PN ABX-710
97958

TEACHER KNOWLEDGE OF THE CONTENT OF PRIMARY SCHOOL CURRICULUM IN MATHEMATICS, SCIENCE, AND LANGUAGES

Sar Biland Khan

AED Consultant

Washington, D.C.

The Directorate of Primary Education

Hayatabad, Peshawar, NWFP

Government of Pakistan

January 31, 1994

TABLE OF CONTENTS

Executive summary iii

Introduction 1

Methodology 3

Results 7

Overall analysis 49

Summary 64

Conclusions and Recommendations 65

Appendices 68

Executive Summary

The purpose of this study was to improve the content knowledge of teacher-trainees in the 18 Government Colleges for Elementary Teachers (GCETs) in the class V level mathematics, science, Urdu, and Pashto (where applicable). For this purpose, initial testing was done in October, 1993 followed by a one-day workshop on teaching some important concepts in mathematics and Urdu language. On the basis of the analysis of the initial performance data, feedback was provided in weak areas of the class V curriculum. A post-test will be administered to the same students at the end of the training program to determine if improvement did occur as a result of the intervention. The present report is based on the analysis of the pre-test data only.

The sample included 1685 trainees (1039 males and 646 females); 60% of the total enrolment in the training colleges. Tests in mathematics and science had already been developed as part of the National Educational Assessment Program (NEAP) and had been used with class V children in NWFP. Tests in Urdu and Pashto were especially developed for the present study. The tests were administered by the staff of the Instructional Materials Development Cell (IMDC) of the Primary Education Directorate. The tests were scored and in the process various item and test statistics were obtained. Additional analyses included analysis of gender difference, between-subject differences, Pashto-area and non-Pashto area differences, within-college gender differences, and correlations among the subject scores and college average scores.

The results revealed relatively poor performance in math (53% average) and science (57% average). Performance in Urdu (79% average) and Pashto (85% average) was in line with would be expected of teacher-trainees assuming that the tests in languages were of appropriate difficulty for class V children.

Weak areas in mathematics included concepts such as prime numbers, decimal fractions, place value, calculation of areas, etc. In science, concepts such as, characteristics of living things, atoms and molecules, water and clouds, etc. were difficult for the teacher-

trainees. There were fewer areas of difficulty in Urdu and Pashto; mainly on concepts such as spelling, (aarab), etc.

Females performed significantly poorly in all content areas when compared to their male counterparts. Within content-areas differences were also significant. Trainees did poorer on the mathematic tests compared to science and Urdu tests. Similarly, they did poorer in science compared to Urdu. Scores on the Pashto test were not compared to scores in other content areas because Pashto test was not given to all trainees. Colleges having Pashto-speaking and non-pashto-speaking students differed in performance only on the Urdu test. Correlations among the subject scores were significant. Mathematics correlated higher with science than with Urdu and Pashto. Science correlated higher with Urdu than with Pashto.

On the basis of the initial results, the recommendations offered were: upgrading of content as part of the training, raising academic qualifications for entry into PTC program, using an admission test as part of the entry requirements, obtaining similar data on trainees enrolled in Short-course Program and Allama Iqbal Open University Program, increasing the duration of the training program, monitoring the performance of GCETs on a regular basis, issue of permanent teaching certificate to trainees after two-years of satisfactory performance at the classroom level, and regular in-service training for active teachers in upgrading their content and methodological skills.

Introduction

School Learning and academic achievement is affected by a variety of factors. These may include, among others, physical facilities and equipment; instructional materials such as textbooks and teacher guides; the quality of teachers in terms of their academic and professional qualifications, commitment and professionalism; student characteristics, such as aptitude, attitudes and motivation, study habits, parental support and commitment; etc. One can formulate the teaching-learning process in the following manner, admittedly somewhat simplistic:

Input----------> Process ----------> Output

The quality of the output, in this case, school learning and achievement (skills, attitudes, and behaviors) assessed through reliable and valid procedures provides an indication of the effectiveness of the combined influences of the input and process characteristics. The teacher plays a pivotal role as an input in the total learning environment of the child. His\her academic qualifications (level and competence), professional qualifications, (pedagogical competence), motivation, and professionalism as relevant characteristics are probably the most significant influences on the learning and achievement of children. An earlier study on the content-know edge of primary school teachers (Teacher Content Knowledge Study by Ellen Van Kalmthout, Directorate of primary Education, NWFP, November, 1992) revealed serious deficiencies on the part of active primary school teachers. That is to say that active primary school teachers possessed inadequate knowledge

of the content especially in mathematics and science. It was concluded that the teacher-training institutions do not pay adequate attention to the upgrading of content-matter and probably concentrate too much on the methodology of teaching. The irony is that no matter how effective teachers' preparation is in the methodology, if they do not know the content to be taught; the preparation in methodology becomes irrelevant in effecting positive change in the content learning of the primary school children. As a consequence, the present study was designed with an overall objective of improving teacher-training in the content knowledge area. Specifically, the study was designed to:

1. Test the PTC students at the time of entry to the training colleges in class V mathematics, science, Urdu, and Pashto (where applicable);
2. Provide guidance to the faculty of each college in the teaching of certain concepts in mathematics, science, and language;
3. Share results of the initial testing with each college, and especially provide information on those content areas which require additional instruction to upgrade the knowledge of the students during training;
4. Test the PTC graduates at the end of their training to determine if there had been any change between the performance at the entry and exit points;
5. Develop a faculty profile for each College and examine the relationships among faculty characteristics and trainees' performance.

Methodology

Instruments

Test were developed based on the class V curriculum in the areas of Mathematics, Science, Urdu, and Pashto. Items in the mathematics and science tests had been field-tested and used as part of the large-scale assessment of the performance of class V children while items in Urdu and Pashto tests were specifically developed for this study and could not be field-tested due to time constraints.

There are two major approaches to examinations and assessments: (i) the essay-type, and (ii) the objective-type. Everyone connected in any way with education, is familiar with the essay approach. Essentially, a few questions are asked on the course content and students are required to write their answers. The marking of the answers is somewhat subjective, tedious, and time-consuming. The objective-type approach consists of tests having a large number of items (the students usually identify the correct option) whose scoring is not open to different interpretations. The scoring is quick and almost mechanical. The essay-type approach becomes unwieldy if one is testing a large number of students. Therefore, for this study, we opted for the objective-type approach (multiple-choice format; a statement or question followed by options among which one is correct) as the study involved testing of a large number of PTC students.

Item-banks had already been developed in class V mathematics and science. Item-bank in mathematics included 315 items and the item-bank in science included 228 items. Items in

these banks were field-tested involving class V children and further revised on the basis of results from the field-testing. Final tests representing each domain of content were developed from these banks, (each test contained 60 multiple-choice items having four options) were used in a large-scale assessment study. The same tests with very minor changes were used in the present study. Items in the tests represented the domains of content from which they were drawn; the domain of content effectively being the class V mathematics and science textbooks.

The Urdu test contained 46 items. The items were measuring concepts such as word meaning (synonyms & antonyms), grammar, spelling, and reading comprehension. Thirty-two items were of multiple-choice type, twelve items required the students to write in the response, and two items required students to put disjointed sentences into meaningful order. The Pashto test contained 50 items having a design similar to the items in the Urdu test. Each item in the language tests had two or three options. Initially, it was desired to have four options for each item, however, the development of plausible distractors presented a formidable task. Therefore, items with fewer options were accepted as adequate with the realization that the probability of choosing the correct option by guessing increased considerably.

Population and Sample

~~There are 18 teacher-training colleges (11 male and 7 female) in NWFP with an~~ enrolment of 1732 males and 1077 females. A sixty percent random sample from each college was chosen for testing. The total sample included 1685 students (1039 males and

646 females). In terms of academic qualifications, the sample included: Matriculation arts (455), Matriculation science (311), F.A.(375), F.Sc (344), B.A. (109), B.Sc (58), M.A. (1), and M.Sc (2). Thirty trainees did not respond to this question. Trainees were also asked if they had any teaching experience as an untrained teacher. The sample included 120 teachers who had teaching experience ranging from one to six years. The majority of them, about 77%, had only upto one year or less teaching experience.

Test Administration

A lot of costly effort in test development can go to waste if the final test is not administered according to recommended procedures. We would like that the test score of a student is a "true" indicator of his/her ability in the content area which the test is designed to measure. Although obtaining a perfectly reliable score for students is "impossible", a close approximation to this value is the one we are striving for. A score may be partly faulty because the test was not administered according to the instructions of the developers. To guard against this possibility, testing was conducted by the staff of the Instructional Materials Development Cell (IMDC) after receiving adequate training. A manual containing directions for the administration of the tests was also prepared and distributed to the testing staff. The person who had the overall responsibility for each testing site was required to complete a report.

Coding and Data Entry

Instructions for coding and data entry were prepared. A five digit ID code was

assigned to each test booklet. Background information on gender, qualifications, etc. was also coded. The test booklets were coded by IMDC staff and then delivered to EMIS for entry of the test responses. The test responses were entered using R-Base. Files in R-Base were then converted to ASCII Files for further analysis.

Analysis

An item-analysis and scoring program was used to analyze and score test-responses for each subject. The item analysis yielded information such as percent of students choosing each option, items-test correlations, and summary statistics such as average, variance, median, average item-test correlations, average difficulty level and alpha reliability coefficient for the total test. The analyses were done for the total sample and separately for each college.

One score file (e.g., mathematics was sorted by ID code and saved as a system file. Similarly, another subject-file (science) for the same students was sorted by ID code saved as a system file, then matched by student ID with the mathematics file, and saved as a combined system file. In this manner, all four subject scores were placed in a combined system file. Statistical analyses such as frequencies, histograms, and other descriptive statistics by college, by gender, correlations among the subject scores, and correlations among the college subject averages were obtained. Secondary analyses involving the total sample were done to test for differences by gender, by college, by academic qualifications, etc.

Results

Results of the pretest analysis are given in this section. Descriptive results are given first for each subject followed by analysis of relationships among the subject scores and differences on key background variables such as gender, college, language, qualifications, etc. Summary statistics for each college are given in Appendices A to D.

Mathematics

Table 1 includes the p-values (percent of students who responded correctly to an item). Arbitrarily, setting a criterion of 50% success rate, students found calculating the area of a triangle as the most difficult (p=9%). It was followed by the identification of prime numbers (p=15%), place value in decimal fractions (p=17%). Concepts having p-values ranging between 21% and

---

Insert Table 1 about here

30% were: Fractions: word problem (21%), decimal fractions:division (23%), calculating area of a rectangle (23%), proper/improper fractions (27%), calculating volume (30%), and equivalent fractions (30%). P-values ranging from 31% to 40% corresponded to these concepts: common divisors (38%), mixed fractions: division (36%), decimal fractions: mixed operations (35%), and rays and angles (35%). Concepts for which p-values fell in the range of 41% to 50% included: fractions: addition, mixed fractions: subtraction and multiplication,

TABLE # 1
PERCENT SUCCESS-RATE ON MATH ITEMS (K=60) BY PROVINCE (N.W.F.P)
PTC STUDENTS

ITEM NO/CONCEPT	*P-VALUE*	*ITEM NO/CONCEPT*	*P-VALUE*
1. Equivalence : Roman Arabic Numerals	85	14. Common Multiples	91
2. Prime Number	15	15. Common Multiples	71
3. Divisibility	51	16. Least Common Multiples (LCM)	56
4. Prime Factorization	62	17. LCM (Word Problem)	76
5. Place Value Whole Number	56	18. Associated Property	63
6. Fractions (Num/Denominator)	71	19. Fractions (Simplification)	62
7. Fractions (Interpretation)	52	20. Mixed Fractions(simplification)	89
8. Proper/Improper Fractions	27	21. Fractions (Addition)	19*
9. Composite Numbers	71	22. Fractions (Addition)	50
10. Common Divisors	38	23. Fractions (Addition)	42
11. Equivalent Fractions	30	24. Fractions(Mixed + whole No)	67
12. Fractions (Interpretation)	58	25. Fractions (Subtraction)	87
13. Greatest Common Divisor	78	26. Fractions (Subtraction)	49

TABLE # 1 (Contd)

PERCENT SUCCESS-RATE ON MATH ITEMS (K=60) BY PROVINCE (N.W.F.P) PTC STUDENTS

ITEM NO/CONCEPT	*P-VALUE*	*ITEM NO/CONCEPT*	*P-VALUE*
27. Mixed Fractions (Subtraction)	50	40. Decimal Fractions (Place Value)	17
28. Fractions (Multiplication)	72	41. Decimal Fractions (Subtraction)	59
29. Mixed Fractions (Multiplication)	50	42. Decimal Fractions(Multiplication)	49
30. Fractions (Division)	50	43. Decimal Fractions(Multiplication)	76
31. Mixed fractions (Division)	38	44. Decimal Fractions(Multiplication)	48
32. Fractions (Word Problem)	21	45. Decimal Fractions (Division)	36
33. Fractions (Mixed Operations)	50	46. Decimal Fractions (Division)	63
34. Fractions (Mixed Operations)	49	47. Decimal Fractions (Division)	23
35. Fractions (Order)	52	48. Decimal Fractions (Mixed opera)	35
36. Decimal Fractions (Addition)	63	49. Unitary Method	77
37. Conversion to Decimal Fractions	49	50. Average	53
38. Conversion to Decimal Fractions	55	51. Rays and Angles	35
39. Conversion to Mixed Fractions	65	52. Area (Triangle)	9

TABLE # 1 (Contd)

PERCENT SUCCESS-RATE ON MATH ITEMS (K=60) BY PROVINCE (N.W.F.P) PTC STUDENTS

ITEM NO/CONCEPT	*P-VALUE*	*ITEM NO/CONCEPT*	*P-VALUE*
53. Area (Square)	28	57. Bar Graphs (Interpretation)	76
54. Area (Rectangle)	23	58. Bar Graphs (Interpretation)	85
55. Area (Word Problem)	46	59. Line Graphs (Interpretation)	66
56. Volume (Cubic)	30	60. Line graphs (interpretation)	67

OVERALL MEAN (NWFP)=31.84, PERCENT= 53, S.D=11.06, N=1685

Note : Item no 21 should not be considered due to printing error

in the stem of the item.

