TECHNIQUES FOR SECTIONING MATHEMATICS STUDENTS IN THE SAULT STE. MARIE RESIDENCE CENTER

by

WALTER WILLIAM FUNKENBUSCH

A THESIS

submitted to

OREGON STATE COLLEGE

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

August 1949

APPROVED:

[redacted]

Professor of Education

In Charge of Major

[redacted]

Head of Department of Education and

Chairman of School Graduate Committee

[redacted]

Dean of Graduate School

ACKNOWLEDGEMENTS

The writer wishes to make grateful acknowledgement to Dr. Riley Jenkins Clinton for his constructive criticism of this work, and for his invaluable counsel and guidance in the preparation of this thesis.

Grateful acknowledgement is also made to Dean Franklin R. Zeran for his encouragement and cooperation in the initial phase of this study.

A debt of gratitude is owed to Resident Director Russell, Professor Franklin Otis, and Professor Stephen Youngs of the Sault Ste. Marie Residence Center for their cooperation in making available the necessary data for this study.

TABLE OF CONTENTS

LIST OF TABLES

TECHNIQUES FOR SECTIONING MATHEMATICS STUDENTS IN THE SAULT STE. MARIE RESIDENCE CENTER

CHAPTER I

INTRODUCTION

The past decade has witnessed a phenomenal growth in the testing movement. This growth was bound to come in time, but the impetus of world conditions brought about its early emergence. The magnitude of the program carried on, either directly or indirectly, by the U. S. Government during the war years was without parallel.

There are at present many tests on the market, the validity and reliability of which are still open to question. Of the different types of tests, the predictive type is one of the more important, and of these the variety that measures a specific aptitude is the most practical. To obtain optimum benefits from using a test in this category, the following principles must be observed:

(1) A test appropriate for the problem at hand must be selected from the many available.

(2) The test must be properly administered and scored.

(3) The results must be correctly interpreted.

(4) Advisement to the testee must be on a personal basis.

Tolley (16, p. 86) wrote in 1948:

> On few other fronts has American Education made more progress in recent years than in the field of testing. Really phenomenal strides have been made in the development of accurate and valid instruments and techniques of measurement. In the course of this development so much light has been thrown upon individual differences and the learning process that sooner or later the testing movement will lead the way to sweeping educational reforms. Already education at every level has been affected in varying degrees. In scattered schools and colleges the change has been fundamental. For most institutions, however, it will be many years before the full impact is felt of the coming revolution of teaching, learning, and evaluation.

He also stated (16, p. 88):

> Despite allowance for the time lag, it is a little discouraging to find in so many schools that the major role of examinations is still the maintenance of standards. There was a time in American education when it was quite essential that primary attention be given to the setting, enforcement, and raising of scholastic standards. The need in our day, however, is for intelligent guidance of students. The purpose of tests is not to pronounce a judgment or to determine a grade, but to improve learning. Evaluation should be done with students rather than to them.

. . . . . .

(16, p. 94)

> One of our emerging needs is an agency to test tests. If the new Educational Testing Service can function as a national test standards bureau, it will render a most important service. About 1940 the American Council on Education held several regional

meetings to consider this matter among others.

Statement of the Problem

The problem of this study was the determination of the relationship that existed between the scores on two tests and the Algebra grades received by the students. The tests, which were administered at entrance or shortly thereafter at the Sault Ste. Marie Residence Center of the Michigan College of Mining and Technology, were:

(1) The Pre-Engineering Inventory by the Measurement and Guidance Project in Engineering Education of the Educational Testing Service (present organization), and

(2) The American Council on Education Psychological Examination prepared under the supervision of the Committee on Measurements and Guidance of the Council.

An additional project was the comparison of the achievement on Test IV - General Mathematical Ability - of the Pre-Engineering Inventory for a re-test given at the completion of two terms of work with the original achievement on the test.

Purpose of the Study

The purpose of this study was the determination of the value of the Pre-Engineering Inventory and the American Council on Education Psychological Examination in predicting probable success in either College Algebra or Remedial Algebra at the Sault Ste. Marie Residence Center of the Michigan College of Mining and Technology. This information was, provided it proved significant, to be used by the Mathematics Department of the Residence Center as a basis upon which to section our entering students into the College Algebra or Remedial Algebra groups. A non-credit Algebra course, Remedial Algebra or "Math. A," is offered for those students not having the entrance requirement of a unit and a half of high school Algebra. Students meeting this requirement take College Algebra, "Math. 1." It was felt by Prof. Otis, Mathematics Department Chairman, that a reliable placement examination given at entrance would be a much better criterion for sectioning. This view was shared by Prof. Youngs of the Department of Psychology, who is the present administrator of the above-mentioned tests given to most entering engineering students.

These two tests were not chosen by the writer

of this paper but were the tests which had been administered by the Veteran's Administration Representative and the Psychometrist at the Residence Center since its founding in October 1946. The writer feels, however, that these two used together form as good a combination as any available. In reply to a requested opinion from the U. S. Office of Education on the Pre-Engineering Inventory and the American Council on Education Psychological Examination, the writer received a letter, dated May 3, 1949, from Henry H. Armsby, Specialist in Engineering Education, Federal Security Agency, Office of Education, which stated in part:

> This is in reply to your letter of April 27, addressed to Director Russell.
>
> The two tests you mention are, I believe, the ones more generally used in engineering colleges, and of the two I believe the majority of institutions would prefer the pre-engineering inventory test.

The purpose in re-giving the Pre-Engineering Inventory, Test IV, General Mathematical Ability, at the end of two terms of work was to measure the change in the mathematical background of the students. As stated above, only the Mathematics section of the test was re-given since the Mathematics Department felt that this would be adequate for their purpose, and also to reduce to a

minimum the testing time which was deducted from regular classroom instruction.

Location of the Study

This study was carried on at the Sault Ste. Marie Residence Center of the Michigan College of Mining and Technology. This "Residence Center" was started in October 1946 in what until that time was an Army post, Fort Brady, by the Michigan College of Mining and Technology at Houghton as an annex to help care for the unprecedented large enrollment. The first two years work only are offered at the Center; the students, at the end of this period, transfer to the main branch at Houghton. During this period, 1946-49, our grading, especially as regards the Mathematics Department, came in for considerable scrutiny from the parent school at Houghton. This condition prompted the Mathematics Department at the Residence Center to begin this study during the past academic year.

Subjects Employed in the Study

The subjects employed in this study were 511 male[1] students, who took either College Algebra or Remedial Algebra at the Sault Ste. Marie Residence Center, since its inception in October 1946, and for whom both the Pre-Engineering Inventory and the American Council on Education Psychological Examination scores were available. There was a small number of students who had taken only one of the two tests. It would therefore have been possible to increase slightly the number of cases in either group, but it was felt that the gain of a few more cases would not outweigh the advantage of having the two groups identical, not only for a comparison of test scores but also for a possible combination of the two in a multiple correlation with Algebra scores.

The 118 students in the re-test project were from the 1948-49 portion of the entire 511.

---

1 Not by selection of cases by writer, but by natural process of male selection of an engineering curriculum.

Limitations of the Study

A major limitation of this study is the lack of high reliability in regard to instructor's marks. Contributing causes were:

(1) This being a new school, instructors as well as students were undergoing a period of psychological adjustment during the first year when facilities were not at a desirable standard. Many rooms were without blackboards until well into the term; there was hammering, and there were other disturbances in the halls.

(2) Concerning the unreliability of instructor's marks in general, E. G. Spacie, Student Counselor of the Michigan College of Mining and Technology, Houghton, stated, in a letter dated May 12, 1949 to Prof. Franklin Otis, Chairman of the Mathematics Department, Sault Ste. Marie Residence Center, the following:

> While I do not know what type of study you are making I can't help but echo Darley's statements concerning grade-test comparisons. He believes that modern test construction methods and the resulting high reliability coefficients have been thoroughly validated by the experience through the war years, that test validity and reliability have reached a degree of effectiveness where grades with their demonstrated unreliability become useless as criteria of test performance. He would hold that the usual study should be inverted, using the standard tests as criteria against grades.

