



Calhoun: The NPS Institutional Archive
DSpace Repository

Others Look at NPS

Articles and Reports about NPS (External)

1950-05-08

Validities of tests given in July 1949 and of
subtests of the Postgraduate School Aptitude
Test Form WNPF for predicting average
grades for first two terms at the U.S. Naval
Postgraduate School

Mollenkopf, William G.

Princeton, New Jersey; Educational Testing Service

<http://hdl.handle.net/10945/55836>

Downloaded from NPS Archive: Calhoun



Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

<http://www.nps.edu/library>

RESEARCH

BULLETIN

VALIDITIES OF TESTS GIVEN IN JULY 1949
AND OF SUBTESTS OF THE POSTGRADUATE SCHOOL APTITUDE TEST
FORM WNPF FOR PREDICTING AVERAGE GRADES FOR FIRST TWO
TERMS AT THE U. S. NAVAL POSTGRADUATE SCHOOL

William G. Mollenkopf

(Report submitted to U. S. Naval Postgraduate School)

Educational Testing Service

May 8, 1950

ABSTRACT

Scores on the seven aptitude and achievement tests given in July 1949 to members of the incoming class at the U. S. Naval Postgraduate School together with scores on the five parts of the Postgraduate School Aptitude Test Form WNPF have been related to the quality-point ratios of the officers in this class for their first two terms at the Postgraduate School. Bar charts have been prepared which show for each of four score intervals on each test what per cent of students earned quality-point ratios of 2.00 or higher. Product-moment coefficients of correlation between scores on each test and the criterion of quality-point ratios have also been computed.

Consideration of the results of these two methods of evaluating the effectiveness of the tests reveals that the most promising tests are those in mechanics, mathematics, physics, experimental science, and engineering. Comparison with previously obtained results shows that consistent results have been found where similar tests were given in 1948 and 1949. Certain of the new tests tried out in 1949 were also found to be worthy of further investigation.

It was concluded that the good progress made in the first year of research had been continued through the second year, and that by the end of the planned three-year program of research it would be possible to indicate an effective test battery for use as part of the means of selecting Postgraduate School students.

Validities of Tests Given in July 1949 and of Subtests of the
Postgraduate School Aptitude Test Form WNPF for Predicting Average
Grades for First Two Terms at the U. S. Naval Postgraduate School

Introduction

In July 1949 seven aptitude and achievement tests were given to the officers in the incoming class at the U. S. Naval Postgraduate School. Most of these officers had previously taken the U. S. Naval Postgraduate School Aptitude Test WNPF, which was sent out by the Bureau of Naval Personnel in the fall of 1948 to applicants for admission to the School. The present report presents the results obtained when scores on each of the seven tests and on each of the five parts of Test WNPF were correlated with average grades achieved by the officers in their first two terms of study at the Postgraduate School.

Descriptions of the Tests

Tests Given in July 1949. Seven tests were administered in July 1949. These were as follows:

1. Physics XNPA1. This test consisted of 80 multiple-choice questions and problems in first-year college physics. (Since a separate test in mechanics was being given, this physics test did not contain items on mechanics.) The time limit was one hour.

2. Experimental Science XNPA2. In this test seven simple experiments were described in some detail. Each description was followed by a set of ten statements pertaining to the experiment. A statement was to be classified under that one of the following categories which best described it:

- A. Agrees with the results of the experiment
- B. Contradicts the results of the experiment
- C. Not relevant to the experiment
- D. Justly criticizes the experiment
- E. Unjustly criticizes the experiment

There were 70 statements to be so classified. The time limit was one hour.

3. Mechanics XNPA3. This test contained 74 multiple-choice questions and problems in mechanics. The time limit was one hour.

4. Reading XNPA4. This was a speed of reading test introduced to supplement the reading comprehension measure also used on the same candidates. There were four seven-minute passages. Questions were interspersed in the passages read; the score was the number of these questions that were answered correctly.

5. Advanced Mathematics XNPA5. This test consisted of 60 multiple-choice items in college algebra, trigonometry, analytic geometry, differential calculus, and integral calculus. The time limit was two hours.

6. Interpretation of Data XNPA6. Items in this test required the test-taker to analyze data presented in the form of charts, tables, and graphs. There were 38 multiple-choice items in the test. The time limit was 70 minutes.

7. Engineering XNPA7. This test consisted of 62 multiple-choice items on engineering. The time limit was one hour.

Subtests of the Naval Postgraduate School Aptitude Test, Form WNPF. In the fall of 1948 arrangements were made through the Postgraduate School for the Bureau of Naval Personnel to send out an experimental test to each officer applying for admission to the Postgraduate School in the group to begin at the School in the summer of 1949. This Aptitude Test consisted of five parts, described below.

8. WNPF Part 1, Reading Comprehension. This test consisted of six passages of scientific reading material each of which was followed by a set of five questions to be answered on the basis of what was stated or implied in the passage. There were thus 30 questions in all in this part of the test.

9. WNPF Part 2, Mathematics Aptitude. This was a mathematical aptitude test consisting of 30 items chosen so as to minimize the effects on the score of length and recency of formal training in mathematics.

10. WNPF Part 3, Verbal Antonyms. This part contained 30 verbal antonyms items similar to those used in the verbal section of the College Board Scholastic Aptitude Test.

11. WNPF Part 4, Spatial Intersections. Each item in this test showed a drawing of a solid figure cut by a plane. The student was called upon to select from five figures presented him the one which correctly represented the shape on the cutting plane of the intersection between the solid figure and this plane. There were 50 items in the part.

12. WNPF Part 5, Elementary Physics. The items in this part were on topics covered in the usual high-school course in physics. There were 30 items in the part.

The Postgraduate School Aptitude Test WNPF was administered with one overall time limit of three hours.

Results of the Analyses

Two techniques have been employed for judging the effectiveness of a test in predicting the academic criterion of quality-point ratio for the first two terms at the Postgraduate School. First, the range of scores for each test was broken into four parts, and the per cent of the group having scores in each of these

quarters who had quality-point ratios of 2.00 or above was calculated. The resulting four percentages were used in preparing the bar charts presented as Figures 1-12. Secondly, there was computed for each test the coefficient of correlation between scores on the test and the quality-point ratios. These coefficients are given both below the captions of the bar charts, and grouped for convenience of comparison, in Table 1 at the end of the report.