[K] indicates the total number of items in the test.

fractions: mixed operations, conversion to decimal fractions, multiplication involving decimal fractions, and calculating cost of a construction job by obtaining area of a rectangle. The students found the concept of common multiples (p=91%), simplification of a mixed fraction (p=89%), subtraction of fractions with the same denominator (p=87%), converting a 2-digit arabic numeral into roman numeral equivalent (p=85%), and interpreting a bar graph (p=85%) as relatively easiest.

Table 2 includes frequency distribution of the p-values (percent of students who responded correctly to an item). About 47% of the items have p-values upto 50%. Only 9% of the items have p-values greater than 80%. The

---

Insert Table 2 about here

---

overall mean of the scores on the mathematics test is 31.84 (53%) with a standard deviation of 11.06 (see Table 3). The median of the scores is 32. The internal consistency reliability (Cronbach's alpha) is 0.91 indicating that the test items are highly intercorrelated having a homogenous structure.

Insert Table 3 about here

The item-total test correlations has an average value of 0.53 in the from of biserial correlations while it is 0.40 in the form of point-biserial correlations. These values indicate that the items are quite valid as a collective measure of the mathematical concepts and skills in the class V curriculum. The average

TABLE # 2

*FREQUENCY DISTRIBUTION OF P-VALUES OF MATH ITEMS (n=1685)**

INTERVAL	*MATH Items=60)*	
	FREQUENCY	*PERCENT*
00 - 10	1	2
11 - 20	3	5
21 - 30	7	12
31 - 40	5	8
41 - 50	12	20
51 - 60	9	15
61 - 70	9	15
71 - 80	9	15
81 - 90	4	7
91 - 100	1	2

* Item no 21 with p=.19 had an error in the stem.

TABLE # 3

Summary Statistics for Mathematics

Statistical Index	Value
No of Items	60
No of Students	1685
Mean	31.84
Variance	122.31
Std.Dev	11.06
Skew	-0.001
Kurtosis	-0.62
Minimum	0.00
Maximum	58.00
Median	32.00
Alpha	0.91
SEM	3.27
Mean p-value	0.53
Mean Point-Biserial	0.40
Mean Biserial	0.53

difficulty level is 53%, slightly higher than the median.

Figure 1 includes a frequency polygon of the scores. An examination of the frequency polygon curve reveals that the score distribution approximates almost a normal distribution. Such a distribution of scores would be expected from testing a representive sample of class V children. This means that the performance of these trainees is similar to what would be expected from average class V children. Various percentile points are given in Figures 2 and 3. The bars in Figure 2 corresponds to actual score points below which certain percent of scores fall. Figure 3 include bars corresponding to percent scores below which certain percent of the students have their scores. For instance, 25% of the students have scores below the score point 23 or about 38%.

Insert Figures 1, 2, and 3 about here

Three-fourths of the students have scored below the actual score point of 40 (approximately 67%).

There is a significant difference between the performance of males and females (see Table 4). Mean for the male students (n=1039) is 34.89 with a standard deviation of 10.76. Mean for the female students (n=646) is 26.91

Insert Table 4 about here

---

with a standard deviation of 9.73. Males have performed significantly better than females ($F_{(1,1683)}=236.87$; $p<0.001$).

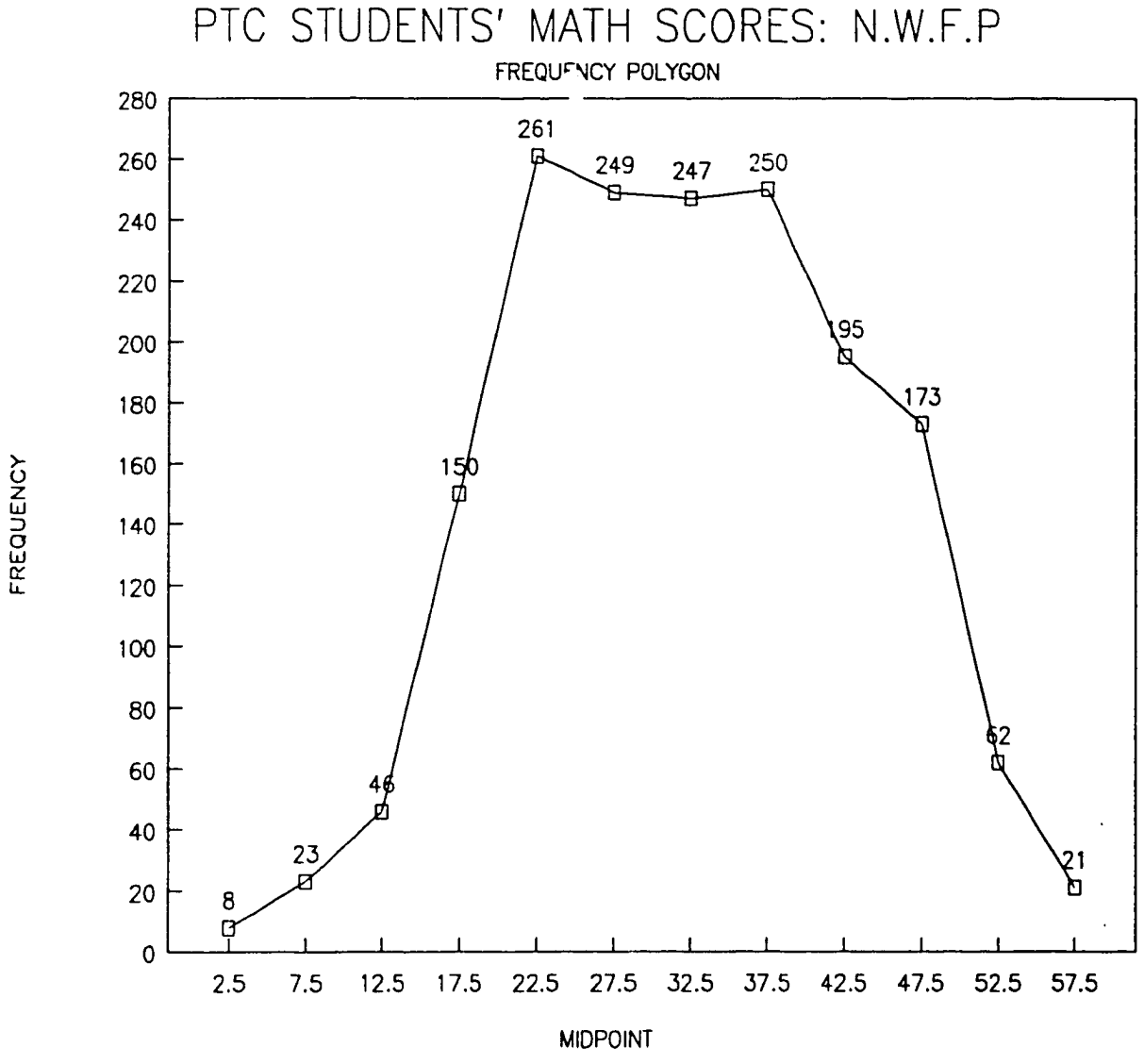
PTC STUDENTS' MATH SCORES: N.W.F.P
FREQUENCY POLYGON
FREQUENCY
MIDPOINT
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
2.5
7.5
12.5
17.5
22.5
27.5
32.5
37.5
42.5
47.5
52.5
57.5
8
23
46
150
261
249
247
250
195
173
62
21

Figure. 1

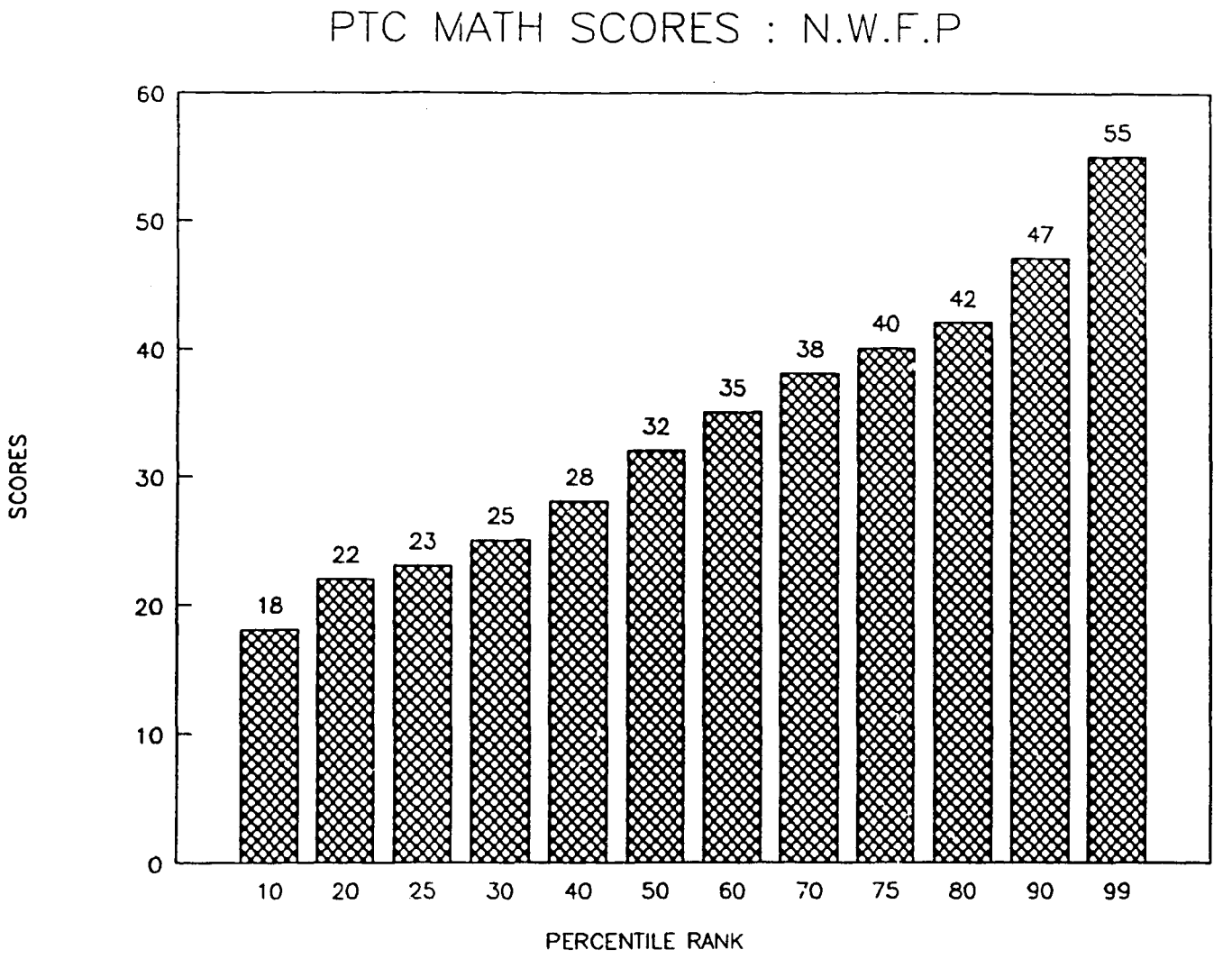
PTC MATH SCORES : N.W.F.P
SCORES
60
50
40
30
20
10
0
18
22
23
25
28
32
35
38
40
42
47
55
10
20
25
30
40
50
60
70
75
80
90
99
PERCENTILE RANK