(3) Since at the Sault Ste. Marie Residence Center, instructors give the Registrar numerical marks rather than letter grades, the frequency distribution has a "saw tooth" appearance. The writer feels that this tendency would be evident from an inspection of the grades issued at any institution where this practice prevails. In regard to this, the author stated in a letter dated April 2, 1949 to Dean Franklin R. Zeran, Oregon State College, that there are large numbers of 93's, 86's, 78's, 70's, and 65's, and very few scores of 92, 85, 77, and scores between 70 and 65. Our grades run as follows:

A	100-93
B	92-86
C	85-78
D	77-70
F	69- 0

The relative high frequencies of 93's, 86's, 78's, and 70's, are felt due to the tendency of the instructors to raise the numerical value of students' grades in order to put them in the next letter grade division. The high frequency of 65's is probably due to the fact that those grades in the range 66-69 that the instructor did not see fit to make into a 70, he moved back to a 65. Also there probably is a tendency to increase scores below 65 up to that value.

An additional limitation of the study is the fact that the two tests are not being used for the purpose for which they were intended. The American Council on Education Psychological Examination claims to measure aptitude for general scholastic success not success in College Algebra. The Pre-Engineering Inventory purports to measure the informational aspects present basic to pursuing the study of engineering, not the study of College Algebra. Although there is good reason for believing that the background necessary for studying Algebra and Engineering have much in common, they are not identical.

CHAPTER II

HISTORICAL BACKGROUND

There have been many studies of the predictive values of various tests, yielding a diversity of results. It has become common practice at most institutions of higher learning to subject the incoming student to a battery of scholastic aptitude tests during his orientation week.

Bloom and Allison (1, p. 210) wrote recently:

> Since the reorganization of the College of the University of Chicago in 1931, placement tests of some kind have been given new students during the orientation week to assist their advisors in planning appropriate programs of study for them.
>
> . . . . . .
>
> Placement tests which could be used to determine the appropriate level at which each student should start his work became necessary.
>
> . . . . . .
>
> (1, p. 212)
>
> Students enter the College of the University of Chicago after having attended any of several hundred schools and after having completed from ten to fifteen years of schooling. Some are as young as thirteen, while others are thirty.

Tyler (17, p. 336) in writing on the subject of predictive tests says:

> For many years, the criteria used to select students for college work have been: (1) completion of a high school program of study, (2) the taking of certain specified courses in high school usually referred to as college - preparatory courses, (3) certification that the student has done satisfactory work in these courses.
>
> For more than 30 years, evidence has been accumulating that these criteria are not wholly satisfactory bases for selecting good college students. For example, Professor Harl Douglass found no correlation between the subjects the students had taken in high school and their success in college. A similar result was also obtained in the Eight-Year Study of the Progressive Education Association. Furthermore, a great many investigations have found that the correlations between grades made in high school and grades made in college, altho positive, are usually less than .40.
>
> In spite of these findings, few colleges had changed their bases for admission until after the experience of World War II, which indicated that, in selecting personnel for various types of training and educational programs, a more efficient selection could be made by using a battery of tests than by depending upon the previous educational record of the student.
>
> The Eight-Year Study had shown that for the students in that investigation five criteria could be used to predict college success more efficiently. These were: (1) score on a scholastic-aptitude test such as the American Council of Education Psychological Test, (2) score on a reading test based on fairly complex reading materials of the sort used in college, (3) score on a test of writing skills, (4) score on a test of simple mathematical operations, largely

> arithmetic, and (5) evidence that the student had carried one subject for at least two years and had made better than average grades in that subject.
>
> This last criterion provides some indication of good work habits and motivation, but in place of using the average high school grade it selects the subject in which the student has been sufficiently motivated and well enough grounded to have carried it for two years. Using these criteria, correlations with college success ranging from .70 to .80 were obtained.

Johnson (4, pp. 19-20), however, takes a different view in regard to scholastic aptitude tests:

> In a study by the author the average scholastic-aptitude-test score of a large number of students disqualified for reason of poor scholarship was only <u>one</u> point lower than the average test score of those remaining in college; that is, there was practically no difference in the average aptitude-test scores of those who failed and those who succeeded.

The Study-Habits Inventory represents an attempt to measure study habits and aptitudes mentioned by Tyler above. Wrenn (20, pp. 1, 2) states:

> The Inventory is a weighted check list of specific study habits and attitudes which high scholarship and low scholarship groups of students possess in differing proportions.
>
> . . . . . .
>
> Many factors contribute to good academic achievement, such as general scholastic aptitude, subject-matter background, motivation or drive, and study habits and attitudes.

> It is these habits of work that determine the effectiveness of the other factors in the study situation.

Prediction of Success in Mathematics Courses

Once the student is on the campus, the practical need in counseling is not to predict general scholastic success but to measure his probability of success within a definite academic area or subject. According to Conant (2, p. 3), "It is generally admitted, too, that for a large majority of students entering college relatively simple tests can provide evidence as to aptitude for study of mathematics or for foreign languages."

Speaking of the demonstrated merit of placement examinations in individual subject fields, Tyler (17, pp. 336-337) observes:

> The common college practice is to enrol a student in the first course in a given field unless he has taken work in this field in high school. On the other hand, if his high school record shows that he has had courses in this subject, he is expected to take a more advanced course in college.
>
> . . . . . .
>
> Placement tests are most commonly used in English, the foreign languages, mathematics, and science, and in a few institutions they are given in all fields. On the basis of these tests, the student is advised as to

> whether he should register for an elementary or a more advanced course, or if he has not yet the equivalent of a complete elementary course, suggestions are made about how he may prepare for the examination in the elementary course without spending full time in class attendance.
>
> The use of placement tests has proved to have two important values.
>
> On the one hand, student failures have been reduced. At the University of Chicago, for example, failures in college have been reduced more than half by these placement tests since students are placed in courses where they are prepared to proceed effectively.
>
> In the second place, the use of placement tests has saved time for students and faculty by reducing the amount of duplication. Where placement tests are used in a comprehensive way, more than 80% of the students are able to demonstrate some competence worthy of recognition in more advanced work which was not on their previous educational record. The saving of time and the elimination of boredom for such a large fraction of the student-body is a very important contribution.

Thurston in 1927 (13, pp. 172-73) reported receiving the following information from Fryer concerning the correlation between the American Council on Education Psychological Examination and grades in College Mathematics.

Arithmetic Test Scores - - - - - - - - - -	.204
Number Completion Test - - - - - - - - - -	.077
Absurdities Scores - - - - - - - - - - - -	.236
Analogies II Scores - - - - - - - - - - -	.347
Arithmetic Test (Engineering Stud. Only)	.129
Analogies II Scores (Engineering Stud. Only)	.157

Odell (10, p. 16) reported a correlation coefficient between the Thorndike Test Score and first semester mathematics grades as .52. He also gave a correlation between high school algebra grades and college algebra grades of .41.

Working on the basis of individual classes rather than students, Krathwohl (6, pp. 234-242) correlated average grades in various mathematics courses with average test scores on the Iowa Mathematics Aptitude Test. He reported the following correlations with mathematics aptitude scores:

College Algebra Grades - - - - - -	0.61
Review Algebra Grades - - - - -	-0.10
Analytical Geometry Grades - - - -	0.52
Calculus I Grades - - - - - - - -	0.59
Calculus II Grades - - - - - - -	0.45

He maintains that students have difficulty in mathematics not because of lack of training but because of lack of ability, and that lack of mathematical ability is certainly a reason, if not the main one, why students drop out of engineering schools.

Cooprider (3, pp. 21-24) lists a group of nine correlation studies, by others, between college mathematics grades and various aptitude tests, having a median correlation of 0.48. The results of his own study carried out at Oregon State College were given in the following table form (3, p. 53).