In these charts a highly discriminating test is one which shows a large percentage of quality-point ratios of 2.00 or above for persons having scores in the top range on the test, and a progressively smaller percentage of quality-point ratios of 2.00 or above for persons having scores in each successively lower range of scores on the test, with the percentage for the lowest range of test scores quite small. Judged by this yardstick, the following tests have high discriminating ability:

- Test 1: Physics XNPA1
- Test 2: Experimental Science XNPA2
- Test 3: Mechanics XNPA3
- Test 5: Advanced Mathematics XNPA5
- Test 6: Interpretation of Data XNPA6
- Test 7: Engineering XNPA7
- Test 9: Part 2 WNPF, Mathematics Aptitude

The following tests were found to have fair discriminating ability:

- Test 8: Part 1 WNPF, Reading Comprehension (Note that the lowest bar represents data for only two persons.)
- Test 12: Part 5 WNPF, Elementary Physics

Reference to Table 1 at the end of the report shows that there is substantial agreement between the results of the two methods of evaluating the tests. The four tests shown by the charts to give the best discrimination are numbers 1, 3, 5, and 7. These same four tests (Physics, Mechanics, Advanced Mathematics, and Engineering) have the highest correlations with the criterion of quality-point ratios. Similarly, the test shown by the bar charts to be least effective is test number 4, Reading Speed. Scores on this test also show the lowest correlation with the criterion.

Combinations of Tests

A method for selecting predictive variables and approximating regression weights developed (but not yet published) by Dr. Paul Horst, Director of Research for the Educational Testing Service, was used for determining what combination of the twelve tests would as a group give the best prediction of the criterion. Like the better-known Wherry-Doolittle test-selection method, Dr. Horst's procedure corrects for the chance error added as each successive test is brought into the combination, and shows when no further test additions are advisable.

When Dr. Horst's method was applied to the correlations and validity coefficients represented in Table 1, it was found that the best combination of tests consisted of the following four:

Test 3: Mechanics
 Test 5: Advanced Mathematics
 Test 2: Experimental Science
 Test 7: Engineering

When scores on these tests are properly weighted and summed, a multiple correlation of .66 is found with the criterion of quality-point ratios for the first two terms at the Postgraduate School.

A composite score was derived for each student from his scores on the four tests listed above. To compute the composite score, the student's score on test 2 was doubled, his score on test 3 was multiplied by 3, his score on test 5 was doubled, and to the sum of these products was added his score on test 7. (These weights reflect the relative importance of the tests in the composite, and were derived from the multiple-regression weights.) Figure 13 shows how well this weighted composite score relates to the per cent of students obtaining quality-point ratios of 2.00 and above. It is clear from a comparison of Figure 13 with previous figures that the composite score is appreciably more effective than any single test.

Comparison with 1948 Results

In the report to the School entitled "Validities of Tests Given in July 1948 for Predicting First Year Average Grades at the U. S. Naval Postgraduate School," dated November 1949, the tests shown to be most effective were the following:

Engineering Achievement Test in Mathematics
 Physics Test
 G.R.E. Advanced Test in Engineering
 Reading Comprehension Test
 Mathematics Test from CEEB Scholastic Aptitude Test
 Mathematics Test from CEEB Intermediate Test

The results for the two years are thus for the most part consistent insofar as similar tests were given. In both years mathematics, physics, engineering, and reading comprehension tests have been found to be quite effective. Some tests given in 1949 were entirely new, and of these the Mechanics and Experimental Science tests turned out to look quite promising. The Interpretation of Data Test was by itself rather effective, but it was found to be too highly related to other tests in the battery for it to add to the predictiveness obtained by use of the Mechanics-Mathematics - Experimental Science-Engineering combination.

Conclusions

The charts and other data presented in this report demonstrate (as was found to be true for tests given in 1948) that mathematics, physics, engineering, and reading comprehension tests are effective for predicting academic performance in the first two terms at the Postgraduate School. Of the new tests tried out in 1949, the Mechanics, Experimental Science, and Interpretation of Data tests were found to give good predictions of this criterion. Reading Speed was found ineffective.

Results of the analysis of the 1949 tests and of the parts of the Postgraduate School Aptitude Test Form WNPF will be used in setting up the final experimental battery to be given in August 1950. It seems only fair to conclude that the promising start made in 1948 has been continued with a good degree of success, and that the additional evidence to be gained from the administration and analysis of tests to be given in 1950 will, when added to results already at hand, enable a suitable test battery to be indicated for use as part of the means of selecting Postgraduate School students.

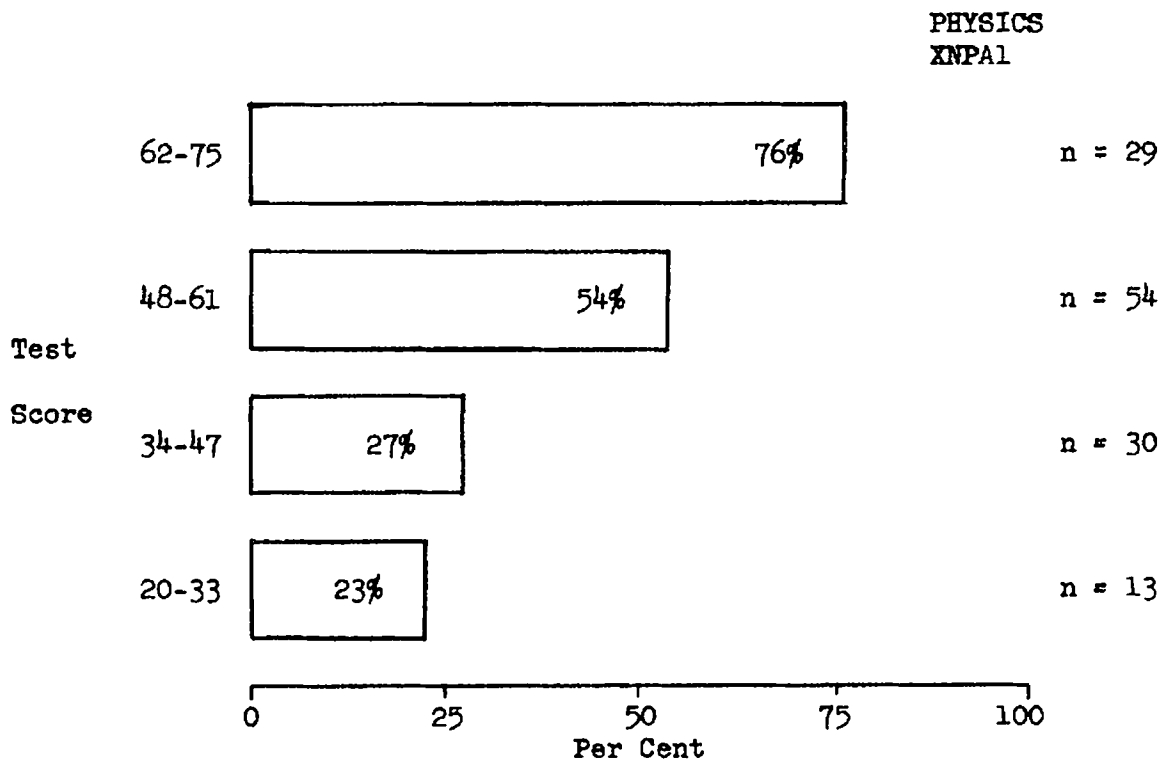


Figure 1. Per cent of students in each of four score ranges on the Physics Test XNPA1 with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 51.9 S.D. = 12.6 Test-Q.P.R. Correlation = .54