Figure. 2

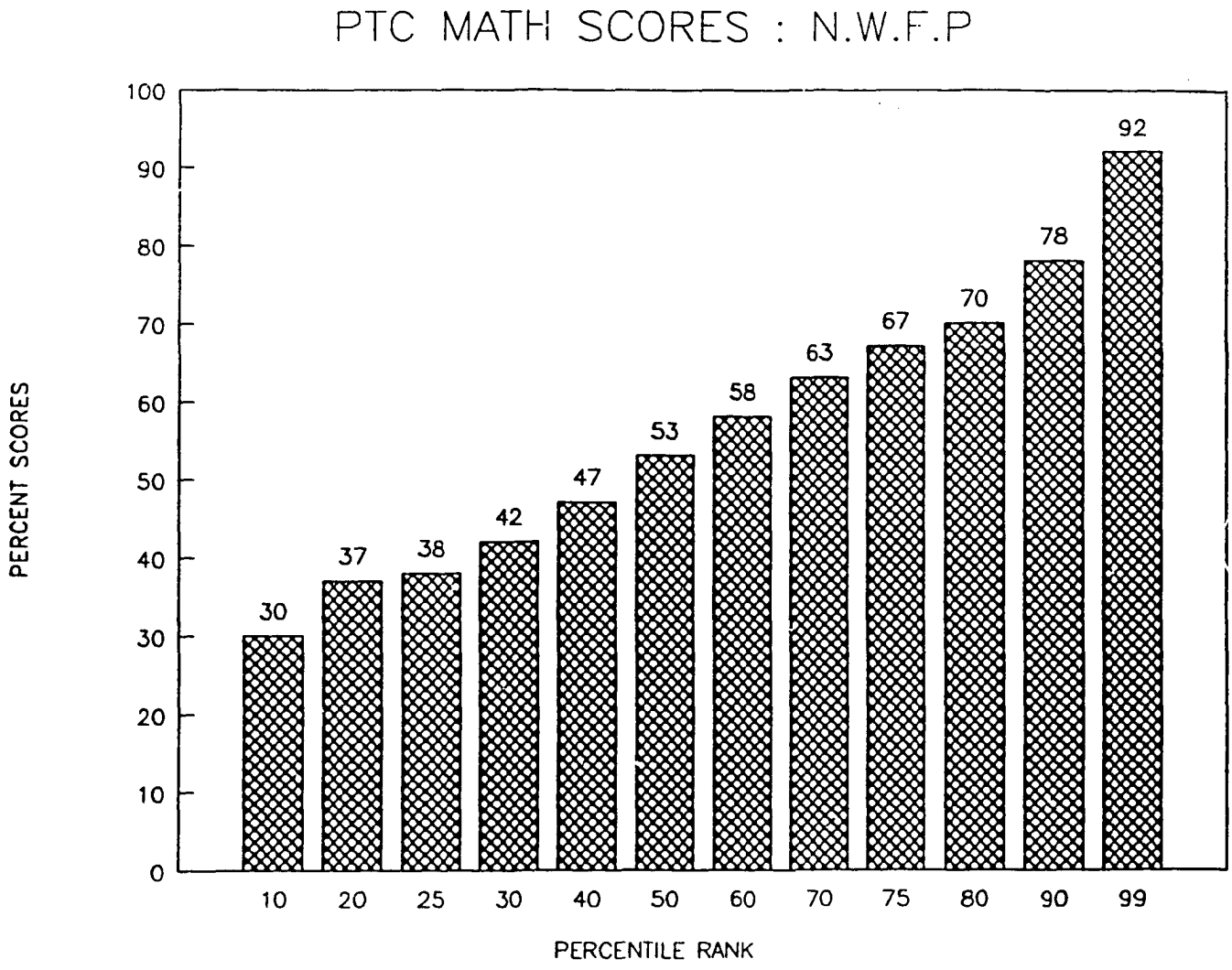
PTC MATH SCORES : N.W.F.P
PERCENT SCORES
100
90
80
70
60
50
40
30
20
10
0
30
37
38
42
47
53
58
63
67
70
78
92
10
20
25
30
40
50
60
70
75
80
90
99
PERCENTILE RANK

Figure. 3

TABLE # 4

Analysis of Gender Differences for the Mathematics Test

	Mean	%	S.D	n	F	p
Overall	31.84	53	11.06	1685		
Males	34.89	58	10.76	1039		
Gender					236.87	<0.001
Females	26.91	45	9.73	646		

Science

As already mentioned, the science test contained 60 multiple-choice items. The p-values (percent responding correctly to an item) are given in Table 5. The frequency distribution of the p-values is also provided (see Table 6).

Insert Tables 5 and 6 about here

About one-third of the items have p-values less than 50%. About eight percent of the items have p-value greater than 80%. Relatively difficult items were: characteristics of living things'; 'plants production and growth'; 'carbohydrate foods'; 'food and health'; 'different things have different atoms'; 'distance between molecules'; 'water level'; 'water purification'; 'clouds'; 'wind movements'; 'energy and photography'; 'mechanical energy'; 'properties of gases'; 'light and color of objects'; 'magnetic energy'; 'production of charge', etc. The students found questions on 'fertilizers'; 'animals which can live in the water'; 'the kind of energy required to use an electric heater when its picture is given'; ' the solar system' relatively the easiest.

The frequency polygon is given in Figure 4. Various percentile points are given in Figures 5 and 6. About 92 percent of the students have scores in 21-50 range (the total possible score is 60). One-fourth of the students have scored below 29 (48%) while three-fourth of the students have scored

TABLE # 5

PERCENT SUCCESS-RATE ON SCIENCE ITEMS (K=60) BY PROVINCE (N.W.F.P) PTC STUDENTS

ITEM NO/CONCEPT	P-VALUE	ITEM NO/CONCEPT	P-VALUE
1. Animals	78	26.Atoms & Molecules	74
2. Animals	78	27.Atoms & Molecules	64
3. Animals	54	28.Atoms & Molecules	45
4. Animals	60	29.Gases	70
5. Animals	64	30.Atoms & Molecules	55
6. Animals	95	31.Atoms & Molecules	60
7. Animals	43	32.Atoms & Molecules	76
8. Animals	39	33.Atoms & Molecules	67
9. Insects	57	34.Water	56
10.Plants	53	35.Atoms & Molecules	48
11.Plants	67	36.Water	78
12.Plants	71	37.Water	35
13.Plants	96	38.Water	29
14.Plants	65	39.Water	28
15.Plants	70	40.Water	14
16.Plants	33	41.Water	70
17.Plants	69	42.Force & Energy	54
18.Plants	39	43.Force & Energy	64
19.Plants	82	44.Force & Energy	34
20.Foods	39	45.Force & Energy	87
21.Foods	63	46.Force & Energy	45
22.Foods	68	47.Force & Energy	38
23.Foods	54	48.Gases	55
24.Foods	37	49.Gases	41
25.Atoms & Molecules	62	50.Light	55

TABLE # 5 (Contd)

PERCENT SUCCESS-RATE ON SCIENCE ITEMS (K=60)
BY PROVINCE (N.W.F.P)
PTC STUDENTS

ITEM NO/CONCEPT	P-VALUE	ITEM NO/CONCEPT	P-VALUE
51.Light	46	56.Magnetism	53
52.Light	49	57.Magnetism	78
53.Light	35	58.Magnetism	58
54.Magnetism	70	59.Electrostatics	38
55.Magnetism	18	60.Solar System	81

Overall Mean(NWFP)=34.11 ,Percent=57, S.D= 8.78, N= 1684

Note: K indicates the total number of items in the test.

TABLE # 6

FREQUENCY DISTRIBUTION OF P-VALUE OF SCIENCE ITEMS (n=1684)

INTERVAL	SCIENCE (Items=60)	
	FREQUENCY	PERCENT
00 - 10	0	0
11 - 20	2	3
21 - 30	2	3
31 - 40	10	17
41 - 50	7	12
51 - 60	13	22
61 - 70	14	23
71 - 80	7	12
81 - 90	3	5
91 - 100	2	3

below 40 or 67 percent (see Figures 4, 5, and 6).

Insert Figures 4, 5, and 6 about here

The mean of the scores on the test is 34.11 (about 57%) with a standard deviation of 8.78. The distribution of scores is slightly negatively skewed (Skewness = -0.35). The performance of the trainees on the science test is similar to that which would be expected from a sample of class V children of average ability. The median of the test scores is 35. The internal consistency reliability is 0.85. The average item-total score correlation in the form of a biserial correlation is 0.42. The reliability of the test, although slightly lower than what one would desire, is quite acceptable. The item-total score correlation indicates that the items in general are quite valid as a collective measure of the concepts included in the class V science curriculum. The average score of 1038 male students is 36.54 (61%) with a standard deviation of 8.42 (see Table 8). The average score of 646 female students is 30.21 (50%) a with standard deviation of 7.92.

Insert Tables 7 & 8 about here

The male students have performed significantly better than female students ($F_{(1,1682)}=236.17$; $p<0.001$).

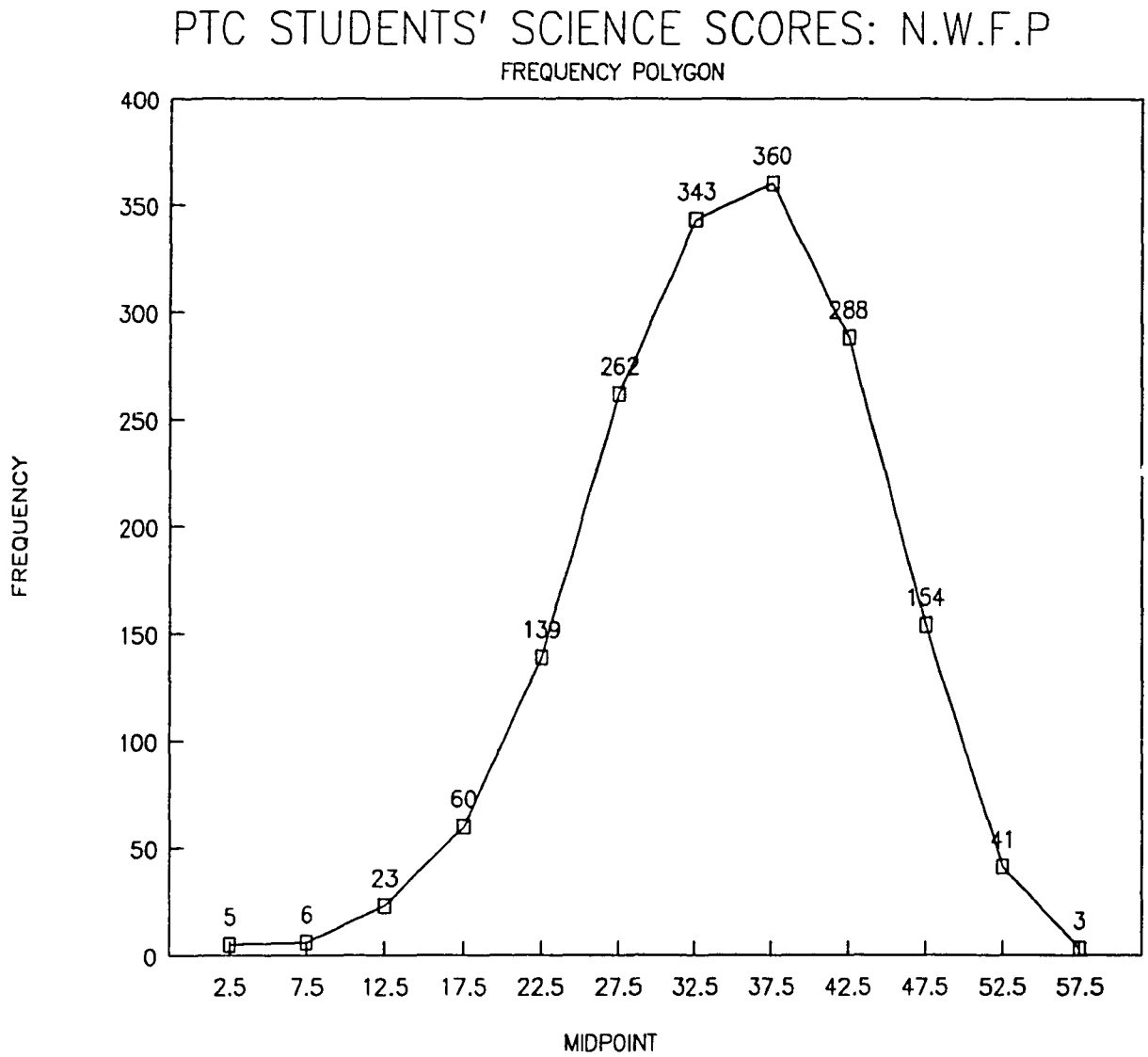
PTC STUDENTS' SCIENCE SCORES: N.W.F.P
FREQUENCY POLYGON
FREQUENCY
MIDPOINT
400
350
300
250
200
150
100
50
0
2.5
7.5
12.5
17.5
22.5
27.5
32.5
37.5
42.5
47.5
52.5
57.5
5
6
23
60
139
262
343
360
288
154
41
3