Correlations with Mathematics Grade-Point Averages

Test	r	P. E.	N
Stanford Scientific Aptitude Test	.33	.05	160
Engineering and Physical Science Aptitude Test	.35	.03	293
A. C. E. Psychological Examination Total Score	.41	.03	325
A. C. E. Psychological Examination Q-Score	.43	.03	325
A. C. E. Psychological Examination L-Score	.31	.03	325
Ohio State University Psychological Test	.26	.04	321

The American Council on Education Psychological Examination

The American Council on Education was established in 1918 and is composed of representatives of national educational associations and of approved universities, colleges, and technological schools. A group of prominent educators representing national educational associations met in Washington D. C. in January 1918 to discuss the possible contribution of schools and colleges to national defense. From this grew the American Council on Education. Under the direction of the Committee on Measurement and Guidance, the Council supervises the Psychological Examination for College Freshmen, the first issue of which was in

1924. This examination represents one of the most comprehensive experiments in testing in the world.

The test has been largely the work of Thurston. The examination has been widely used and already by 1936 had been given to more than two million students.

In 1927 Thurston (13, p. 165) gave the following correlation coefficients between Test Scores and Freshmen Scholarship based on twenty six colleges using the 1925 form.

Completion Test - - - - -	.372
Arithmetic - - - - - - -	.341
Artificial Language - - -	.338
Analogies I - - - - - -	.284
Analogies II - - - - -	.291
Number Completion - - - -	.274
Absurdities - - - - - -	.295
Opposites - - - - - - -	.385

Thurston and Thurston (15, p. 170) stated that Dean W. S. Allen of Baylor University in a report to the President gave a correlation of $.331 \pm .051$ between marks made by students in mathematics during the fall quarter 1929 and scores on the American Council on Education Psychological Examination 1929. They (15, p. 172) also report that Odom had indicated that a correlation of 0.35 for 155 cases was found between the American Council on Education scores and grades in mathematics.

Seder (12, p. 99) working in the Secondary field gave the following results:

Correlation between Cooperative Achievement Test in Algebra and A.C.E. Psychological Examination Q Score

	Grade 10	Grade 11	Grade 12
Algebra 1st. year	.468	.187	
Algebra 2nd. year		.508	.446

In a Master's Thesis, Jones (5, p. 13-16) tells of five investigations by others in which the correlations between A. C. E. Test Scores and General Scholastic success were 0.45, 0.50, 0.45, 0.41, and 0.54.

Votaw (19, p. 216) found a correlation of 0.53 between Grade Points and the A. C. E. Score.

Thurston and Thurston (14, p. 2-3) wrote in the Manual of Instructions for the American Council on Education Psychological Examination for College Freshmen 1946 Edition as follows:

> The purpose of the Psychological Examination of the American Council on Education is to appraise what has been called scholastic aptitude or general intelligence, with special reference to the requirements of most college curricula. A large number of different tests have been used for this purpose. It has been found that, in general, linguistic tests give higher correlations with scholarship in the liberal arts colleges than do quantitative

tests. This higher correlation is probably, in part, due to the fact that most of the freshman courses in the liberal arts colleges depend more upon linguistic abilities than upon the abilities involved in quantitative thinking. For the scientific and technical curricula the quantitative tests may be more significant. These considerations indicate the desirability of isolating the several mental abilities and of appraising them separately. A complete program of this kind would be the ideal basis for educational and vocational counseling, and it is the objective toward which psychological research should be directed. A program of tests for seven primary mental abilities was made available by the American Council on Education for the first time in 1938 in the form of an experimental edition. This program required about four hours of student time, and it was necessarily more expensive than the one-hour test in current use. In the 1938 edition of the general Psychological Examination we introduced a practical compromise whereby two subscores were obtained from the one-hour examination, namely, a subscore for three linguistic tests, and another subscore for three quantitative tests. These two subscores do not represent primary mental abilities significant for college curricula that are dominantly linguistic or technical.

. . . . . .

The examination consists of the six tests that have been used for several years. The order of the tests has been arranged to alternate linguistic and quantitative tests because of the fatigue element. All of the tests have been included in several test experiments with factorial analyses to determine the primary mental abilities. These studies have justified the grouping of the six tests in two general classes, as follows:

Quantitative Tests: (the Q-score)
- Arithmetical Reasoning
- Number Series
- Figure Analogies

Linguistic Tests: (the L-score)
- Same-Opposite
- Completion
- Verbal Analogies

It is not recommended that the six separate test scores be used for any counseling, but there seems to be justification for using the two principal subscores as well as the total or gross score in this manner.

The test forms should be found useful in handling those problems in which it is advisable to distinguish a student's mental abilities from his high-school preparation and his industry. Faculty action in the case of a student who is failing can be intelligently guided if one has some means of knowing to what extent the student has applied himself to his college work, to what extent his high-school training meets the requirements of his college course, and what his mental abilities are. Very different faculty action can be taken, depending on which of these three factors may be held primarily responsible for a student's failure. It is to be hoped that these psychological test forms may lead to the early discovery of bright students. In those colleges where sectioning of classes in accordance with ability or preparation is customary, these test forms may serve as part of the evidence upon which the sectioning is based.

Perhaps a word should be said about different interpretations of test scores. Those who have used psychological tests and who have become convinced of their merits sometimes overestimate the significance of the test scores. While the scores do show roughly the mental

> alertness of the student, they should not be thought of as measuring mentality with high accuracy. The scores are roughly indicative of the level of mental alertness of the student, but they should not be taken so seriously as to exclude other evidences of intelligence and talent in individual cases. On the other hand, it is undeniably true that the psychological test scores tell us much more about the mental alertness of students than could be ascertained in a personal interview. In those situations where the scholarship standards of the high schools vary considerably, one is safer in judging the abilities of freshmen by psychological tests than by high-school records. Generally, the best usefulness of the tests is in combination with other evidences of ability such as grades in high school and in content examinations that are given uniformly to all students.

The Pre-Engineering Inventory

The Pre-Engineering Inventory was a joint project of The Engineers' Council for Professional Development, The Society for the Promotion of Engineering Education (now the American Society for Engineering Education), and the Carnegie Foundation for the Advancement of Teaching. It is now (since reorganizational changes) administered by The Measurement and Guidance Project in Engineering Education of the Educational Testing Service. It is, as previously stated, a test which purports to measure the informational aspect of the student who plans to pursue the study of Engineering.

Quoting from Memorandum No. 1, to Deans and Examiners, Date-July 15, 1946, issued by the Measurement and Guidance Project in Engineering Education (7, p. 1):

> After three years of experimentation with and refinement of the Pre-Engineering Inventory, the validity of this series of tests has been demonstrated sufficiently to insure its great value in programs of freshman selection and guidance in all types of engineering colleges. During this initial period, the Inventory was administered to more than 36,000 students and extensive researches have been conducted in thirty-five colleges of engineering which were carefully selected to provide a representative cross-section of all accredited colleges of engineering. The results in these institutions have demonstrated clearly that the Inventory should be made available to all colleges of engineering for use in their guidance programs.
>
> . . . . . .
>
> The services of the Measurement and Guidance Project in Engineering Education are now available to all colleges of engineering accredited by the Engineers' Council for Professional Development.

Excerpts from Memorandum No. 2, to Deans and Examiners, dated November 15, 1946, (9, p. 1-2) follow:

> Method of Scoring the Test
>
> In the Project Office, all of the tests in the Pre-Engineering Inventory are scored by I.B.M. electric test scoring machines. Each paper is scored twice

and, when any discrepancy between the first and second scoring arises, the test paper is hand scored to determine the correct score to be assigned.