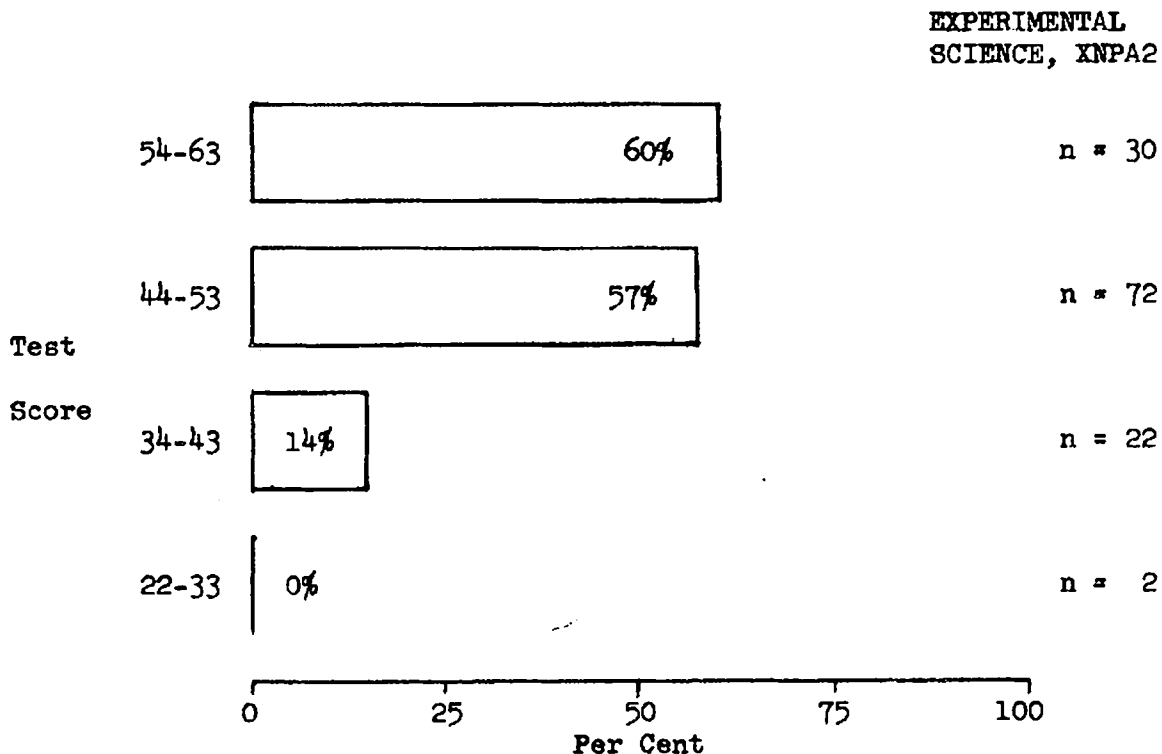


Figure 2. Per cent of students in each of four score ranges on the Experimental Science Test XNPA2 with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 48.5 S.D. = 6.6 Test-Q.P.R. Correlation = .49

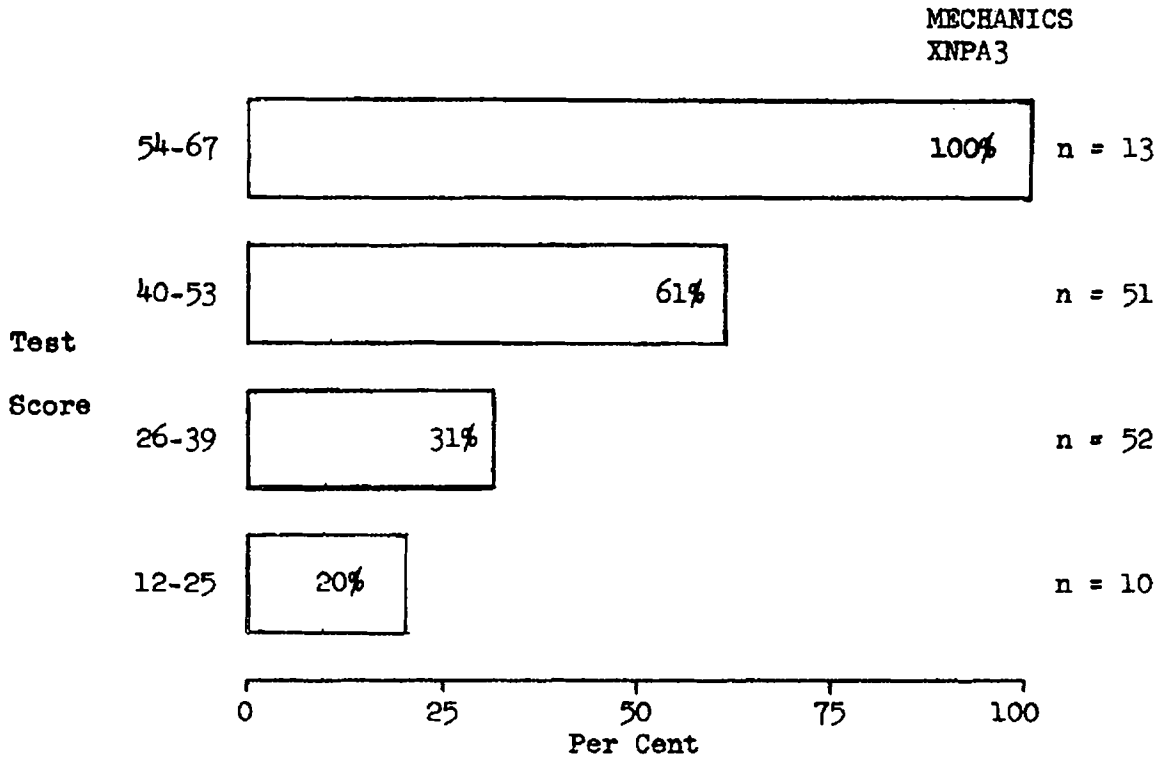


Figure 3. Per cent of students in each of four score ranges on the Mechanics Test XNPA3 with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 34.8 S.D. = 10.5 Test-Q.P.R. Correlation = .61