Figure. 4

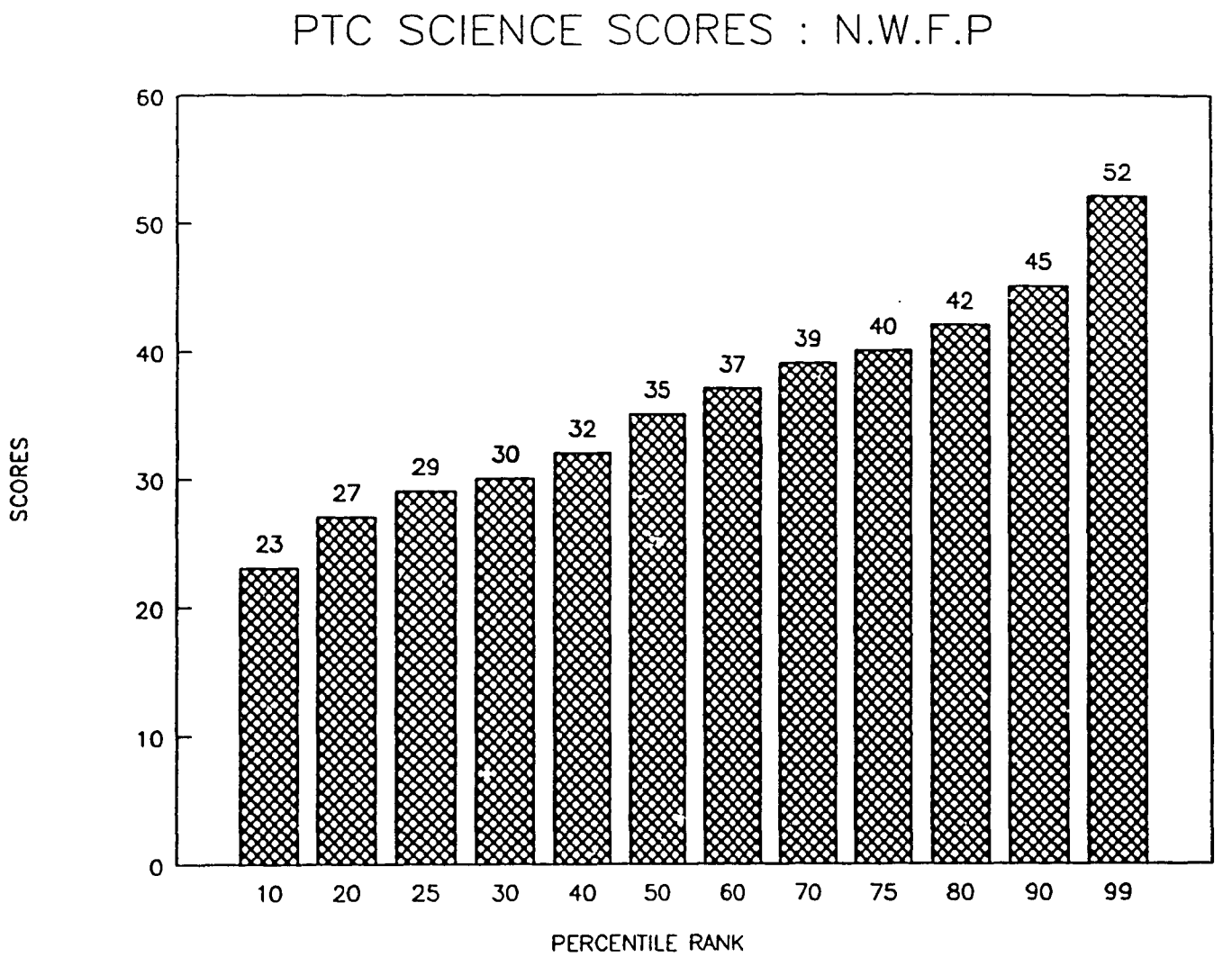
PTC SCIENCE SCORES : N.W.F.P
SCORES
0
10
20
30
40
50
60
23
27
29
30
32
35
37
39
40
42
45
52
10
20
25
30
40
50
60
70
75
80
90
99
PERCENTILE RANK

Figure. 5

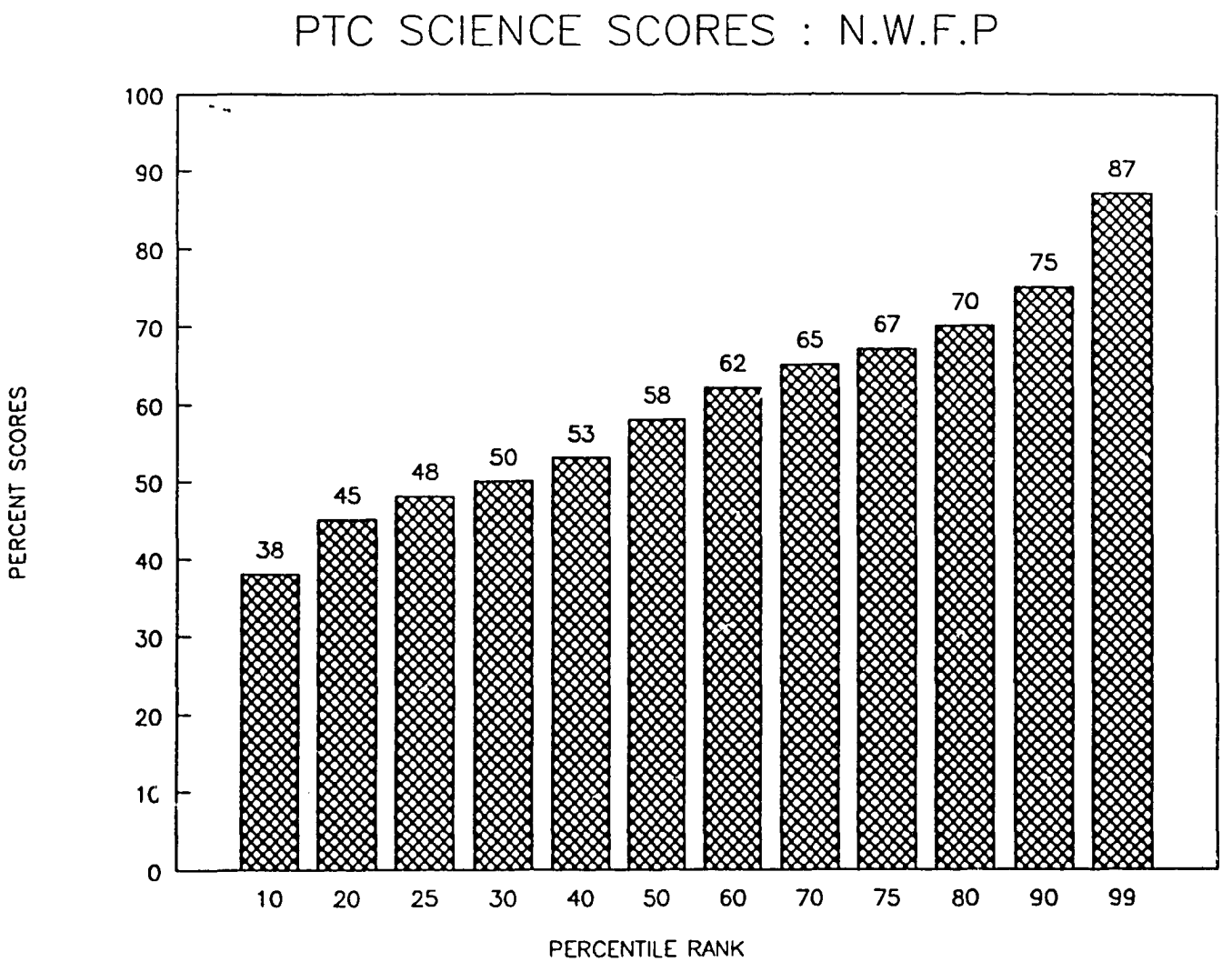
PTC SCIENCE SCORES : N.W.F.P
PERCENT SCORES
100
90
80
70
60
50
40
30
20
10
0
38
45
48
50
53
58
62
65
67
70
75
87
10
20
25
30
40
50
60
70
75
80
90
99
PERCENTILE RANK

Figure. 6

TABLE # 7

Summary Statistics for Science

Statistical Index	Value
No of Items	60
No of Students	1684
Mean	34.11
Variance	77.04
Std.Dev	8.78
Skew	-0.35
Kurtosis	0.08
Minimum	1.00
Maximum	56.00
Median	35.00
Alpha	0.85
SEM	3.42
Mean p	0.57
Mean point-biserial	0.32
Mean biserial	0.42

TABLE # 8

Analysis of Gender Differences for the Science Test

	Mean	%	S.D	n	F	P
Overall	34.11	57	8.78	1684		
Males	36.54	61	8.42	1038		
Gender					236.17	<0.001
Female	30.21	50	7.92	646		

Urdu

The p-values of 46 items in the Urdu test are given in Table 9. Using the criterion of 50% success rate, there are only four items which have

Insert Table 9 about here

p-value less than 50%. These items refer to concepts such as word meaning (interpretive synonyms), word usage(grammar), and word usage(spelling). Following the arbitrary criterion of p-values greater than 80% indicating mastery, 29 items out of 46 (63%) fall in this category.

The frequency distribution of the p-values is given in Table 10. The

Insert Table 10 about here

frequency polygon and percentile graphs for the Urdu test scores are given in Figures 7, 8, and 9. These figures indicate relatively somewhat better

Insert Figures 7, 8, and 9 about here

performance on the Urdu test. Summary statistics are provided in Table 11. The average score of the students is 36.47 (79%) with a standard deviation of

Insert Tables 11 & 12 about here

5.76. The median value of the scores is 38. The score distribution is negatively skewed (skewness= -2.13) indicating scores to be clustering towards

TABLE # 9

PERCENT SUCCESS-RATE ON URDU ITEMS (K=46) BY PROVINCE (N.W.F.P)

PTC STUDENTS

ITEM NO/CONCEPT	*P-VALUE*	*ITEM NO/CONCEPT*	*P-VALUE*
1.Word Meaning Interpretive Synonyms	77	14.Word Usage Grammar	89
2.Word Meaning Interpretive Synonyms	63	15.Word Usage Grammar	97
3.Word Meaning Interpretive Synonyms	92	16.Word Usage Grammar	97
4.Word Meaning Interpretive Synonyms	53	17.Word Usage Grammar	81
5.Word Meaning Interpretive Synonyms	62	18.Word Usage Grammar	65
6.Word Meaning Interpretive Synonyms	66	19.Word Usage Grammar	90
7.Word Meaning Interpretive Synonyms	47	20.Word Usage Grammar	49
8.Word Meaning Interpretive Antonyms	95	21.Word Usage Grammar	57
9.Word Meaning Interpretive Antonyms	84	22.Word Usage Grammar	88
10.Word Meaning Interpretive Antonyms	80	23.Word Usage Spelling	26
11.Word Meaning Interpretive Antonyms	89	24.Word Usage Spelling	44
12.Word Meaning Interpretive Antonyms	55	25.Word Usage Spelling	75
13.Word Usage Grammar	92	26.Word Usage Spelling	55

TABLE # 9 (Contd)

PERCENT SUCCESS-RATE ON URDU ITEMS (K=46) BY PROVINCE (N.W.F.P)

PTC STUDENTS

ITEM NO/CONCEPT	*P-VALUE*	*ITEM NO/CONCEPT*	*P-VALUE*
27.Reading Comprehension (Title)	85	37.Sentence Completion (Grammar)	96
28.Deduction	97	38.Sentence Completion (Grammar)	92
29.Deduction	94	39.Sentence Completion (Grammar)	91
30.Reading Comprehension (Title)	53	40.Sentence Completion (Grammar)	86
31.Deduction	94	41.Sentence Completion (Grammar)	97
32.Deduction	88	42.Sentence Completion (Grammar)	96
33.Sentence Completion (Grammar)	94	43.Sentence Completion (Grammar)	92
34.Sentence Completion (Grammar)	96	44.Sentence Completion (Grammar)	87
35.Sentence Completion (Grammar)	93	45.Structure :Meaningful Order	84
36.Sentence Completion (Grammar)	86	46.Structure :Meaningful Order	77

Overall Mean (NWFP)= 36.47, Percent = 79, S.D=5.76, N= 1685

Note: K indicates the total number of items in the test.

TABLE # 10

FREQUENCY DISTRIBUTION OF P-VALUE OF URDU ITEMS (n=1685)

INTERVAL	URDU (Items=46)	
	FREQUENCY	PERCENT
00 - 10	0	0
11 - 20	0	0
21 - 30	1	2
31 - 40	0	0
41 - 50	3	7
51 - 60	5	11
61 - 70	4	9
71 - 80	4	9
81 - 90	12	26
91 - 100	17	37

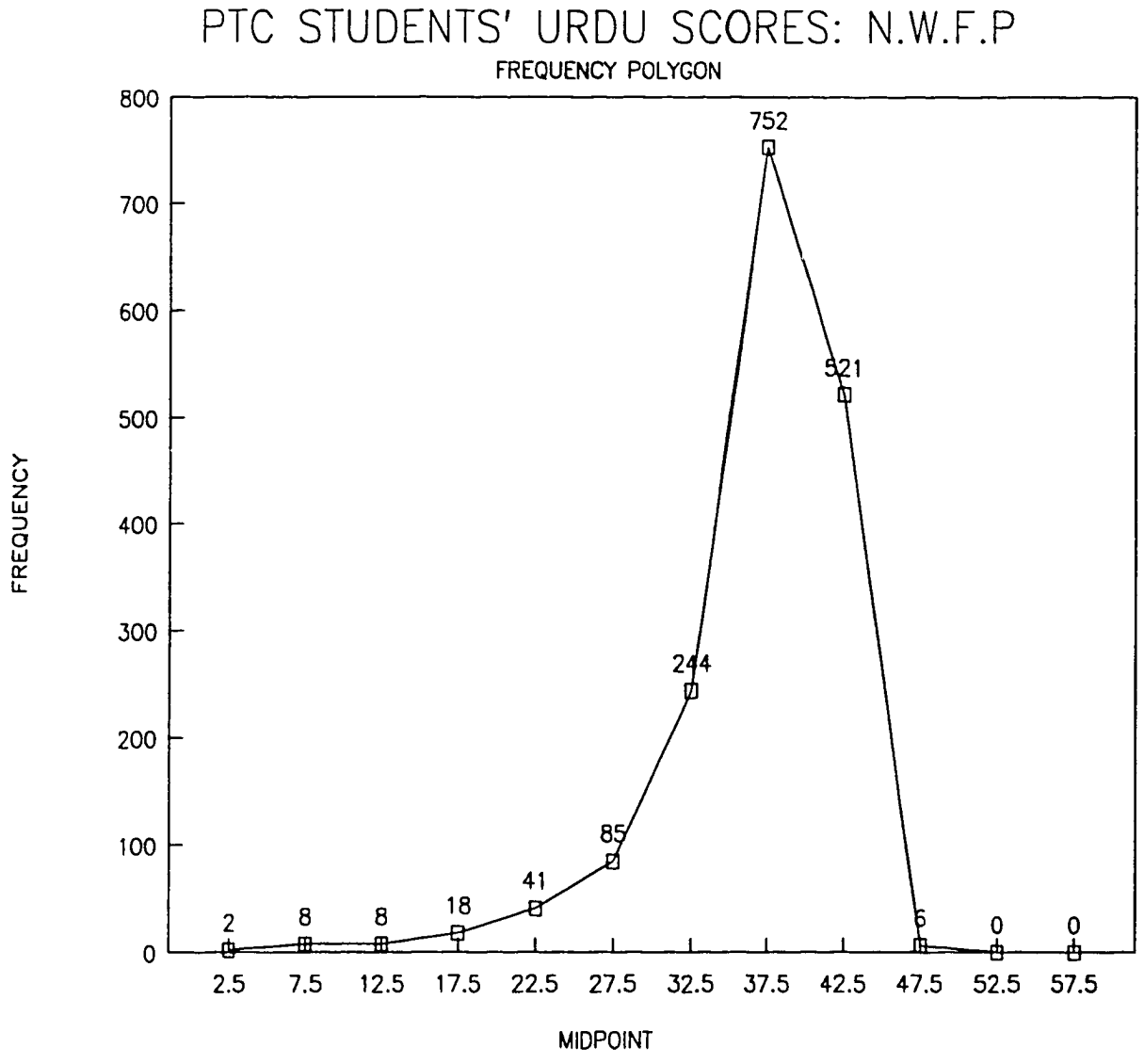
PTC STUDENTS' URDU SCORES: N.W.F.P
FREQUENCY POLYGON
FREQUENCY
800
700
600
500
400
300
200
100
0
752
521
244
85
41
18
2
8
8
6
0
0
2.5
7.5
12.5
17.5
22.5
27.5
32.5
37.5
42.5
47.5
52.5
57.5
MIDPOINT