In the scoring process, all tests in the Pre-Engineering Inventory are corrected for guessing, i. e., a certain fraction of the wrong answers is subtracted from the number of right answers. In the case of five-response items, this subtraction is one-fourth (1/4) of the wrong answers and, in the case of four-response items the subtraction is one-third (1/3) of the wrong answers. Although there seems to be little or no indiscriminate guessing on the part of pre-engineering students, there is good reason to employ the correction for guessing technique in the scoring of the Pre-Engineering Inventory tests. The few individuals who do guess wildly are thus identified and penalized.

The Scores Provided

The Pre-Engineering Inventory Revised Form A derives eight (8) scores: one (1) score for each of the seven (7) tests, and a Composite Score. All of these scores are raw scores, i.e., each score is the number of questions answered correctly minus the correction for wrong answers. The Composite Score is the sum of the raw scores for Tests II, III, and IV.[1]

After considerable preliminary research, it was discovered that the combined scores of three tests - - Test II - Technical Verbal Ability, Test III - Comprehension of

---

1 Compare with "The Composite Score" on page 27.

> Scientific Materials, and Test IV - General Mathematical Ability - - resulted in a prediction almost as effective as that obtained from the multiple regression equation. With this finding in hand, these three tests were constructed each with approximately the same number of items but with different time limits so that the obtained means and standard deviations of raw scores were almost exactly the same for each test when administered to a large representative group of freshman engineering students.
>
> . . . . . .
>
> On the I.B.M. Individual Report Cards, the only scores reported by the Project Office are the raw scores for each test and the Composite Score. Space is provided on the card for the Examiner or the Dean's Office to record the particular national percentiles and local percentiles which are most meaningful to the institution's purposes; e.g., College A may wish to use the public college national norms, while College B prefers to use the national norms for private colleges. Still other colleges may prefer the combined norms for public and private colleges. Since these needs cannot be anticipated by the Project Office, only the raw scores are reported.

In order to obtain percentile ranks, raw scores are first converted into scaled scores by a conversion table, and scaled scores are then converted into percentile ranks by the use of one of the following percentile tables (1) All Freshmen Engineering Students, (2) Students Enrolled in Public Colleges, No College Study, (3) Students Enrolled in Private Colleges, No College Study, (4) All Freshmen

Engineering Students, No College Study, (5) All Freshmen Engineering Students, Previous College Study, (6) Students Enrolled in Public Colleges, Non-veterans, (7) Students Enrolled in Public Colleges, Veterans, (8) Students Enrolled in Private Colleges, Non-veterans, or (9) Students Enrolled in Private Colleges, Veterans.

Norms Bulletin No. 2, May 1947, for use with the P. E. I., 1946 Edition (8, p. 2) discusses the Scaled Score System and Percentile Tables as follows:

> The Scaled Score System
>
> The system of scaled scores adopted for use with the Pre-Engineering Inventory provides a convenient basis for interpretation of test performance. In this system, the scores are strictly comparable from test to test, i.e., a score of 60 on the General Verbal Ability Test, which contains 100 items, represents the same relative status with respect to the standardization group as a score of 60 on the Spatial Visualization Test, which includes 56 items. Arithmetic in character, scaled scores may be added or multiplied. This practice is not strictly valid in the case of percentile ranks or raw scores. Perhaps of even greater significance is the fact that the scaled score system provides units which are in themselves comparable and are not influenced by irregularities in the distribution of raw scores as are percentile ranks. In this scaled score system, the average score of the standardization group is set equal to 50 and the standard deviation of scores is set equal to 15. Scores are provided at each one-fifteenth (1/15th) of a standard deviation, e.g.,

50, 51, etc. Thus a score of 65 or 35 represents one standard deviation from the average. Scores of 20 and 80 represent two standard deviations below and above the average respectively.

The Standardization Sample

The standardization group on which the scaled score system is based included 7,187 freshman engineering students. All members of this group were accepted by colleges of engineering cooperating in the Project and no student had been previously enrolled in college. Of this group 5,144 students were accepted by public colleges and 2,043 were enrolled in private colleges. Within the public college group, 1,382 students were non-veterans and 3,729 were veterans. In the private colleges, 658 students were non-veterans and 1,376 were veterans.

The standardization group is a highly representative sample of engineering freshmen with no previous college training who were accepted by colleges of engineering in 1946.

The Composite Score

The Composite Score now provided in the scaled score system is strictly comparable with the Composite Score as calculated previously. The conversion tables provided in Norms Bulletin No. 3 present the scaled score equivalents of each raw score. The Composite Score, which was formerly the sum of the scores obtained on Tests II, III, and IV, is still obtained by this procedure, but the resultant raw score is scaled and presented as a standard score. Consequently, the Composite Score is no longer expressed as a sum of the scores on Tests II, III, and IV, although it is

derived by their addition.[2]

Interpretation

In interpreting the scaled score system used with the Pre-Engineering Inventory, it may be helpful to bear in mind the general distribution of scores for students with no previous college training. In general, such students received scores distributed as follows:

Above	80	2 per cent
	70 and above	10 per cent
Above	60	25 per cent
	50 -- Average Score --	
Below	40	25 per cent
	30 and below	10 per cent
Below	20	2 per cent

The Percentile Tables

Despite their values, scaled scores are not without limitations. The greatest difficulty in interpretation is usually encountered in translating scaled scores into percentile ranks. The tables provided in the following pages of this report present the per cent of students receiving scores lower than the designated scaled scores. For example, in the case of the entire group of freshman engineering students (N = 9,994), 44 per cent of the total number of students in this group received a score on Test I of less than 50--the average score of all

---

2 The scores now given on the I.B.M. cards are scaled scores; Norms Bulletin No. 2 superseded Memorandum No. 2. What is designated as the Composite P. E. I. Score in this report is the one defined above, i.e. it is the scaled score equivalent of the sum of the raw scores on Tests II, III, and IV.

students who had had no previous college training.[3]

Vaughn (18, p. 114-120), Director, The Measurement and Guidance Project in Engineering Education, proposed that a Pre-Engineering Test, to be given at the end of the sophomore year high school, be developed. In an article written for the Journal of Engineering Education he says:

> From the point of view of the engineering curriculum, the sophomore year in high school is an appropriate point to begin exploring the potentialities of individuals with respect to the possible study of engineering.
>
> . . . . . .
>
> While the elective system within the last two years of high school is relatively restricted as compared with college curricula, there is in most cases an opportunity for the student to elect further courses in science and mathematics during the last two years of high school.
>
> . . . . . .
>
> The purpose of the proposed engineering aptitude test is to provide measures which are predictive of the student's success in advanced mathematics and science courses in the last two years of high school study. It is assumed, and there is much evidence to support this assumption, that abilities

---

3 Appendix I, page 83.

developed in these subjects are most predictive of initial success in the first stages of the engineering curriculum. This fact has been borne out repeatedly in researches of the effectiveness of the Pre-Engineering Inventory.

Secondly, the performance on the engineering aptitude test must be related to success on the Pre-Engineering Inventory. Specifically, the purpose of the engineering aptitude test is two-fold in character: (1) performance on the test must predict, with good precision, the student's later success in scientific subjects in the high school; and (2) individual performance on the test must also be highly related to a performance on certain sections of the Pre-Engineering Inventory.

. . . . . ; . . . .

It is certain that the engineering aptitude test must include a section dealing with mathematics. Our first problem is to measure as dependably as possible the student's general ability in this area.

. . . . . .

The three abilities which will definitely be sampled by the test are: (1) ability to comprehend and interpret scientific materials, including scientific vocabulary; (2) general mathematical ability; and (3) ability to comprehend mechanical principles. The optional test is spatial visualizing ability. In the Pre-Engineering Inventory, the tests which measure the ability to comprehend the meaning of important words in the vocabulary of high school science and mathematics and the ability to comprehend and interpret reading materials in scientific fields have consistently proved that their results provide a good prediction of a student's success in engineering subjects.

. . . . . .

The test of general mathematical ability will measure the student's ability to solve problems ranging in difficulty from arithmetic through plane geometry, to interpret graphs and tables, and to apply mathematical thinking to the solution of new problem types which the student has not previously encountered in his formal study, but for which he presumably will have developed the ability to proceed in an orderly manner. This test will measure both achievement and aptitude, and may properly be called a general ability test.