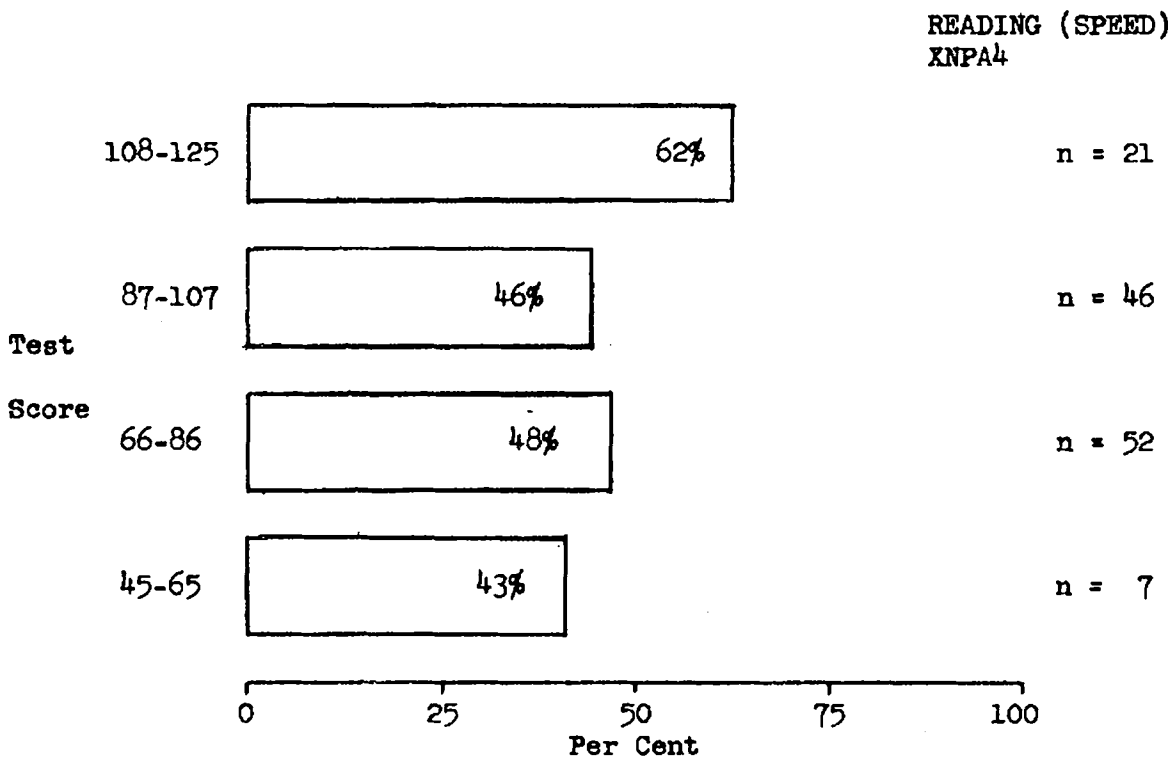


Figure 4. Per cent of students in each of four score ranges on the Reading (Speed) Test XNPA4 with Quality Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 89.4 S.D. = 16.2 Test-Q.P.R. Correlation = .16

ADVANCED MATHEMATICS, XNPA5

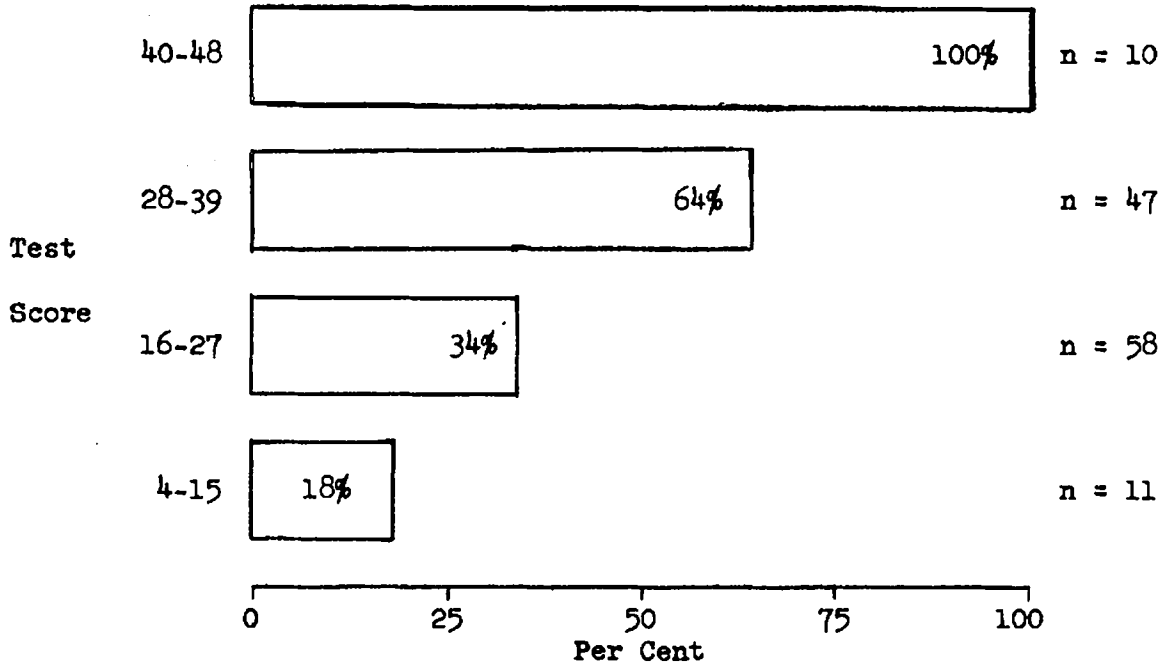


Figure 5. Per cent of students in each of four score ranges on the Advanced Mathematics Test XNPA5 with Quality Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 26.3 S.D. = 8.3 Test-Q.P.R. Correlation = .59