Figure. 7

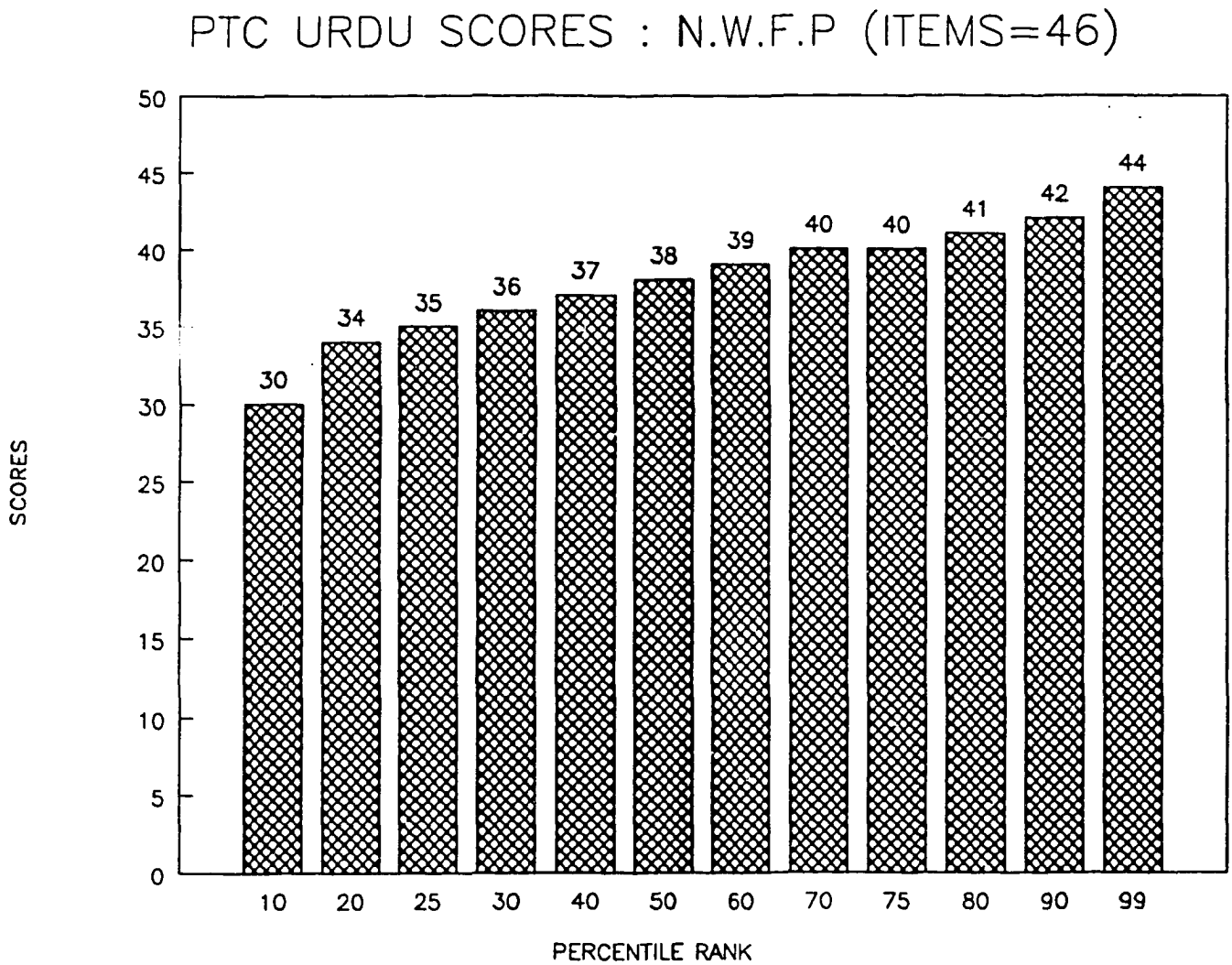
PTC URDU SCORES : N.W.F.P (ITEMS=46)
SCORES
50
45
40
35
30
25
20
15
10
5
0
30
34
35
36
37
38
39
40
40
41
42
44
10
20
25
30
40
50
60
70
75
80
90
99
PERCENTILE RANK

Figure. 8

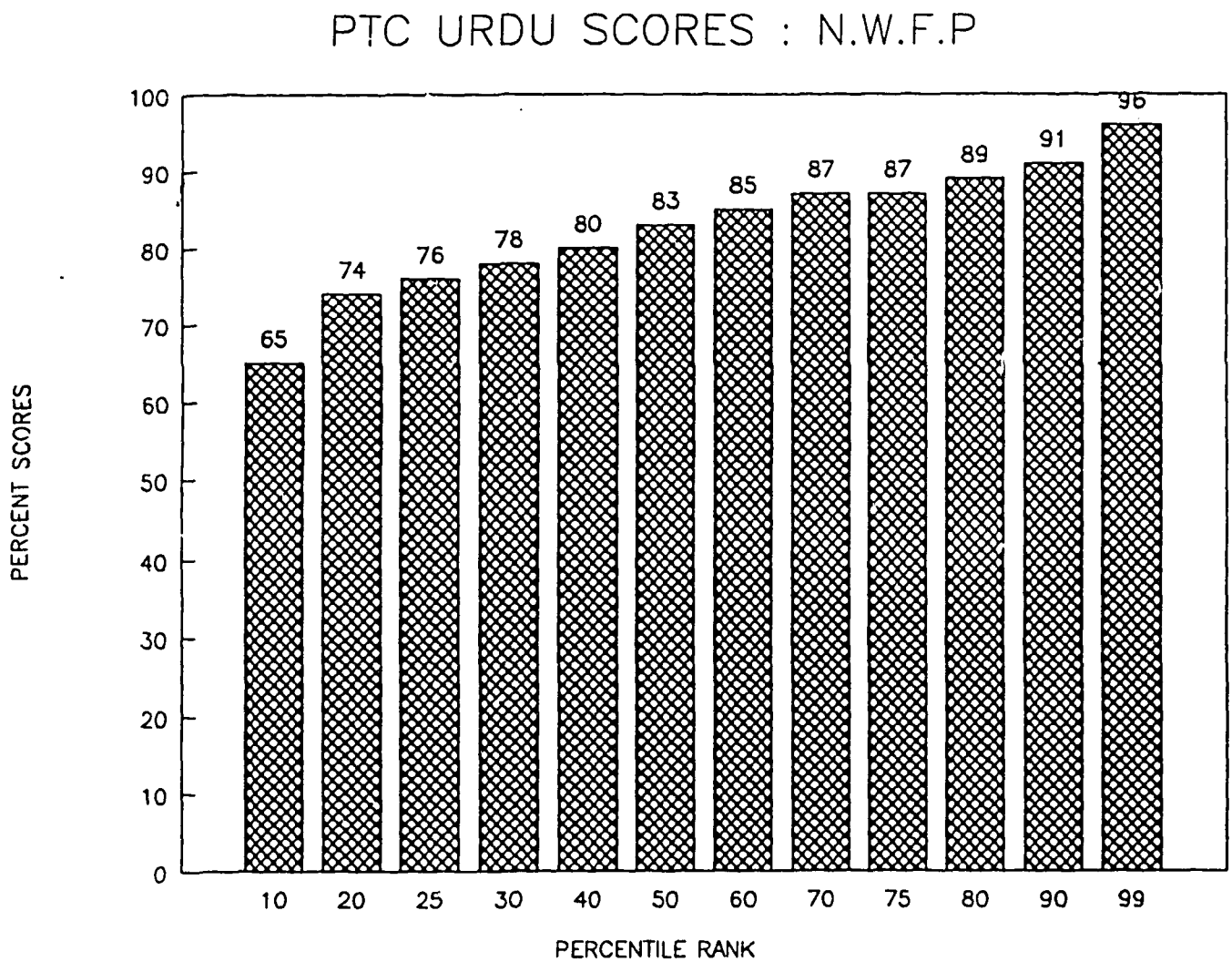
PTC URDU SCORES : N.W.F.P
PERCENT SCORES
100
90
80
70
60
50
40
30
20
10
0
65
74
76
78
80
83
85
87
87
89
91
96
10
20
25
30
40
50
60
70
75
80
90
99
PERCENTILE RANK

Figure. 9

TABLE # 11

Summary Statistics for Urdu

Statistical Index	Value
No of Items	46
No of Students	1685
Mean	36.47
Variance	33.17
Std.Dev	5.76
Skew	-2.13
Kurtosis	6.47
Minimum	0.00
Maximum	46.00
Median	38.00
Alpha	0.83
SEM	2.34
Mean p	0.79
Mean point-biserial	0.39
Mean biserial	0.66

TABLE # 12

Analysis of Gender Differences for the Urdu Test

	Mean	%	S.D	n	F	p
Overall	36.47	78	5.76	1685		
Male	36.80	80	5.61	1039		
Gender					9.00	<0.005
Female	35.94	78	5.95	646		

the right-end of the distribution. The average item-total test correlation in the form a biserial correlation is 0.66 indicating that, in general, each item is measuring the same skill as the total test. The internal consistency reliability coefficient is 0.83 which is within the acceptable range.

There is a significant difference, although slight, in the performance of male and female students. The average score for the males is 36.80 with a standard deviation of 5.61 (n=1039) and the average score for females is 35.94 with a standard deviation of 5.95 (n=646). The male students, in general, have performed better than female students ($F_{(1,1683)}=9.00$; $p<0.005$)

Pashto

The pashto test was given in those colleges which were located in predominantly Pashto-speaking areas. There are nine such colleges. Urdu-speaking students in these colleges were excused from taking the Pashto test. As a result, the test was given to 754 students. The p-values for the pashto items are given in Table 13. There are only two p-values which are below 50%.

Insert Table 13 about here

These two values correspond to items on spelling (the use of aarab zabar, zer, pesh, etc.). Actually all four items toward the end items # (47-50) of the test on this concept have relatively low p-values; ranging from 21% to

TABLE # 13
PERCENT SUCCESS-RATE ON PASHTO ITEMS (K=50) BY PROVINCE (N.W.F.P)
PTC STUDENTS

ITEM NO/CONCEPT	*P-VALUE*	*ITEM NO/CONCEPT*	*P-VALUE*
1.Word Meaning Interpretive Synonyms	92	15.Word Usage Grammar	97
2.Word Meaning Interpretive Synonyms	78	16.Word Usage Grammar	97
3.Word Meaning Interpretive Synonyms	89	17.Word Usage Grammar	97
4.Word Meaning Interpretive Synonyms	84	18.Word Usage Grammar	71
5.Word Meaning Interpretive Antonyms	97	19.Word Usage Grammar	81
6.Word Meaning Interpretive Antonyms	98	20.Word Usage Grammar	96
7.Word Meaning Interpretive Antonyms	95	21.Word Usage Grammar	92
8.Word Meaning Interpretive Antonyms	91	22.Word Usage Grammar	95
9.Word Meaning Interpretive Antonyms	80	23.Word Usage Pronoun	93
10.Sentence completion (Reasoning)	89	24.Word Usage Pronoun	95
11.Sentence completion (Reasoning)	96	25.Word Usage Pronoun	96
12.Sentence completion (Reasoning)	96	26.Word Usage Pronoun	92
13.Sentence completion (Reasoning)	84	27.Word Usage Pronoun	93
14.Word Usage Grammar	98	28.Word Usage Pronoun	90

TABLE # 13 (Contd)

PERCENT SUCCESS-RATE ON PASHTO ITEMS (K=50) BY PROVINCE (N.W.F.P)

PTC STUDENTS

ITEM NO/CONCEPT	*P-VALUE*	*ITEM NO/CONCEPT*	*P-VALUE*
29.Reading Comprehension (Deduction)	95	40.Word Usage Grammar	73
30.Reading Comprehension (Deduction)	94	41.Word Usage Grammar	89
31.Reading Comprehension (Deduction)	94	42.Word Usage Grammar	92
32.Structure (Meaningful Order)	67	43.Word Usage Grammar	93
33.Structure (Meaningful Order)	59	44.Word Usage Grammar	90
34.Structure (Meaningful Order)	40	45.Word Usage Grammar	95
35.Structure (Meaningful Order)	77	46.Word Usage Grammar	93
36.Word Usage Grammar	94	47.Word Usage Spelling	21
37.Word Usage Grammar	95	48.Word Usage Spelling	32
38.Word Usage Grammar	93	49.Word Usage Spelling	64
39.Word Usage Grammar	93	50.Word Usage Spelling	50

Overall Mean (NWFP)=42.49 Percent=85 S.D=5.63 N=754

Note: K indicates the total number of items in the test.