. . . . . .

The decisions concerning the content and emphasis of the engineering aptitude test are based largely on experimentation with engineering freshmen. These experiments were conducted in connection with the development and validation of the Pre-Engineering Inventory. It will be of interest to present here data concerning the predictive value of the seven tests included in the Pre-Engineering Inventory.

Lacking uniform criteria among the colleges of engineering participating in the Engineering Project, our initial researches on the predictive value of the Pre-Engineering Inventory must necessarily be based upon the relationship of test scores to grade-point averages as reported by the college of engineering. As is generally recognized, this procedure is not entirely adequate.

. . . . . .

When these factors are taken into account, a battery of tests which produces such a correlation of .50 to .60 is considered adequate for purposes of guidance and, in many cases, selection. Indeed, most colleges and universities utilize test batteries whose validity coefficients (as based on the grade-point criterion) are very much less than .50. When test batteries whose composite scores correlate with grade-point average during the first semester or first year of study of the

> order of .60 or higher, the battery is usually classed as a superior instrument. Individual tests which correlate with this criterion or course grades in excess of .50 can be considered as most useful for purposes of predicting success in the subject or area.

The colleges of engineering in the following ten schools (18, p. 119) gave the Pre-Engineering Inventory:

The California Institute of Technology

The University of California at Los Angeles

The Carnegie Institute of Technology

Georgia Institute of Technology

Massachusetts Institute of Technology

The University of Michigan

Newark College of Engineering

North Carolina State College of Agriculture and Engineering

Oklahoma Agricultural and Mechanical College

The University of Texas

All seven tests were given in all schools except one where Test VII was omitted. The following results were obtained:

Correlation Coefficients between Test Scores and First Semester Grade-point Averages

		Range	Median
Test I	General Verbal Ability	.16-.50	.38
Test II	Technical Verbal Ability	.25-.55	.47
Test III	Comprehension of Scientific Materials	.41-.65	.55
Test IV	General Mathematical Ability	.51-.71	.62
Test V	Comprehension of Mechanical Principles	.30-.55	.39
Test VI	Spatial Visualizing Ability	.22-.42	.35
Test VII	Understanding of Modern Society	.25-.53	.41
Composite Score		.44-.68	.62

The author has called some of the preliminary results of his survey to the attention of the Educational Testing Service. The following quote is from a return letter dated May 6, 1949:

> Although we have obtained correlations between Pre-Engineering Inventory scores and grade average for quite a number of schools, we have not gotten correlations with individual course grades except in two schools. In one school the correlation of mathematics score with first-term grades in analytic geometry and calculus were respectively 0.42 and 0.39 for about 300 students. In the other school the correlation with first-term grades in calculus was 0.55 for about 400 students.

We will appreciate very much a copy of your survey as soon as it is completed, and would like to refer to it when we publish data on the correlations of Pre-Engineering Inventory scores with individual course grades. We will certainly inform you when this is done.

Thank you again for the information you have given us, and let us know of the further results that you find.

Very sincerely,

Frederic M. Lord
Test Program Statistics

CHAPTER III

THE STUDY

The primary purpose of the study was the determination of certain relationships between two tests and grades in either College Algebra or Remedial Algebra at the Sault Ste. Marie Residence Center. This relationship was measured through correlations using the product-moment method in the calculation of the coefficients.

A secondary project was to determine the gain in student mathematical background after two terms of work at the Center. This was measured by re-giving Test IV - General Mathematical Ability - of the Pre-Engineering Inventory.

Tests Used in the Study

One of the tests used in this study was the American Council on Education Psychological Examination for College Freshmen, 1946 form. This test attempts to estimate a student's scholastic aptitude or general intelligence, taking into consideration the requirements of most college curricula. It is composed of six sub-tests, namely (1) Arithmetic, (2) Completion,

(3) Figure Analogies, (4) Same-Opposite, (5) Number Series, and (6) Verbal Analogies. These six sub-tests are divided into two groups:

I. Quantitative Tests: (the Q-score)
   Arithmetical Reasoning
   Number Series
   Figure Analogies

II. Linguistic Tests: (the L-score)
   Same-Opposite
   Completion
   Verbal Analogies

It is recommended by the American Council on Engineering that only the two principal sub-scores as well as the total score be used for any counseling. The test scores used in this study were the Q-score and the total or T-score. These scores were taken from a list given to the writer by Prof. Youngs of the Counseling Service at the Residence Center.

The other test was the Pre-Engineering Inventory, Revised Form A, 1946, of the Educational Testing Service. This test attempts to evaluate the student's background relative to his pursuing the study of engineering. It is composed of seven sub-tests, namely (1) General Verbal Ability, (2) Technical Verbal Ability, (3) Ability to Comprehend Scientific Materials, (4) General Mathematical Ability, (5) Ability to Comprehend Mechanical Principles, (6) Spatial Visualizing Ability, and (7) Understanding of Modern

Society. The test scores used in this study were the Scaled Score for Test IV - General Mathematical Ability - and the total or Composite Scaled Score. These scores were obtained through the cooperation of Prof. Youngs of the Counseling Service who read them from his file of I. B. M. cards.

It should be pointed out again that, although neither of these tests claims to predict success in College Algebra, they were used for that purpose in this study.

In Table I, page 38, it is noticeable that the average American Council on Education Test performance was lower in 1946-47 than in either of the other two years.

Table II, page 39, shows that the average Pre-Engineering Inventory Test scores were also lower in 1946-47 than in the other two years.

A possible explanation of lower test performance during the first year was the fact that testing conditions were not all they should have been. The confusion of building alterations was still in progress, thereby tending to lower test performance.

All four test measures showed a greater coefficient of variation for the initial year (1946-47). The

Table I

A. C. E. Test, College Algebra Group

Means, Standard Deviations and Coefficients of Variation for Test Scores

Year	1946-47	1947-48	1948-49	3 years
$\bar{Q}$	44.8	49.2	49.8	47.2
$S. D._q$	9.2	9.9	9.0	9.6
V_q	21	20	18	20
$\bar{T}$	112.5	113.9	117.8	114.4
$S. D._t$	20.2	23.1	18.8	20.3
V_t	18	20	16	18
N	221	96	131	448

Q - A. C. E. (Q-score)

T - A. C. E. (T-score)

1946-47 data tends to control the entire three year study period data because it accounts for 221 cases out of the 448 in the College Algebra study. This decrease does not reflect a diminishing enrollment at the Sault Ste. Marie Residence Center. In the initial year the student body of the Residence Center

Table II

P. E. I. Test, College Algebra Group

Means, Standard Deviations and Coefficients of Variation for Test Scores

Year	1946-47	1947-48	1948-49	3 years
$\bar{Y}$	34.0	52.5	48.0	42.1
$S.D._y$	12.8	8.1	10.5	13.7
V_y	38	15	22	33
$\bar{Z}$	40.1	50.6	49.8	45.2
$S.D._z$	12.9	10.2	9.2	12.4
V_z	32	20	18	27
N	221	96	131	448

Y - P. E. I. Scaled Score on Test IV - General Mathematical Ability

Z - P. E. I. Composite Scaled Score

was composed entirely of freshmen. In fact a second freshmen group was started early in January 1947. This group finished late in the summer, just in time to start their sophomore year in the fall. Since the fall of 1947 the Residence Center has offered two years of work.