INTERPRETATION OF DATA, XNPA6

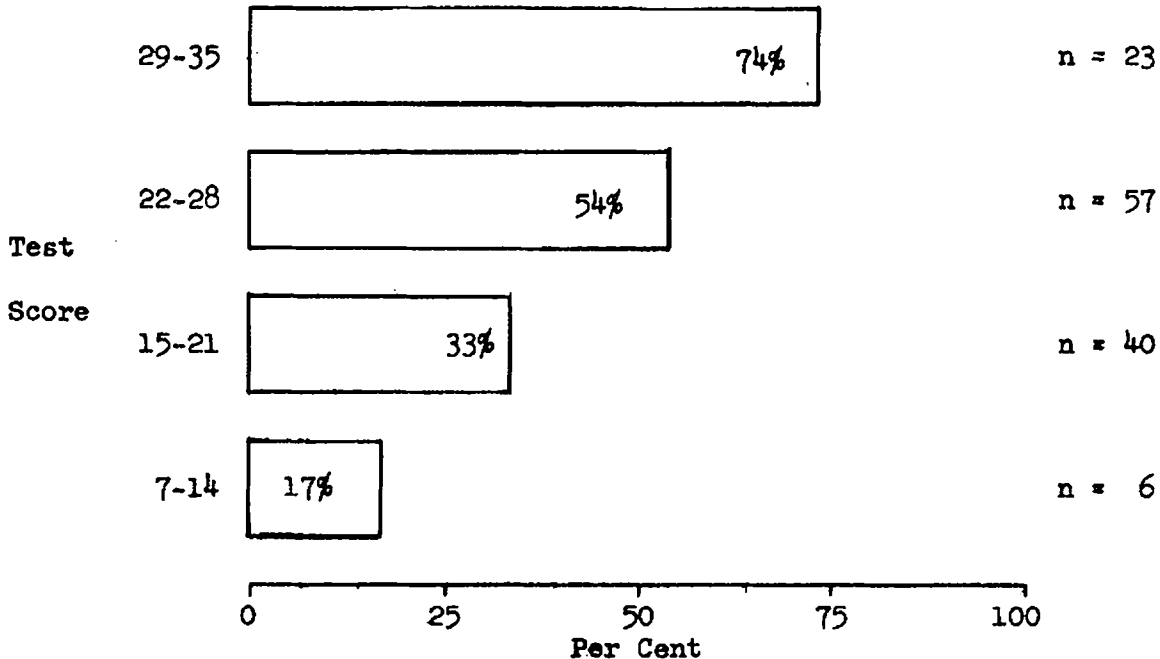


Figure 6. Per cent of students in each of four score ranges on the Interpretation of Data Test XNPA6 with Quality Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 23.1 S.D. = 5.5 Test-Q.P.R. Correlation = .47

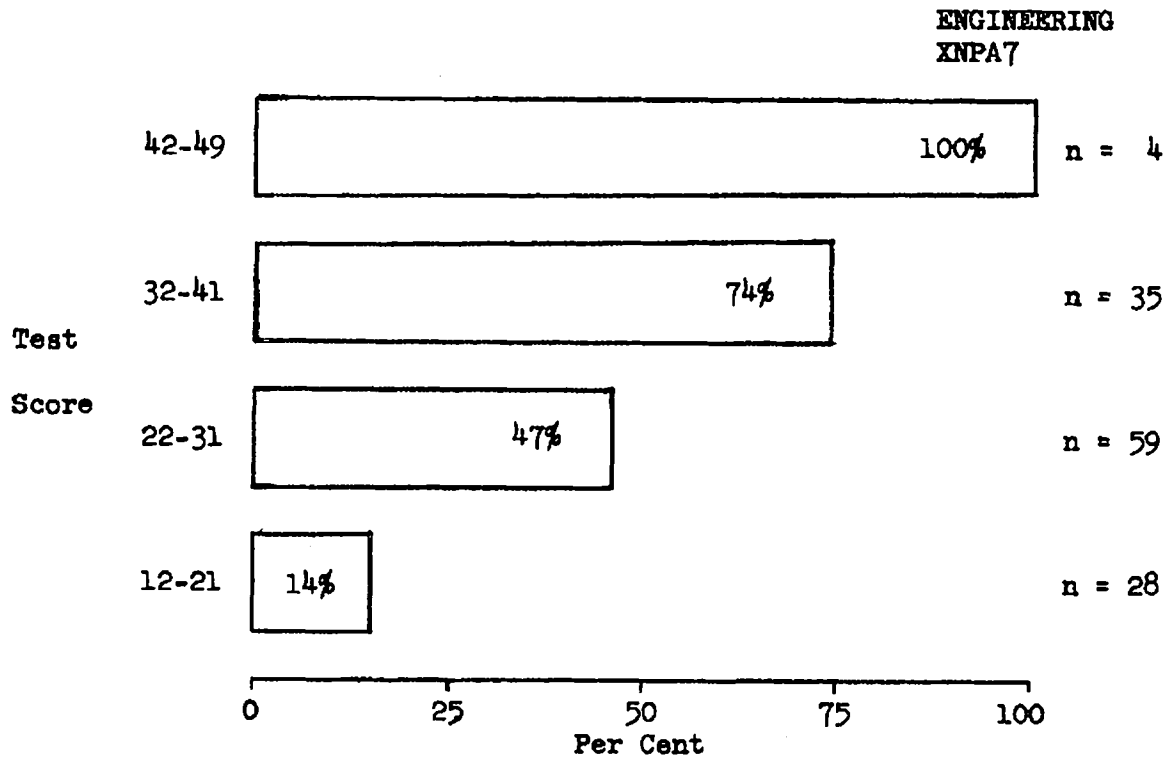


Figure 7. Per cent of students in each of four score ranges on the Engineering Test XNPA7 with Quality Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 27.7 S.D. = 7.6 Test-Q.P.R. Correlation = .56