64%. The frequency distribution of the p-values is given in Table 14. Most

Insert Table 14 about here

of the p-values are in the upper range.

The frequency polygon of the pashto test scores is given in Figure 10. It is negatively skewed (skewness=-2.77) indicating a very large number of students obtaining higher scores on this test. Score values corresponding to different percentile ranks are given in Figures 11 and 12. Only one-quarter

Insert Figures 10, 11, and 12 about here

of the students have scored below 41 (82%). In other words, three-quarter of the students have scored above 82%. The average score on the Pashto test is 42.49 (85%). The median value is 44. The average item-total correlation is 0.74 which is quite high indicating the validity of the items as a measure of Pashto language skills assuming that the test is of appropriate difficulty as a measure of concepts and skills in the class V curriculum. The alpha reliability coefficient is 0.86 indicating

Insert Table 15 about here

homogenous structure of the content and skills covered in the test. Male students, in general, have performed better than female students ($F_{(1,752)}$=44.65;

TABLE # 14

FREQUENCY DISTRIBUTION OF P-VALUE OF PASHTO ITEMS (n=754)

INTERVAL	PASHTO (Items=50)	
	FREQUENCY	PERCENT
00 - 10	0	0
11 - 20	0	0
21 - 30	1	2
31 - 40	2	4
41 - 50	1	2
51 - 60	1	2
61 - 70	2	4
71 - 80	5	10
8 - 90	8	16
9 - 100	30	60

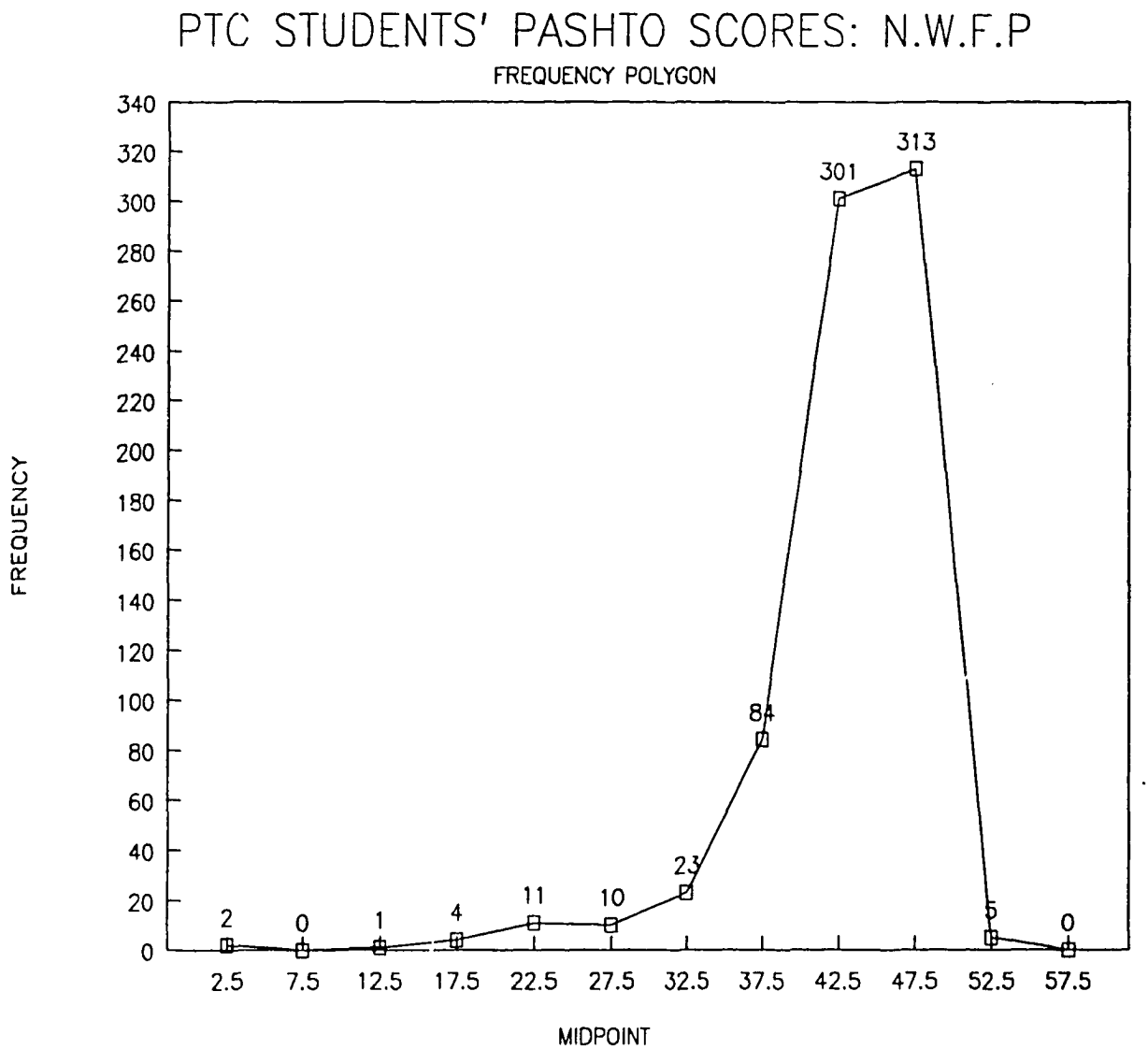
PTC STUDENTS' PASHTO SCORES: N.W.F.P
FREQUENCY POLYGON
FREQUENCY
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
0
2
0
1
4
11
10
23
84
301
313
5
0
2.5
7.5
12.5
17.5
22.5
27.5
32.5
37.5
42.5
47.5
52.5
57.5
MIDPOINT

Figure. 10

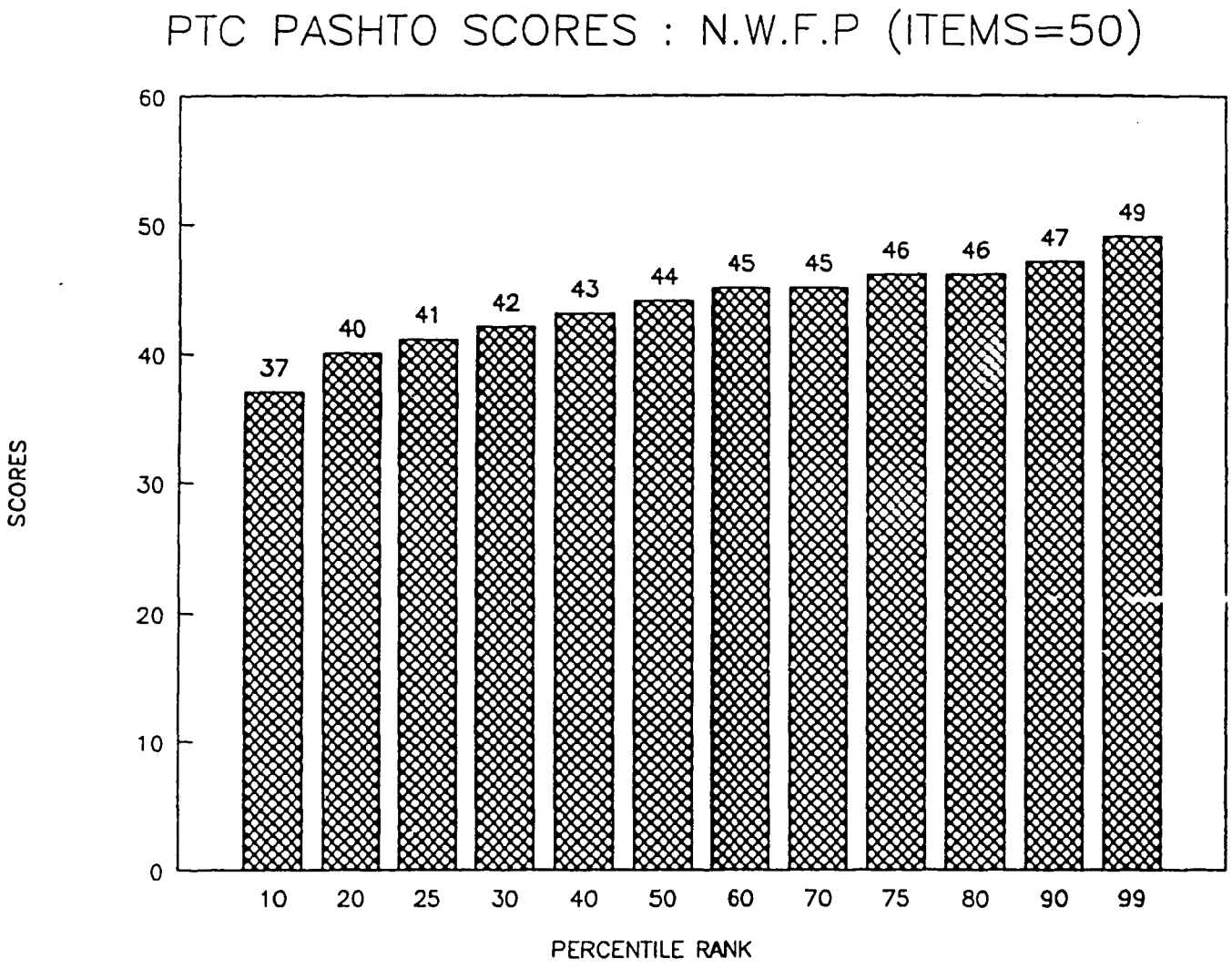
PTC PASHTO SCORES : N.W.F.P (ITEMS=50)
SCORES
60
50
40
30
20
10
0
37
40
41
42
43
44
45
45
46
46
47
49
10
20
25
30
40
50
60
70
75
80
90
99
PERCENTILE RANK

Figure. 11

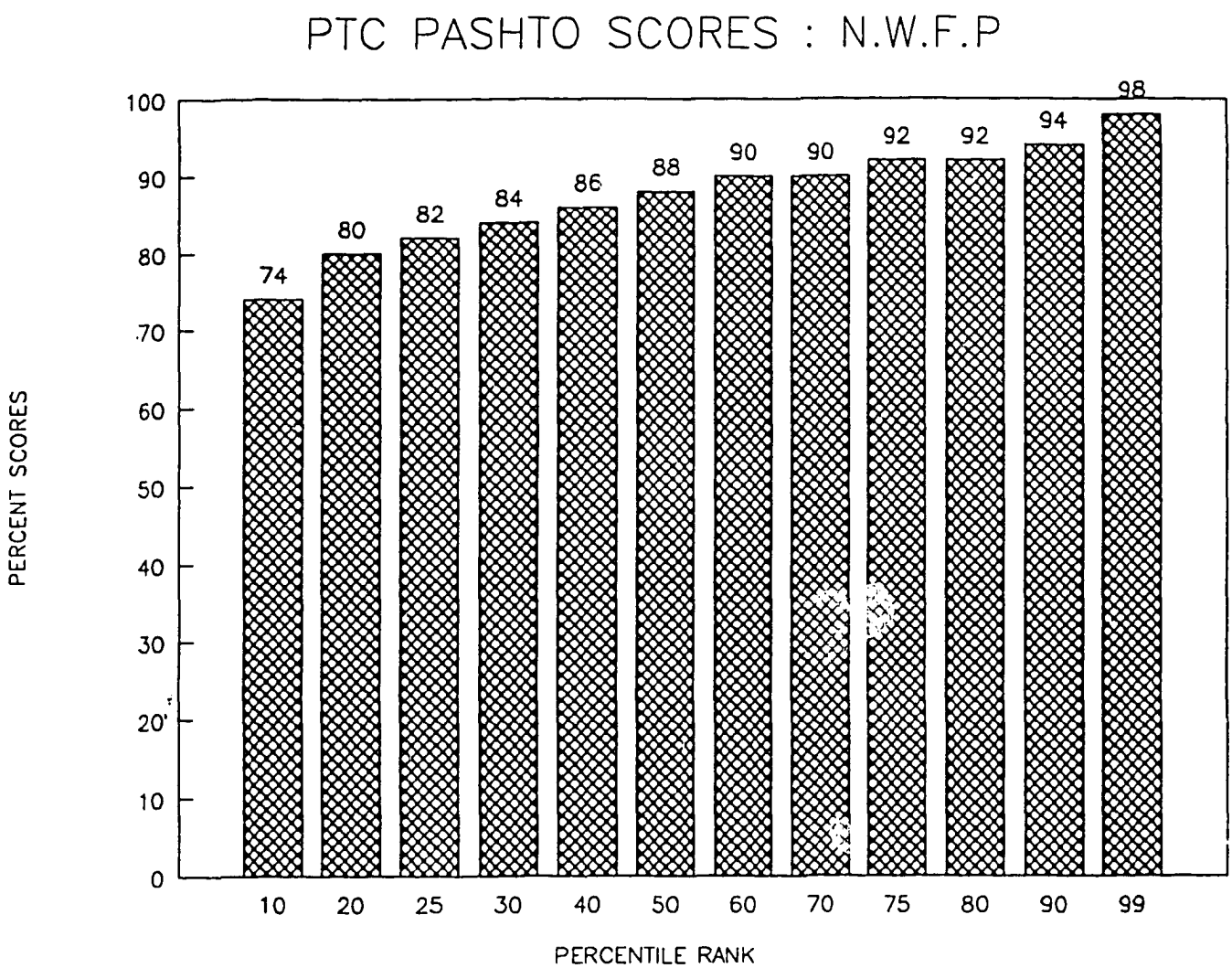
PTC PASHTO SCORES : N.W.F.P
PERCENT SCORES
100
90
80
70
60
50
40
30
20
10
0
74
80
82
84
86
88
90
90
92
92
94
98
10
20
25
30
40
50
60
70
75
80
90
99
PERCENTILE RANK

Figure. 12

TABLE # 15

Summary Statistics for Pashto

Statistical Index	Value
No of Items	50
No of Students	754
Mean	42.49
Std.Dev	5.63
Skew	-2.77
Kurtosis	12.33
Minimum	0.00
Maximum	50.00
Median	44.00
Alpha	0.86
Mean p	0.85
Mean point-biserial	0.40
Mean biserial	0.74

$p<0.001$). The average score for males (n=460) is 43.55 (87%) with a standard

---

Insert Table 16 about here

---

deviation is 4.61. The average score for females (n=294) is 40.82 (82%) with a standard deviation of 6.60.