Table III

A. C. E. Test, Remedial Algebra Group

Means, Standard Deviations and Coefficients of Variation for Test Scores

Year	1946-47	1947-48	1948-49	2 years
$\bar{Q}$		43.7	44.2	43.9
$S.D._q$		8.6	7.8	8.4
V_q		20	18	19
$\bar{T}$		103.2	108.7	105.6
$S.D._t$		20.8	20.5	20.6
V_t		20	19	20
N	0	36	27	63

Q - A. C. E. (Q-score)

T - A. C. E. (T-score)

In Tables III and IV it is shown that the two yearly groups in the Remedial Algebra study were equal in test performance on both the American Council on Education Psychological Examination and the Pre-Engineering Inventory. The reader will also notice

Table IV

P. E. I. Test, Remedial Algebra Group

Means, Standard Deviations and Coefficients of Variation for Test Scores

Year	1946-47	1947-48	1948-49	2 years
$\bar{Y}$		38.3	33.9	36.4
$S.D._y$		10.9	8.1	10.1
V_y		28	24	28
$\bar{Z}$		38.9	39.5	39.1
$S.D._z$		11.5	10.1	10.4
V_z		30	26	27
N	0	36	27	63

Y - P. E. I. Scaled Score on Test IV - General Mathematical Ability

Z - P. E. I. Composite Scaled Score

that the coefficients of variation for the two years are roughly equivalent.

Grades Used in the Study

The grades used in the study were all from College Algebra or Remedial Algebra courses taken at the Sault Ste. Marie Residence Center from the time it was founded in October 1946 to the end of the present academic year, June 1949. They were obtained from the record book of the Mathematics Department at the Residence Center.

Grades given to the Registrar by instructors are numerical. They are converted to letter grades by the Registrar, before distribution to the students, on the following conversion scale: A - 100-93, B - 92-86, C - 85-78, D - 77-70, and F - 70-0. In figuring Grade-point Averages in the Registrar's office, students who drop a course after the time limit of four weeks from starting are given a numerical grade of 65. In such cases instructors report only "F, dropped after end of fourth week of term."

It is noticeable, especially in Table V, that the frequency distribution for grades had a multi-modal tendency. These "modes" occur at 93, 86, 78, 70, and 65. They are caused by the tendency of instructors to raise a student's numerical grade slightly in order to put him in the next letter division. The

Table V

Frequency Distribution of Grades in College Algebra

Grade	Freq. 1946-1947	Freq. 1947-1948	Freq. 1948-1949	3 years	Grade	Freq. 1946-1947	Freq. 1947-1948	Freq. 1948-1949	3 years
97			4	4	57	5			5
96	4	2	2	8	56				
95	4	2		6	55	4		1	5
94	6	2	1	9	54		1		1
93	14	2	7	23	53		1		1
92		2	1	3	52				
91	3		4	7	51				
90	6	5	4	15	50	6	1	3	10
89	6	4	3	13	49				
88	3	3	7	13	48	3			3
87	5	5	3	13	47				
86	11	4	5	20	46	2			2
85	3	2	2	7	45	3			3
84	6	2	5	13	44				
83	5	3	6	14	43				
82	5	5	6	16	42	1			1
81	11	4	6	21	41	1			1
80	16	2	7	25	40				
79	9	6	2	17	39	2			2
78	16	6	9	31	38				
77		1	1	2	37	1			1
76	6	1	2	9	36				
75	3	3	7	13	35				
74	3	1	3	7	34				
73	1	3	4	8	33				
72	4	4	3	11	32				
71	3	2	4	9	31				
70	13	11	7	31	30				
69	1			1	29				
68					28				
67			4	4	27	1			1
66			2	2	26				
65	15	4	2	21	25				
64			1	1	24	2			2
63			2	2	23				
62	1	1	1	3	22				
61	3			3	21				
60	2	1		3	20				
59					19				
58	1			1	18	1			1

Table VI

Frequency Distribution of Grades
in Remedial Algebra

Grade	Freq. 1947-1948	Freq. 1948-1949	2 years	Grade	Freq. 1947-1948	Freq. 1948-1949	2 years
95	1		1	70	5	2	7
94	1	1	2	69			
93	2	2	4	68			
92	1		1	67			
91		1	1	66			
90		1	1	65	3	1	4
89				64		1	1
88				63			
87	1		1	62			
86	3	1	4	61			
85	1	1	2	60			
84	3	1	4	59			
83	2	3	5	58			
82	3	2	5	57			
81	1	1	2	56		1	1
80	2	1	3	55		1	1
79	1	1	2	54			
78		1	1	53	1		1
77		1	1	52			
76		1	1	51	1		1
75				50			
74	1	1	2	49	1		1
73	1		1	48			
72				47			
71	1		1	46		1	1

accumulation of 65's is caused by (a) the fact that those numerical grades between 66 and 69 inclusive that the instructors did not see fit to make into a 70, were moved back to a 65, (b) the fact that there is probably a tendency by the instructors to increase

Table VII

Means, Standard Deviations and Coefficients of Variation for Grades

Year	1946-47	1947-48	1948-49	Combined
$\bar{X}$	75.4	79.4	79.7	77.5
S. D.$_x$	16.0	9.8	10.4	13.2
V_x	21	12	13	17
N_x	221	96	131	448
$\bar{A}$		77.7	77.8	77.7
S. D.$_a$		11.6	11.8	11.9
V_a		15	15	15
N_a	0	36	27	63

X - Grade in College Algebra

A - Grade in Remedial Algebra

scores below 65 up to that value, and (c) the fact that 65 is the Registrar's F.

This "saw tooth" tendency in the grade distribution should be kept in mind in viewing the results obtained in the study.

The median grades in College Algebra were 79 in

1946-47, 80 in 1947-48, 81 in 1948-49, and 80 for the three year period.

The median grades in Remedial Algebra were 82 in 1947-48, 81 in 1948-49, and 81 for the two year period.

It is noticeable that the greater variability of test scores for the initial year, as shown in Tables I and II, is reflected in a greater variability of grades as shown in Table VII. This points out that in the instructors' opinions, at least, they were a more variable group as well as a lower group academically in Algebra.

Subjects Employed in the Study

This study involved 511 Sault Ste. Marie Residence Center students who took either College Algebra or Remedial Algebra in one of the three academic years, 1946-47, 1947-48, or 1948-49. These are divided into two groups, 448 in the College Algebra group and 63 in the Remedial Algebra group. The two groups were thus distinct and no individual was included in both the College Algebra and Remedial Algebra groups. The reason for this was that since the Pre-Engineering Inventory claims to be an inventory test, not an aptitude test, if the subjects who had previously

taken Remedial Algebra were included in the College Algebra group, not only would there occur the possibility but (we hope) the probability that the grades received in College Algebra would be affected. The students were sectioned, as stated previously, on the basis of the entrance requirement of a unit and a half of high school algebra.

The 118 students who took a re-test, at the end of two terms of study, on Test IV - General Mathematical Ability - in the Pre-Engineering Inventory were from the 1948-49 academic year and included students from both groups mentioned above.

Procedure

The main project seemed to be naturally divisible into three parts. First there was the correlation problem between grades and the American Council on Education Psychological Examination. Second, there was the correlation problem between grades and the Pre-Engineering Inventory. Third, there was the multiple correlation problem with the American Council on Education Psychological Examination and the Pre-Engineering Inventory as the dependent variables. At the start of the investigation this third problem was tentative

upon the findings in the previous two. The coefficients of correlation were determined by the product-moment method. The probable variability of the obtained coefficients were computed in terms of the probable error. The correlations were computed on a yearly basis as well as on the total survey period with the College Algebra group, but this was not the case with the Remedial Algebra group because of the questionable value of any results which might be obtained, due to the relatively small size of the groups. Regression lines for the College Algebra grades with each of the four test scores were found, and their Standard Errors of Estimate computed. The Regression Plane for the College Algebra grades with the two "mathematical" test scores was determined as well as the Standard Error of Estimate for it. For the Remedial Algebra group regression lines were developed for only the two P. E. I. test measures, since the correlation coefficients were not significant for the two A. C. E. measures.

In addition to this correlation work, tables were constructed showing the increase in percentage of students passing either College Algebra or Remedial Algebra, corresponding to an increase in the P. E. I. - Test IV - Scaled Score or the A. C. E. Q-score.