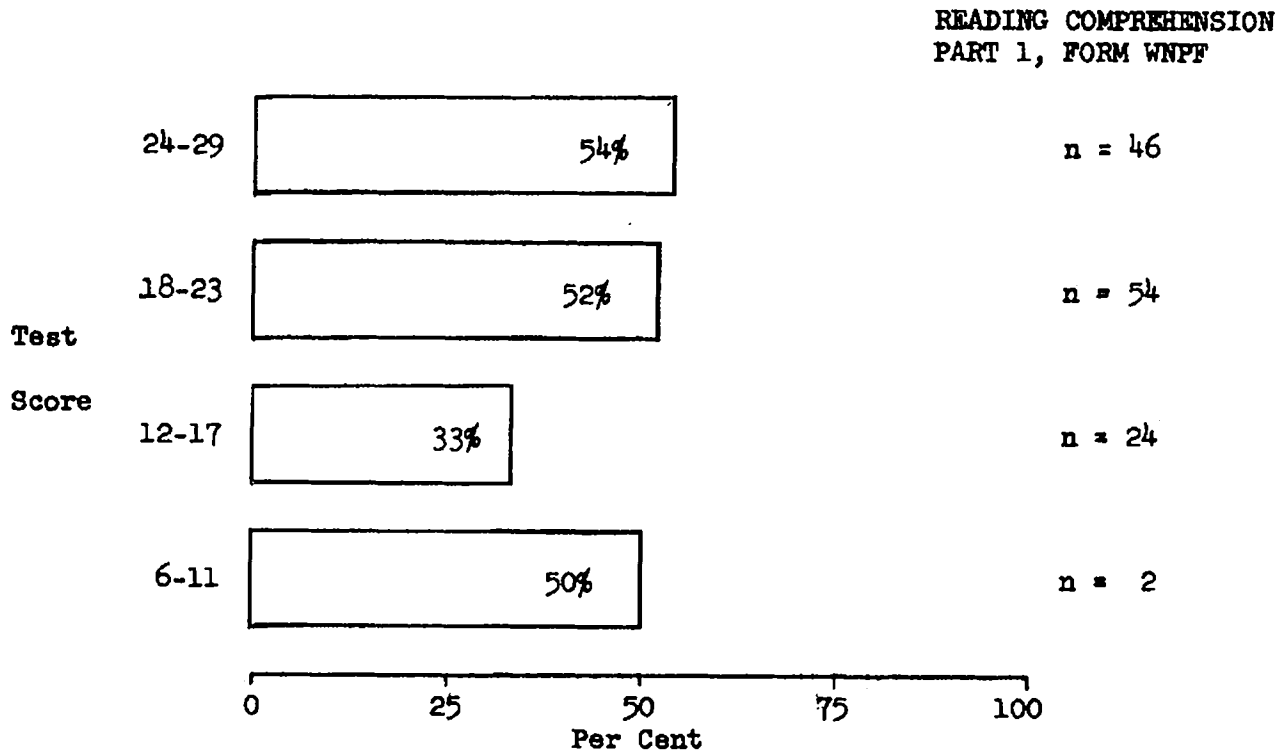


Figure 8. Per cent of students in each of four score ranges on the Reading Comprehension part of Form WNPF with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 21.1 S.D. = 4.6 Test-Q.P.R. Correlation = .37

MATHEMATICS APTI-
TUDE, PART 2, WNPF

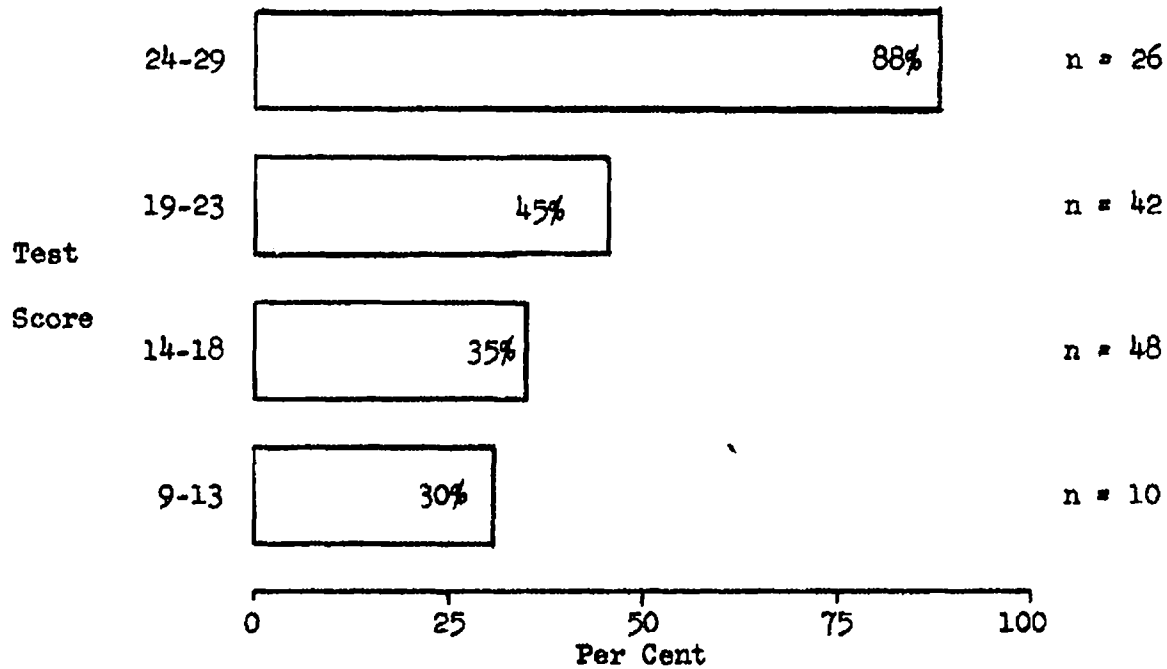


Figure 9. Per cent of students in each of four score ranges on the Mathematics Aptitude part of Form WNPF with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 19.6 S.D. = 4.5 Test-Q.P.R. Correlation = .46