TABLE # 16

Analysis of Gender Differences for the Pashto Test

	Mean	%	S.D	n	F	p
Overall	42.49	85	5.63	754		
Male	43.55	87	4.61	460		
Gender					44.65	<0.001
Female	40.42	82	6.60	294		

Overall analysis

Intercorrelations among Subject Scores

Table 17 includes the intercorrelations among the subject scores. The magnitude of these correlations attest to the concurrent validity of the tests. The correlation between mathematics and science is 0.67 which is quite larger than the correlations of mathematics with Urdu (r=0.47) and with Pashto (r=0.42). It should be expected because conceptually mathematics and science

Insert Table 17 about here

are closer to each other than mathematics and languages. On the other hand, the correlation of science with Urdu is 0.60 and with Pashto 0.42. The correlation of science with Urdu is higher than its correlation with Pashto because performance on the science test also involves comprehension of the Urdu language. The correlation between the two language score is 0.51. All these correlation are significant beyond 0.001 level.

Intercorrelations among College Averages

Correlations among the average scores for the colleges are given in Table 18. The magnitudes of the correlations indicate that, in general, if a college is

Insert Table 18 about here

TABLE # 17

Intercorrelation among Subject Scores of PTC Students

Variables	Math	Science	Urdu	Pashto
Math	1.00	0.68[1] (1684)[2]	0.47 (1684)	0.42 (753)
Science	0.68 (1684)	1.00	0.60 (1683)	0.42 (752)
Urdu	0.47 (1684)	0.60 (1683)	1.00	0.51 (753)
Pashto	0.42 (753)	0.42 (752)	0.51 (753)	1.00

[1] All correlations are significant beyond 0.001 level

[2] Number of Cases

TABLE # 18

Intercorrelations among the Colleges Averages of PTC students

Variables	MathMN	ScienceMN	UrduMN	PashtoMN
MathMN	1.00	0.94[1] (18)[2]	0.70 (18)	0.89 (9)
ScienceMN	0.94 (18)	1.00	0.76 (18)	0.91 (9)
UrduMN	0.70 (18)	0.76 (18)	1.00	0.81 (9)
PashtoMN	0.89 (9)	0.91 (9)	0.81 (9)	1.00

e.g., MN = Mean

[1] All correlations are significant beyond 0.05 level

[2] Number of Cases

weaker or stronger in one subject, it has a similar performance level in the other subjects. All correlations are significant 0.05 level.

Differences among Subject Scores

It was of interest to see if differences exited between the scores of the students on mathematics, science, and Urdu tests. Comparison with the Pashto scores was not done because the test was not given in all colleges. Table 19 includes the results using a paired t-test technique. Performance in mathematics is significantly lower than performance in science ($t_{(1683)}$= -10.80, $p<0.001$) and Urdu ($t_{(1683)}$= -64.41, $p<0.001$).

Insert Table 19 about here

As well, performance in science is significantly lower than performance in Urdu ($t_{(1683)}$=-74.93, $p<0.001$). The results indicates that mathematics was the most difficult test for the teacher-trainees followed by science.

Differences in Pashto and Non-Pashto Speaking Areas

In NWFP, there are nine colleges in Pashto-speaking areas (where the pashto test was administered (n=881) and nine colleges in non-pashto speaking areas (n=804). It was of interest to determine if there were differences between the performance of the students falling in these two groups. There were no significant differences in performance on mathematics and science tests. However, there was a significant difference in performance on the Urdu

TABLE # 19

Differences Between Subject Scores of PTC Students

		Mean	S.D	diff	t	p
Math		31.85	11.05			
	Science	34.11	8.78	-2.26	-10.83	<0.001
	Urdu[1]	47.57	7.51	-15.73	-64.41	<0.001
Science		34.11	8.78			
				-13.46	-74.93	<0.001
Urdu		47.57	7.51			

[1] The actual number of items in this test was 46. For the purpose of this analysis, the Urdu test was brought to the same scale as mathematics and science tests.

test ($F_{(1,1683)}=21.39$; $p<0.001$). The students from the colleges in non-pashto speaking areas performed significantly better (Mean=37.14) than students from colleges in Pashto-speaking areas (Mean=35.85).

Within-Gender Differences

Differences were also analyzed within the male colleges and female colleges (See Tables 20, 21, and 22). There are eleven male and seven female colleges. There were significant differences in performance on mathematics ($F_{(10,1028)}=12.49$; $p<0.001$),

---

Insert Tables 20, 21, and 22 about here

---

science ($F_{(10,1027)}=10.38$; $p<0.001$), and Urdu ($F_{(10,1028)}=18.12$; $p<0.001$). College Drosh (male) has performed significantly lower on mathematics when compared the remaining ten male colleges. As well, students in college Thana (male) did poorer on mathematics when compared to students from six of the ten colleges. In science, students in college Drosh (male) have performed significantly poorer than students in seven out of the ten colleges.

On the Urdu test, college Thana (male) has a significantly lower average score than the remaining nine colleges. Students in college Drosh (male) have done poorly on the same Urdu test compared to students in seven out of ten colleges. There were five male colleges ir which the Pashto test was administered. There were no significant differences in performance of the students from these five colleges on the Pashto test.

Similarly, differences were analyzed among the seven female colleges (see Tables 23, 24,

TABLE # 20

Analysis of Differences among Male Colleges for Mathematics[1]

College Name	Mean	%	S.D	n
Ghoriwala	36.62	61	9.35	120
Drosh	22.58	38	7.88	48
D.I.Khan	35.83	58	10.59	90
Karak	36.49	61	9.18	72
Kohat	32.47	54	9.62	72
Thana	30.61	52	11.91	145
Gulbahar	36.83	61	9.89	96
Gulbahar In-Service	37.02	62	10.27	132
Mathra	37.03	62	9.76	72
Barikot	37.45	62	11.54	84
Haripur	36.28	60	9.65	108

[1] $F_{(10,1028)} = 12.49;\ p < 0.0001$

TABLE # 21

Analysis of Differences among Male Colleges for Science[1]

College Name	Mean	%	S.D	n
Ghoriwala	37.06	62	7.73	120
Drosh	29.77	50	7.46	48
D.I.Khan	36.04	60	7.05	90
Karak	35.72	60	7.23	72
Kohat	35.42	59	7.72	72
Thana	32.89	55	10.80	144
Gulbahar	39.60	66	7.45	96
Gulbahar In-Service	39.30	66	8.43	132
Mathra	38.33	64	6.60	72
Barikot	36.70	61	7.00	84
Haripur	38.17	64	7.60	108

[1] $F_{(10,1027)} = 10.38$; $p < 0.0001$

TABLE # 22

Analysis of Differences among Male Colleges for Urdu[1]

College Name	Mean	%	S.D	n
Ghoriwala	36.66	80	5.72	120
Drosh	32.58	71	5.04	48
D.I.Khan	36.99	80	5.15	90
Karak	36.18	79	5.60	72
Kohat	37.79	82	4.54	72
Thana	32.74	71	8.46	145
Gulbahar	38.36	83	3.06	96
Gulbahar In-Service	38.77	84	3.10	132
Mathra	38.24	83	3.66	72
Barikot	36.57	79	4.73	84
Haripur	39.28	85	3.63	108

[1] $F_{(10,1028)} = 18.12;\ p < 0.0001$

25, and 26). There were significant differences in performance on mathematics

---

Insert Tables 23, 24, 25, and 26 about here

($F_{(6,639)}=16.79$; $p<0.001$), science ($F_{(6,639)}=19.12$; $p<0.001$), Urdu ($F_{(6,639)}=37.93$; $p<0.001$), and Pashto ($F_{(3,290)}=9.21$; $p<0.001$). In all cases, the college located at Khawaza Khela, Swat, did poorer compared to the other colleges.

Differences Due to Academic Qualifications

Although there were eight academic groups in the sample, scores for six groups were analyzed because there was only one trainee holding an M.A. degree and two trainees holding an M.Sc degree. These three trainees were excluded from the analysis. There were significant differences (See Tables 27, 28, 29, and 30) for mathematics ($F_{(5,1646)}=124.04$; $p<0.0001$); science ($F_{(5,1645)}=127.89$, $p<0.0001$); Urdu ($F_{(5,1646)}=49.03$; $p<0.0001$), and Pashto ($F_{(5,738)}=14.72$, $p<0.0001$).

Insert Tables 27, 28, 29, and 30 about here

Trainees with matriculation certificate (arts) did poorly on the mathematics test compared to trainees holding other qualifications. Trainees holding an intermediate-level certificate with science and a bachelor's degree with science did significantly better than trainees with other qualifications. Similar results were obtained for science. There was one exception: trainees with matriculation(science) did better than trainees with an F.A. certificate. In Urdu, trainees with matriculation (art)s did poorer than

TABLE # 23

Analysis of Differences among Female Colleges for Mathematics[1]

College Name	Mean	%	S.D	n
D.I.Khan	28.83	48	9.52	106
Kohat	26.34	44	8.39	68
Mansehra	29.73	50	9.10	120
Dargai	25.32	42	8.29	75
Dabgari Gate	30.13	50	8.92	88
Charsadda Road	30.39	51	8.76	90
Khawaza Khela	19.42	32	9.74	99

[1] $F_{(6,639)} = 16.79; p < 0.0001$

TABLE # 24

Analysis of Differences among Female Colleges for Science[1]

College Name	Mean	%	S.D	n
D.I.Khan	31.47	52	7.04	106
Kohat	30.53	51	6.50	68
Mansehra	31.92	53	7.64	120
Dargai	31.39	52	7.22	75
Dabgari Gate	29.66	49	7.51	88
Charsadda Road	33.29	55	6.47	90
Khawaza Khela	23.37	39	8.27	99

[1] $F_{(6,639)} = 19.12; p < 0.0001$

TABLE # 25

Analysis of Differences among Female Colleges for Urdu[1]

College Name	Mean	%	S.D	n
D.I.Khan	37.53	82	4.52	106
Kohat	35.94	78	4.78	68
Mansehra	38.17	83	3.95	120
Dargai	36.08	78	3.62	75
Dabgari Gate	36.10	78	5.21	88
Charsadda Road	38.30	83	3.11	90
Khawaza Khela	29.10	63	8.52	99

[1] $F_{(6,639)} = 37.93;\ p < 0.0001$

TABLE # 26

Analysis of Differences among Female Colleges for Pashto[1]

College Name	Mean	%	S.D	n
Dargai	41.97	84	6.91	75
Dabgari Gate	42.10	84	4.55	71
Charsadda Road	42.59	85	3.93	51
khawaza Khela	38.07	76	7.83	97

[1] $F_{(3,293)} = 9.20;\ p < 0.0001$

TABLE # 27

Analysis of Differences for Level of Qualifications for Mathematics[1]

Qualifications	Mean	%	S.D	n
Matriculation (arts)	24.25	40	9.66	455
Matriculation(science)	33.25	55	10.34	311
F.A	31.14	52	9.06	375
F.Sc	40.01	67	8.83	344
B.A	31.59	53	9.54	109
B.Sc	41.59	69	8.76	58

[1] $F_{(5,1646)} = 124.03; \quad p < 0.0001$

TABLE # 28

Analysis of Differences for Level of Qualifications for Science[1]

Qualifications	Mean	%	S.D	n
Matriculation (arts)	27.91	46	8.07	454
Matriculation(science)	35.40	59	7.97	311
F.A	33.32	56	7.02	375
F.Sc	40.48	67	6.81	344
B.A	35.48	59	7.21	109
B.Sc	41.45	69	6.11	58

[1] $F_{(5,1645)} = 127.89; \quad p < 0.0001$

TABLE # 29

Analysis of Differences for Level of Qualifications for Urdu[1]

Qualifications	Mean	%	S.D	n
Matriculation (arts)	33.37	72	7.64	455
Matriculation(science)	36.28	79	5.56	312
F.A	37.54	82	4.14	375
F.Sc	38.73	84	3.27	343
B.A	38.17	83	3.41	109
B.Sc	38.67	84	3.47	58

[1] $F_{(5,1646)} = 49.03;\ p < 0.0001$

TABLE # 30

Analysis of Differences for Level of Qualifications for Pashto[1]

Qualifications	Mean	%	S.D	n
Matriculation (arts)	40.17	80	6.47	204
Matriculation(science)	41.84	84	6.83	129
F.A	43.41	87	4.61	174
F.Sc	44.64	89	2.82	159
B.A	42.43	85	6.23	46
B.Sc	44.31	89	3.16	32

[1] $F_{(5,738)} = 14.72;\ p < 0.0001$

trainees with other qualifications. There were fewer significant differences among the academic groups on the Pashto test. Trainees with F.A./F.Sc and B.Sc did significantly better than trainees with matriculation (arts & science) and B.A. qualifications.