In the secondary project of analyzing the re-test results of the P. E. I. - Test IV, two types of groupings of students were used. In one the students were grouped according to the courses they had either successfully or unsuccessfully attempted. In the other the students were grouped according to their relative standing in the group on the test and on the re-test.

Results of the Study

The correlations between College Algebra grades and A. C. E. scores given in Table VIII are very similar to those reported by Cooprider at Oregon State College (3, p. 53) between A. C. E. scores and Mathematics Grade Point averages. He reported correlations of .43 for the A. C. E. (Q-score) and .41 for the A. C. E. total score.

The correlations between College Algebra grades and P. E. I. scores, given in Table IX, agree fairly well with those reported by Vaughn (18, p. 119) in 1947. His report, based on ten colleges, states that between P. E. I. Test IV and first semester grade point averages there was a median correlation of .62 with a range of correlations from .51 to .71, and that between P. E. I. Composite Score and first semester

Table VIII

Correlations between College Algebra Grades and A. C. E. Scores

Variables	r	P. E.	N	Year
XQ	.401	.04	221	1946-47
XQ	.379	.06	96	1947-48
XQ	.337	.05	131	1948-49
XQ	.396	.03	448	3 years
XT	.410	.04	221	1946-47
XT	.435	.06	96	1947-48
XT	.229	.06	131	1948-49
XT	.376	.03	448	3 years

X - Grade in College Algebra

Q - A. C. E. (Q-score)

T - A. C. E. Total Score

grade point averages there was a median correlation of .62 with a range of correlations from .44 to .68.

It appears from Table X that the A. C. E. test is

Table IX

Correlations between College Algebra Grades and P. E. I. Scores

Variables	r	P. E.	N	Year
XY	.479	.03	221	1946-47
XY	.724	.03	96	1947-48
XY	.655	.03	131	1948-49
XY	.528	.02	448	3 years
XZ	.489	.03	221	1946-47
XZ	.617	.04	96	1947-48
XZ	.655	.03	131	1948-49
XZ	.550	.02	448	3 years

X - Grade in College Algebra

Y - P. E. I. Scaled Score on Test IV

Z - P. E. I. Composite Scaled Score

of little, if any, use in predicting attainment in Remedial Algebra. It appears however that the P. E. I. Test IV Scaled Score might be used.

Table X

Correlations between Remedial Algebra Grades
and Test Scores

Two year Period 1947-48 and 1948-49
N = 63

Variables	r	P. E.
AY	.511	.07
AZ	.305	.08
AQ	.115	.08
AT	.136	.08

A - Grade in Remedial Algebra

Y - P. E. I. Scaled Score on Test IV

Z - P. E. I. Composite Scaled Score

Q - A. C. E. (Q-score)

T - A. C. E. Total Score

Regression Lines and Plane for College Algebra based on 448 cases for a three year period are as follows:

$X = 0.51Y + 56.03$

with Standard Error of Estimate of 11.2

$X = 0.59Z + 50.83$

with Standard Error of Estimate of 11.0

$X = 0.54Q + 52.01$

with Standard Error of Estimate of 12.1

$X = 0.24T + 50.04$

with Standard Error of Estimate of 12.2

$X = 0.43Y + 0.16Q + 51.66$

with Standard Error of Prediction of 11.1 and with the multiple correlation coefficient of .544

X - Grade in College Algebra

Y - P. E. I. Scaled Score on Test IV

Z - P. E. I. Composite Scaled Score

Q - A. C. E. (Q-score)

T - A. C. E. Total Score

Regression Lines for Remedial Algebra Grade based on 63 cases for a two year period are as follows:

$A = 0.60Y + 55.86$

with Standard Error of Estimate of 10.2

$A = 0.35Z + 64.02$

with Standard Error of Estimate of 11.3

A - Grade in Remedial Algebra

Y - P. E. I. Scaled Score on Test IV

Z - P. E. I. Composite Scaled Score

It appears that the A. C. E. is not of much use in predicting grades in Remedial Algebra, the correlation with the Q-score being .115 and with the Total score .136. Therefore, no Regression Plane was determined. The multiple correlation coefficient for Remedial Algebra on the P. E. I. Test IV Scaled Score and A. C. E. (Q-score) is .528 which is only .017 higher than the simple correlation between Remedial Algebra and P. E. I. Test IV Scaled Score.

Table XI shows definitely that the chance of passing College Algebra increases with the A. C. E. (Q-score). This percentage increases with the exception of one case in the 0-9 interval.

Table XI

Relationship of Chance of Passing College Algebra to A. C. E. (Q-score)

A. C. E. (Q-score)	Number of Cases Pass	Fail	% Pass
70-80	3	0	100
60-69	39	2	95
50-59	124	17	88
40-49	149	30	83
30-39	45	23	66
20-29	6	6	50
10-19	1	2	33
0-9	1	0	
All groups	368	80	82

While not as pronounced as in Table XI, it is evident in Table XII that there is a general tendency for chance of passing Remedial Algebra to increase with the A. C. E. (Q-score).

Table XIII illustrates effectively that increased chance of passing College Algebra goes with increased P. E. I. Test IV Scaled Score. However, one student

Table XII

Relationship of Chance of Passing Remedial Algebra to A. C. E. (Q-score)

A. C. E. (Q-score)	Number of Cases Pass	Fail	% Pass
70-80	0	0	--
60-69	2	0	100
50-59	10	1	91
40-49	28	8	78
30-39	7	1	88
20-29	5	1	82
10-19	0	0	--
0-9	0	0	--
All groups	52	11	83

out of two passed in the 0-9 range. In case the question comes up, the student passed in the 0-9 range on the A. C. E. (Q-score) is not the same individual who passed in the 0-9 range on the P. E. I. Test IV Scaled Score. The student reported in Table XI had a Q-score of 9, a P. E. I. Test IV Scaled Score of 57 and received a low D in College Algebra. The passing student reported in Table XIII had a P. E. I.

Table XIII

Relationship of Chance of Passing College Algebra to P. E. I. Test IV Scaled Score

P. E. I. Test IV Scaled Score	Number of Cases Pass	Fail	% Pass
80-90	3	0	100
70-79	6	0	100
60-69	37	0	100
50-59	89	6	94
40-49	109	10	92
30-39	78	22	78
20-29	39	22	64
10-19	6	19	24
0-9	1	1	
All groups	368	80	82

Test IV Scaled Score of 6, a Q-score of 38 and received a "modal" 70 in College Algebra. The failing student reported in Table XIII was low on both test measures, making P. E. I. and A. C. E. scores of 9 and 14 respectively and receiving a correspondingly low numerical grade of 24 in College Algebra.

Table XIV

Relationship of Chance of Passing Remedial Algebra to P. E. I. Test IV Scaled Score

P. E. I. Test IV Scaled Score	Number of Cases Pass	Fail	% Pass
80-90	0	0	--
70-79	0	0	--
60-69	1	0	100
50-59	4	0	100
40-49	16	1	94
30-39	23	3	88
20-29	6	7	46
10-19	1	0	
0-9	1	0	
All groups	52	11	83

From Table XIV it can be seen that a definite "break" in chance of passing occurs at a P. E. I. Test IV Scaled Score of 30.

In Table XV the wide range of individual percentiles within each group points to the fact of much overlapping between the groups.

Table XV

Re-test Results, Students Grouped According to Courses Taken

Group	No. Cases	Test P. Range	Test Median	Re-test P. Range	Re-test Median	Change P. in P.	$\bar{Q}$
I	87	97-6	42	100-24	80	38	51
II	8	63-8	24	89-28	61	37	43
III	8	32-5	16	89-38	60	44	47
IV	15	63-1	10	86- 4	44	34	44
All groups	118	97-1	32	100- 4	71	39	50

P - national percentile on P. E. I. Test IV

$\bar{Q}$ - mean A. C. E. (Q-score)

Groups

I Students who have passed college algebra, advanced algebra, and trigonometry.