VERBAL ANTONYMS
PART 3, WNPF

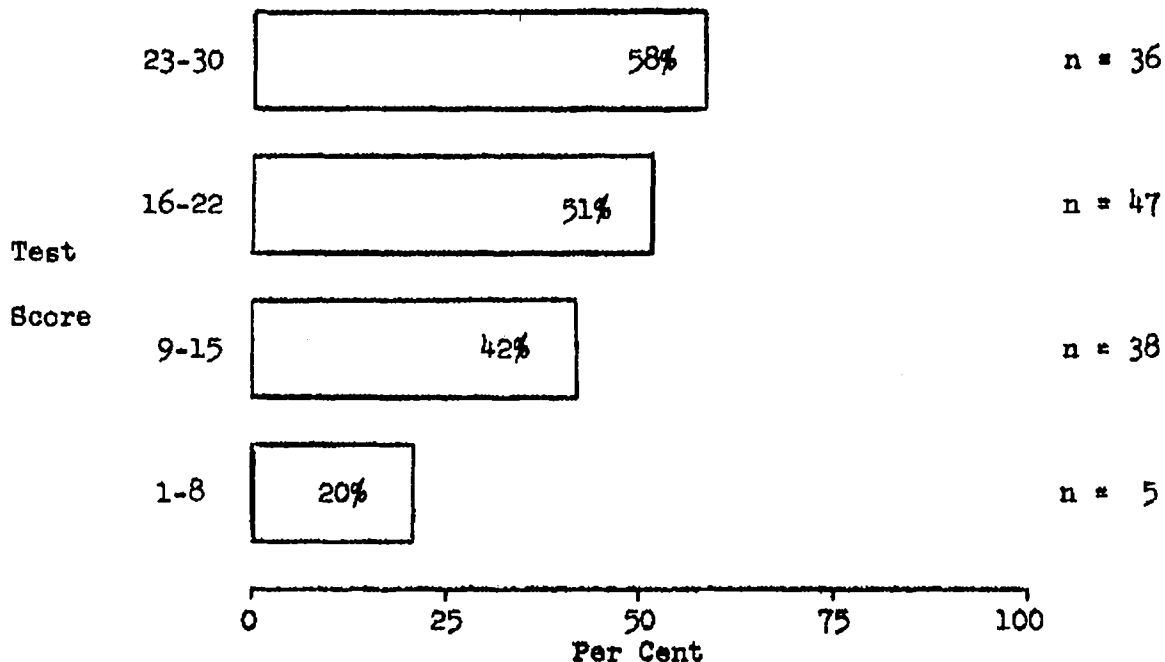


Figure 10. Per cent of students in each of four score ranges on the Verbal Antonyms part of Form WNPF with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 18.4 S.D. = 5.9 Test-Q.P.R. Correlation = .23

SPATIAL INTERSECTIONS
PART 4, WNPF

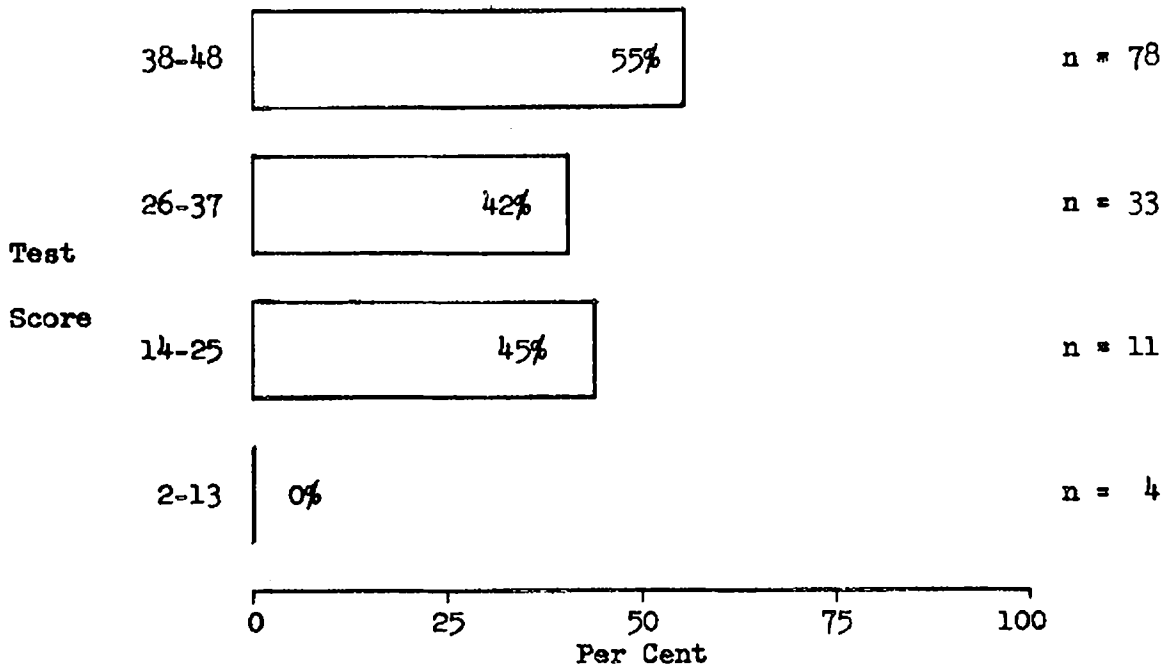


Figure 11. Per cent of students in each of four score ranges on the Spatial Intersections part of Form WNPF with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 37.2 S. D. = 9.5 Test-Q.P.R. Correlation = .44

ELEMENTARY PHYSICS
PART 5, WNPF

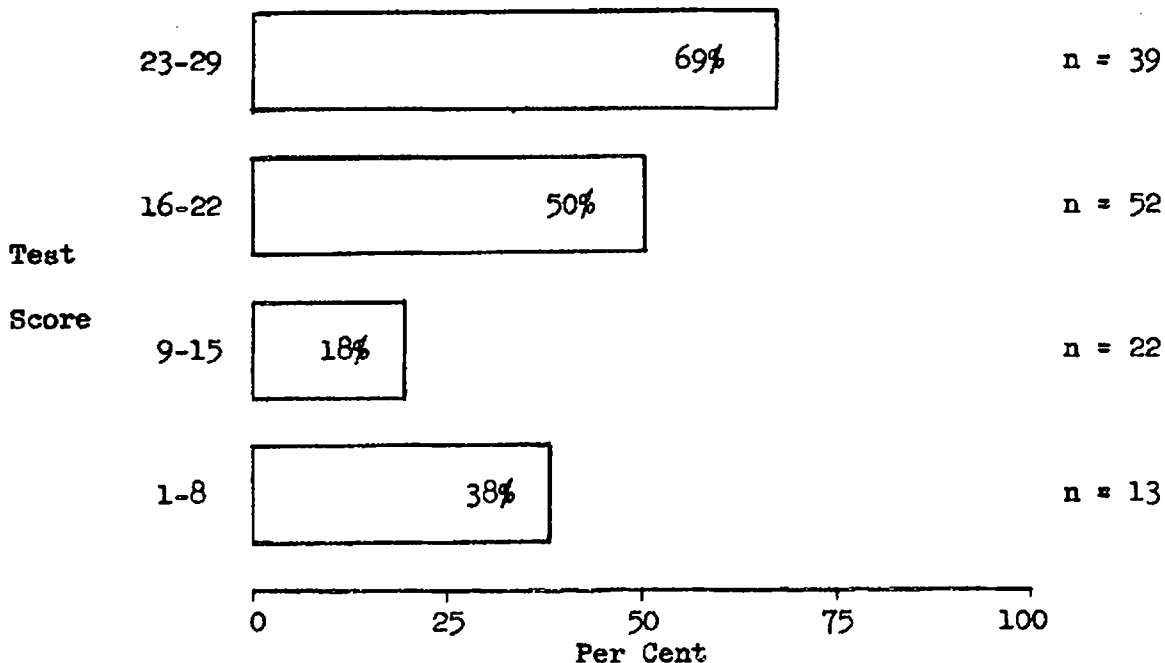


Figure 12. Per cent of students in each of four score ranges on the Elementary Physics part of Form WNPF with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Test Score = 18.4 S.D. = 6.8 Test-Q.P.R. Correlation = .48