Summary

The results of the initial testing revealed serious weaknesses in the training of the PTC teachers in content preparation. These trainees do not even possess mastery of content at class V level, especially in mathematics and science. The situation is worse in female colleges than male colleges. There exists a serious problem in teacher-training which should be addressed immediately by studying its causes and implementing a systematic program of improvement. The quality of the teacher may not be the sole cause of falling standards in primary education but it is one of the important causes and a start for improvement can be made at the training stage.

Conclusions and Recommendations

Trainees in these training colleges have at least a matriculation certificate with higher marks in the matriculation examination. These trainees do include individuals with higher academic qualifications. The expectation from these trainees would normally be mastery of the content in the subject areas of the class V curriculum. That is, the distribution of scores in each content area would have been expected to be negatively skewed; most of the scores clustering toward the right end of the score continuum. Unfortunately, this did not occur in mathematics and science. Conclusions on the basis of the results of the initial assessment are:

1. Students, in general, are weak in mathematics and science;
2. Female trainees have performed poorly in all subject areas compared to their male counterparts;
3. Trainees with higher academic qualifications (especially from the science stream) have generally done better than trainees with lower academic qualifications;
4. Performance in various content areas is significantly correlated, that is students who are weak in one content area seem to be weak in other content areas also;
5. Performance of colleges (college being the unit of analysis) is highly correlated, that is, if a college is showing poor performance in one area, it is also showing poor performance in other content areas;
6. There are significant differences within male colleges and within female colleges. That is, some male colleges are better than others; similarly, some female

colleges are better than other female colleges;

7. There are no significant differences between the colleges (as a category) located in Pashto in speaking areas and non-Pashto speaking areas except on the Urdu test.

On the basis of the above conclusions, the following recommendations are offered:

1. Content upgrading especially in mathematics and science should become an integral part of the PTC curriculum;
2. Entrance qualifications to the PTC program should be upgraded, that is, the minimum qualifications may be the completion of an intermediate-level program with appropriate level of competency;
3. An admission test should be developed and administered to all applicants and the performance on the admission test should be used as one of the criteria for entry to the program;
4. Students enrolled in the PTC short-course and in the course offered by Allama Iqbal Open University should also be assessed to determine the level of their competency in the key content areas of the class V curriculum and compare performance of these students to students enrolled in the regular PTC program;
5. Teacher-training colleges should be regularly monitored to improve the training offered by these colleges;
6. The current duration of the PTC training seems to be too short for improving the content-knowledge level of trainees and pedagogical skills. Consideration should be

given to increasing the length of the course. As well, short-courses and private sitting in PTC and other professional examinations in education should be eliminated. The PTC course offered by Allama Iqbal Open University should be evaluated to determine if this program is meeting the objective of producing quality teachers;

7. A provisional certificate should be issued after the completion of the training program which should be converted into a permanent certificate only after the teacher has shown competence in teaching (content and method), say, over a two-year period;
8. A maintenance of certification process should be initialed whereby teachers would complete a re-certification examination every five years to determine whether they have sustained the same level of skills and competencies.
9. Regular in-service training should be available to active teachers who lack appropriate content knowledge and pedagogical skills and such teachers should be regularly monitored in the field after they have received such training.

APPENDIX - A

SUMMARY STATISTICS OF SCORES ON MATHS TEST (K = 60) FOR EACH COLLEGE

College Name (M&F)	Mean	Median	S.D	Range	rel	$\bar{p}$	$\widehat{pbis}$	$\widehat{bis}$	n
Ghoriwala (M)	36.62	36.00	9.31	17 - 58	0.88	0.61	0.35	0.48	120
Drosh,Chitral (M)	22.58	22.00	7.80	6 - 49	0.83	0.38	0.30	0.41	48
D.I.Khan (M)	35.83	36.00	10.54	14 - 58	0.91	0.60	0.39	0.53	90
Karak (M)	36.49	36.00	9.12	16 - 53	0.88	0.61	0.35	0.48	72
Kohat (M)	32.47	32.00	9.55	14 - 55	0.88	0.54	0.35	0.48	72
Thana,Malakand (M)	30.61	31.00	11.87	2 - 56	0.92	0.51	0.42	0.55	145
Gulbahar,Pesh (M)	36.83	38.00	9.84	16 - 54	0.90	0.61	0.37	0 51	96
Gulbahar In-ser(M)	37.02	37.00	10.23	14 - 55	0.91	0.62	0.39	0.50	132
Mathra, Pesh (M)	37.03	37.00	9.69	17 - 56	0.90	0.62	0.37	0.50	72
Barikot Swat (M)	37.45	40.00	11.48	8 - 55	0.93	0.62	0.44	0.60	84
Haripur (M)	36.28	36.00	9.60	11 - 55	0.89	0.60	0.36	0.48	108

APPENDIX - A (Contd)

SUMMARY STATISTICS OF SCORES ON MATHS TEST (K = 60) FOR EACH COLLEGE

College Name (M&F)	Mean	Median	S.D	Range	rel	$\bar{p}$	$\overline{pbis}$	$\overline{bis}$	n
D.I.Khan (F)	28.83	29.00	9.47	4 - 51	0.88	0.48	0.34	0.46	106
kohat (F)	26.40	25.00	8.33	8 - 47	0.84	0.44	0.30	0.39	68
Mansehra (F)	27.08	25.00	9.06	7 - 50	0.87	0.46	0.33	0.45	120
Dargai,Mkd (F)	25.32	24.00	8.23	5 - 47	0.85	0.42	0.30	0.41	75
Dabgari Gate (F)	30.14	29.00	8.88	13 - 53	0.86	0.50	0.32	0.43	88
Charsadda (F)	30.39	29.00	8.71	12 - 53	0.86	0.51	0.32	0.42	90
Khawaza Khela (F)	19.42	18.00	9.69	0 - 43	0.89	0.32	0.36	0.48	99

Note : K indicates the total number of items in the test,

and [n] indicates the number of students.

rel: indicates the internal consistency estimate of the reliability

APPENDIX - B

SUMMARY STATISTICS OF SCORES ON SCIENCE TEST (K=60) FOR EACH COLLEGE

College Name (M&F)	Mean	Median	S.D	Range	rel	$\overline{p}$	$\overline{pbis}$	$\overline{bis}$	n
Ghoriwala (M)	37.06	37.00	7.70	2 - 51	0.82	0.62	0.29	0.41	120
Drosh, Chitral (M)	29.77	29.00	7.38	14 - 47	0.79	0.50	0.27	0.36	60
D.I.Khan (M)	36.04	37.00	7.01	18 - 52	0.77	0.60	0.26	0.35	90
Karak (M)	35.72	36.00	7.18	13 - 48	0.78	0.59	0.27	0.36	72
Kohat (M)	35.42	36.00	7.66	18 - 52	0.81	0.59	0.28	0.39	72
Thana (M)	32.89	34.00	10.76	4 - 53	0.90	0.55	0.38	0.50	144
Gulbahar, Pesh (M)	39.60	40.00	7.41	17 - 56	0.81	0.66	0.28	0.38	96
Gulbahar In-Ser(M)	39.30	40.00	8.40	9 - 55	0.85	0.65	0.32	0.46	132
Mathta, Pesh (M)	38.33	39.00	6.55	21 - 52	0.75	0.64	0.25	0.34	72
Barikot (M)	36.70	38.00	6.95	8 - 48	0.77	0.61	0.26	0.36	84
Haripur (M)	38.17	39.00	7.57	14 - 54	0.81	0.64	0.29	0.40	108

APPENDIX - B (Contd)

SUMMARY STATISTICS OF SCORES ON SCIENCE TEST (K=60) FOR EACH COLLEGE

College Name (M&F)	Mean	Median	S.D	Range	rel	$\bar{p}$	$\overline{pbis}$	$\overline{bis}$	n
D.I.Khan (F)	31.10	32.00	7.01	12 - 49	0.75	0.52	0.25	0.34	106
Kohat (F)	41.36	30.00	6.45	15 - 47	0.71	0.51	0.24	0.32	68
Mansehra (F)	31.92	32.00	7.60	2 - 46	0.80	0.53	0.28	0.37	120
Dargai (F)	31.39	31.00	7.18	13 - 48	0.77	0.52	0.26	0.35	75
Dabgari Gate (F)	29.66	29.00	7.47	12 - 47	0.78	0.49	0.27	0.36	88
Charsadda Road (F)	33.21	33.00	6.37	20 - 48	0.71	0.55	0.23	0.31	90
Khawaza Khela (F)	23.37	23.00	8.23	1 - 45	0.83	0.39	0.30	0.39	99

Note : K indicates the total number of items in the test,
and [n] indicates the number of students.

rel : indicates the internal consistency estimate of the reliabilty

APPENDIX - C

SUMMARY STATISTICS OF SCORES ON URDU TEST (K=46) FOR EACH COLLEGE

College Name (M&F)	Mean	Median	S.D	Range	rel	$\overline{p}$	$\overline{pbis}$	$\overline{bis}$	n
Ghoriwala (M)	36.66	38.00	5.70	0 - 44	0.84	0.80	0.42	0.67	120
Drosh, Chitral(M)	32.58	33.00	4.99	18 - 42	0.73	0.71	0.28	0.42	48
D.I.Khan (M)	36.99	38.00	5.13	22 - 44	0.79	0.80	0.33	0.54	90
Karak (M)	36.18	37.00	5.56	20 - 44	0.82	0.79	0.36	0.59	72
Kohat (M)	37.79	38.00	4.51	20 - 45	0.76	0.82	0.35	0.62	72
Thana (M)	32.74	36.00	8.43	1 - 42	0.91	0.71	0.47	0.68	145
Gulbahar,Pesh (M)	38.36	39.00	3.05	26 - 45	0.52	0.83	0.24	0.42	96
Gulbahar In-Ser(M)	38.77	39.00	3.09	25 - 45	0.54	0.84	0.23	0.45	132
Mathra,Pesh (M)	38.24	39.00	3.63	27 - 44	0.64	0.83	0.24	0.43	72
Barikot (M)	36.57	37.00	4.70	18 - 43	0.78	0.79	0.33	0.58	84
Haripur (M)	39.28	40.00	3.62	23 - 45	0.67	0.85	0.30	0.56	108

APPENDIX - C (Contd)

SUMMARY STATISTICS OF SCORES ON URDU TEST (K=46) FOR EACH COLLEGE

College Name (M&F)	Mean	Median	S.D	Range	rel	$\overline{p}$	$\overline{pbis}$	$\overline{bis}$	n
D.I.Khan (F)	37.53	38.00	4.50	10 - 46	0.76	0.82	0.35	0.59	106
Kohat (F)	35.94	37.00	4.75	15 - 45	0.76	0.78	0.31	0.53	68
Mansehra (F)	38.17	39.00	3.94	14 - 44	0.72	0.83	0.35	0.62	120
Dargai (F)	36.08	36.00	3.59	24 - 43	0.58	0.78	0.22	0.35	75
Dabgari Gate (F)	36.10	37.00	5.18	20 - 43	0.79	0.78	0.35	0.60	88
Charsadda Road (F)	38.30	39.00	3.10	26 - 44	0.53	0.83	0.26	0.47	90
Khawaza Khela (F)	29.10	31.00	8.48	6 - 42	0.89	0.63	0.42	0.58	99

Note : K indicates the total number of items in the test,
and [n] indicates the number of students.
rel : indicates the internal consistency estimate of the reliability

APPENDIX - D

SUMMARY STATISTICS OF SCORES ON PASHTO TEST (K = 50) FOR EACH COLLEGE

College Name (M&F)	Mean	Median	S.D	Range	rel	$\overline{p}$	$\overline{pbis}$	$\overline{bis}$	n
Thana, Malakand(M)	43.30	44.00	5.38	12 - 50	0.87	0.87	0.43	0.73	103
Gulbahar Pesh (M)	43.53	45.00	4.67	19 - 50	0.83	0.87	0.38	0.69	93
Gulbahar In-ser(M)	44.18	45.00	3.50	30 - 50	0.70	0.88	0.25	0.52	108
Mathra ,Pesh (M)	43.24	45.00	4.98	23 - 50	0.84	0.86	0.38	0.65	72
Barikot, Swat (M)	43.37	44.00	4.32	22 - 49	0.78	0.87	0.33	0.61	84
Dargai,Malakand(F)	41.97	44.00	6.87	0 - 49	0.91	0.84	0.51	0.79	75
Dabgari Gate (F)	41.99	43.00	4.59	21 - 49	0.79	0.84	0.34	0.57	72
Charsadda Road (F)	42.59	43.00	3.89	29 - 49	0.72	0.85	0.30	0.55	51
Khawaza Khela (F)	38.07	40.00	7.79	0 - 49	0.90	0.76	0.45	0.70	97

Note : K indicates the total number of items in the test,
and [n] indicates the number of students.
rel : indicates the internal consistency estimate of the reliability