students (for open ended questions) sitting close to each other; changes in patterns of response to questions, for example, well-written answers with few errors alongside hurriedly written answers with many grammatical and syntactical errors.

PEAS

# MARKING AND CODING ASSESSMENT INSTRUMENTS

For the 2004 pilot, marking and coding methodologies were developed by each of the subject specialists including rubrics for the constructed response items (mostly for language). For the constructed response items (mainly in language) exemplars for the correct answers were provided and agreement was reached as to what was acceptable for a mark to be allocated.

Coding sheets for each subject were further developed in 2005 using the Excel program. Each possible answer was given a specific code. The markers did not mark questions right or wrong. If the first possible answer was chosen a code of 1 was given; for answer 2 a code of 2 was given; for answer 3 a code of 3 was given for answer 4 a code of 4 was given. Where a student had not answered a code of 8 was given and where a student has not yet reached the question a code of 9 was given.

Students from the Federal College, Islamabad and teachers from the Federal Directorate of Education, Islamabad were hired for the manual marking and coding. This involved marking and coding of approximately 45,000 assessment items. They were instructed on how to enter the data on the coding sheets before the start of the marking and coding process. They were paid for the completion of each booklet. Due to the large volume of work to be completed in a limited time, delays and problems in marking and coding and later in data entry were encountered.

The staffs of the NEAS, PEACEs, AEACs and ATCs were involved in the monitoring of this activity, checking 10% of the data. Checking the data involved taking every sheet and checking two out of the 20 students on each sheet. This was too onerous a task so it was not possible to check every single sheet. Where mistakes were found the students/teachers employed were asked to recheck their sheets and correct the mistakes. There appeared to be a lack of understanding of the need for rigors in this work and it appears that the majority of the scorers and coders were mainly interested in the quantity of booklets they could complete rather than in doing the task well.

It was noted that when some of the NEAS staff re-checked the data:

- There were mistakes in the manual data entry. The codes were not accurately entered;
- As only two students out of 20 were checked on each sheet, 18 students' data entry was not checked. Also where there were, for example, only 18 students on a sheet or 50 students on a sheet, only two students were checked. This does not give an accurate representation of a 10% sample.
- There were also some mistakes in the data entry of the scores but these were negligible.

- When the SPSS data was checked there were also some additional mistakes made. Data cleaning in itself also resulted in mistakes. Even some missing figures were found to have been entered.  
Marking and coding test booklets is an onerous task. The difficulties that have arisen from the methodology used for marking and coding should be reviewed. More provincial/area participation should be encouraged as recommended in the World Bank Aide Memoir, July 2005 and the training of the markers and coders should be more thorough – besides explaining the methodology, trial runs of entering the data should take place and where the scorers and coders have difficulty their participation should be discontinued. As well as this the monitoring of the scoring and coding should be much more rigorous and manageable with 10 -20% of the sheets being monitored rather than 10% of 20 students. More monitors should be appointed so that the ratio of monitors to coders and scorers is 1:5.

# ITEM AND BACKGROUND DATA ANALYSES

Data analyses at NEAS takes place in three Phases:

- Phase-I consists of Microsoft Excel for sample selection;
- Phase II includes item and reliability analysis of the pilot administration for selection of test items; while,
- Phase-III consists of achievement and background data analysis to determine the correlates and determinants of achievement among the students population.

The following is a description of softwares used for each phase of analysis.

### **Phase I: Microsoft Excel**

This is used at the stage of sample selection and is quite helpful in basic arrangements of sampling frames, calculations of sampling intervals, generating random number seeds and sorted lists of sampled schools.

### **Phase II: Item Analysis**

Item analysis is carried out using the following programs

- a) **ITEMAN** is used for conducting “classical item analysis” for the purpose of item selection. This program provides information on two parameters; item difficulty and item discrimination that could be used in item selection along with information on performance of distractors for each item.
- b) **SPSS** is used mainly for data input and some basic descriptive and graphical analysis to supplement ITEMAN outputs.

### **Phase III: Achievement and Background Data Analysis**

- c) **WINSTEPS** is based on Rasch IRT analysis. For the NAT 2005 WINSTEPS was used for one parameter (item difficulty) IRT analysis as a basis for scaling. IRT parameters are sample independent and provide the basic statistical information about each individual item.
- d) For two and three parameter IRT analysis, either **Parscale** or **Bilog-MG** is used. These programs have not been introduced at NEAS yet but they could be used for linking the NEAS results to international assessments such as TIMSS.
- e) **SYSTAT** is used to check the consistency of data, editing of data and for some basic descriptive analysis. Because of its more flexible options (e.g., bootstrap) and better all round performance it is sometimes preferred over SPSS.
- f) **AM** software by American Research Institute is used for a limited range of analysis at NEAS. Particularly AM is used for application of tests of significance, and computation of raw frequencies. The program has a provision for automatic weighting and estimation of standard errors using Jackknife.