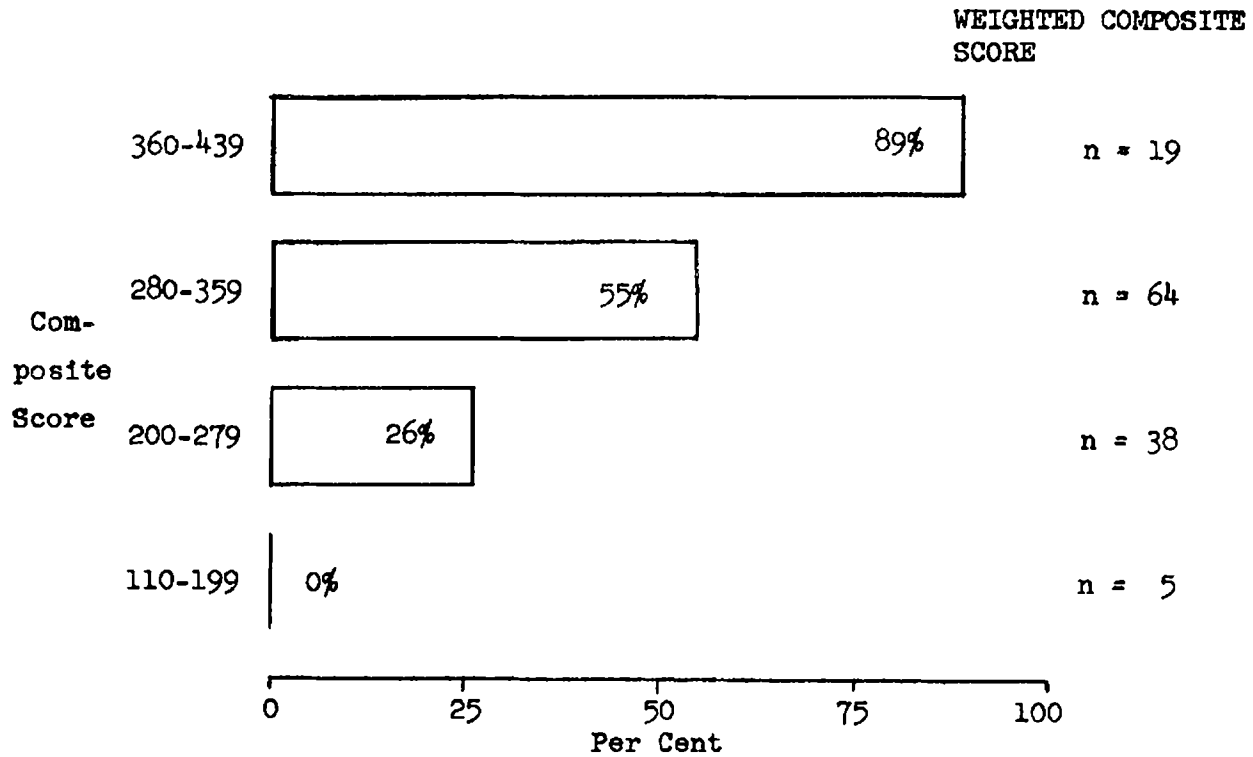


Figure 13. Per cent of students in each of four score ranges on the Weighted Composite Score with Quality-Point Ratios of 2.00 or above.

N = 126 Mean Composite Score = 298 S.D. = 60 Test-Q.P.R. Correlation = .67

TABLE 1

Correlations with Quality-Point Ratio (First Two Terms), Inter-correlations, Means, and Standard Deviations of Five Parts of the Aptitude Test WNPF and the Seven Tests of the July 1949 Battery
(N = 126)

	1 Physics	2 Exper. Science	3 Mechanics	4 Reading (Speed)	5 Advanced Mathematics	6 Interpretation of Data	7 Engineering	8 WNPF1 Reading Comp.	9 WNPF2 Math. Aptitude	10 WNPF3 Verbal Antonyms	11 WNPF4 Spatial Intersections	12 WNPF5 Elem. Physics	Q.P.R. - First Two Terms
1 Physics XNPA1	--	.66	.81	.27	.55	.60	.77	.48	.55	.43	.57	.74	.54*
2 Exper. Science XNPA2	.66	--	.58	.24	.44	.62	.56	.50	.48	.38	.41	.51	.49*
3 Mechanics XNPA3	.81	.58	--	.19	.70	.62	.74	.49	.69	.37	.59	.66	.61*
4 Reading (Speed) XNPA4	.27	.24	.19	--	.24	.15	.17	.30	.02	.46	.07	.13	.16
5 Adv. Math. XNPA5	.55	.44	.70	.24	--	.52	.57	.40	.62	.29	.40	.46	.59*
6 Interpretation of Data XNPA6	.60	.62	.62	.15	.52	--	.54	.49	.68	.43	.51	.55	.47*
7 Engineering XNPA7	.77	.56	.74	.17	.57	.54	--	.33	.54	.40	.45	.64	.56*
8 WNPF1 Reading Comprehension	.48	.50	.49	.30	.40	.49	.33	--	.48	.49	.27	.27	.37*
9 WNPF2 Math Aptitude	.55	.48	.69	.02	.62	.68	.54	.48	--	.31	.44	.48	.46*
10 WNPF3 Verbal Antonyms	.43	.38	.37	.46	.29	.43	.40	.49	.31	--	.24	.38	.23
11 WNPF4 Spatial Intersections	.57	.41	.59	.07	.40	.51	.45	.27	.44	.24	--	.50	.44*
12 WNPF5 Elem. Physics	.74	.51	.66	.13	.46	.55	.64	.27	.48	.38	.50	--	.48*
No. of Items	80	70	74	--	60	38	62	30	30	30	50	30	
Mean	51.9	48.5	34.8	89.4	26.3	23.1	27.7	21.1	19.6	18.4	37.2	18.4	
σ	12.6	6.6	10.5	16.2	8.3	5.5	7.6	4.6	4.5	5.9	9.5	6.8	

Mean of Q.P.R. = 1.91

Standard Deviation = 0.66

*A validity coefficient as great as this would occur by chance less than once in a hundred times when the true correlation was zero, for an N of 126.