II Students who have passed college algebra and failed, either or both, advanced algebra and trigonometry.

III Students who have failed college algebra and took it again whether passing or failing in the second attempt.

IV Students who passed remedial algebra and then either passed or failed college algebra.

From Table XV (admittedly based on too few cases) it would appear that it is better to fail a student in College Algebra and have him take it twice than to pass him and have him fail in the next course. Group III end up for all practical purposes equal to Group II in Pre-Engineering preparation as measured by the re-test and were eight percentile points lower than Group II on the original test measure. This greater increase in group percentile may to a certain extent, however, be a reflection of the higher mean A. C. E. (Q-score) for the group, as seen in Table XV.

It should be kept in mind in viewing the best group, Group I, in regard to the gain in percentile, that in this group are found those students whose gain is for practical purposes limited by the 100 percentile ceiling. Perhaps a better comparative index of achievement would be the ratio of the change in group median percentile to the mathematically possible change in group median percentile. Under this index of achievement, Group I would rank first with a .67, with the indices for Groups II, III, and IV being .49, .52, and .38 respectively.

Table XVI

Re-test Results, Students Grouped According to Relative Standing on a National Basis

Group	Number of Cases	Mean A. C. E. (Q-Score)
plus plus	31	57
plus zero	2	54
plus minus	4	53
minus plus	22	51
minus zero	3	54
minus minus	56	44
All groups	118	50

Explanation of Grouping Above

The first sign indicates whether the student's national percentile on the P.E.I. Test IV is above (plus), equal to (zero) or below (minus) the mean (or median) national percentile of the national group, which of course is 50.

The second sign indicates whether the student's national percentile on the re-test of the P.E.I. Test IV is above, equal to, or below the median national

percentile of the local group, which was 71. (See Table XV)

Table XVII

Re-test Results, Students Grouped According to Relative Standing on a Local Basis

Group	Number of Cases	Mean A. C. E. (Q-score)
plus plus	43	55
plus zero	3	48
plus minus	7	50
zero plus	7	55
zero zero	1	56
zero minus	4	43
minus plus	3	50
minus zero	1	69
minus minus	49	44
All groups	118	50

Explanation of Grouping Above

The first sign indicates whether the student's national percentile on the P. E. I. Test IV is above (plus), equal to (zero), or below (minus) the median

national percentile of the local group, for the original test, which was 32. (Table XV).

The second sign indicates whether the student's national percentile on the re-test of the P. E. I. Test IV is above, equal to, or below the median national percentile of the local group, which was 71. (See Table XV).

Nothing especially significant is apparent from a random glance at Tables XVI and XVII. The various groups in both tables appear to show but a slight variation in the mean Q-score. In an attempt to draw some conclusions from these dats, however, the writer combined these small groups into the larger groups, the Gainers, the Holders, and the Losers.

In Table XVI, the Gainer Group composed of the minus plus and zero plus sub-groups has a mean Q of 51, the Holder Group composed of the plus plus and minus minus sub-groups has a mean Q of 49, and the Loser Group composed of the plus zero and plus minus sub-groups has a mean Q of 53. This information does not seem to lend any clarity to the picture. Interpretations made from the scores of the Holder Group are open to question since this group is composed of some individuals who could not be gainers and some individuals who could not be losers.

Table XVII presents a different break up of the subjects. Here standing is based entirely upon local norms. If, as in the above paragraph, these minor groups be combined into the Gainers, Holders, and Losers, the mean Q for the Gainers, Holders, and Losers, respectively, would be 55, 49, and 48.

In an effort to measure the reliability of the P. E. I. Test IV Scaled Score the writer determined the correlation between the test and re-test measures to be .780 with a probable error of .02. The means for the test and re-test were 46.9 and 63.9 respectively. The Standard Deviations were for all purposes equal, being 12.0 for the test and 12.1 for the re-test. The test showed a coefficient of variation of 25 compared to a coefficient of variation of 19 for the re-test.

An Additional Phase of the Study

In attempting to digest the data given in the previous pages, the writer was impressed with the difference between the data given for the initial year and the two following years. Noticeable differences occur between the P. E. I. scores and the grades.

The mean P. E. I. Test IV Scaled Score for the first year was 34.0 compared to a 52.5 and a 48.0

for the other two years. The Standard Deviation is larger the first year than for either of the other two years, being 12.8, 8.1, and 10.5 respectively. (See Table II). Table II shows that the Coefficient of Variation was noticeably higher for the 1946-47 year than for either of the other two years. The values were 38, 15, and 22 respectively. For the P. E. I. Composite Score we find a significant difference, as well between the initial year and the two following years. The initial year also has the greatest Standard Deviation of the three, and a Coefficient of Variation of 32 compared to 20 and 18 for the other two years.

For the grades it is also noticeable that the data for the year 1946-47 was different. The mean grade is not a great deal lower for the first year; however, this difference becomes more pronounced when we consider the fact that, of the eighty failures reported in Table XIII, fifty five were from the initial year. Considering the fact that the College Algebra group the first year was approximately equal to the group for the other two years combined, the percentage of failure the first year was 25%, compared to 9% and 12% for the two succeeding years. Could

this high percentage of failure the first year be explained by the possibility that everyone was put in College Algebra regardless of the regular requirement of a unit and a half of high school algebra? An inquiry by the writer to the Registrar's Office received a negative answer.

The correlation between the P. E. I. Test IV Scaled Score and Grades was noticeably lower in 1946-47, being .479, compared with .724 and .655 for the two following years. This same difference is noticeable in the correlations between P. E. I. Composite Scores and Grades which were .489, .617, and .655 respectively.

The correlations between the A. C. E. Measures and Grades for the first year were not out of line with the other yearly correlations, being the highest of all for the A. C. E. (Q-score). See Table VIII. The A. C. E. correlations were all lower than the P. E. I. correlations. The highest recorded A. C. E. correlation was .435 compared to the lowest P. E. I. correlation of .479. It appears that the College Algebra grades can be predicted from the P. E. I. scores alone almost as well as by a multiple correlation with the A. C. E. scores.

The multiple correlation for College Algebra grades based on the P. E. I. Test IV Scaled Score and the A. C. E. (Q-score) was .544 compared to a simple correlation coefficient of .528 based on the P. E. I. Test IV Scaled Score only.

The writer made the following conclusions at this time:

(1) The P. E. I. measures have a marked superiority over the A. C. E. measures in predicting success in College Algebra.

(2) A simple P. E. I. correlation is for practical purposes as good as a multiple A. C. E. - P. E. I. correlation in predicting success in College Algebra.

(3) The first year was not a normal year either for P. E. I. test results or grades.

With these considerations in mind, the writer determined to compile certain data for a two year College Algebra Group covering the latter two years.

The correlation coefficient between College Algebra Grades and the P. E. I. Test IV Scaled Score was .667 with a probable error of .02 for the two year group compared to a previously stated correlation of .528 with a probable error of .02 for the three year group. The correlation coefficient between College Algebra Grades and the P. E. I.

Table XVIII

Comparative Means, Standard Deviations and Coefficients of Variations of Grades and P. E. I. Scores for the Three Year and Two Year College Algebra Groups

	Three Year Group	Two Year Group
$\bar{X}$	77.5	79.6
$S.D._x$	13.2	9.9
V_x	17	13
$\bar{Y}$	42.1	49.9
$S.D._y$	13.7	9.8
V_y	33	20
$\bar{Z}$	45.2	50.1
$S.D._z$	12.4	9.8
V_z	27	20

X - Grade in College Algebra

Y - P. E. I. Scaled Score on Test IV

Z - P. E. I. Composite Scaled Score

Composite Scaled Score was .638 with a probable error of .03 for the two year group compared to a previously stated correlation of .550 with a probable error of .02 for the three year group.

Testing the two year College Algebra Group correlations against the null hypothesis, we find that they are both very significant. For 225 degrees of freedom a correlation need be only .132 to be significant at the 5% level and .173 to be significant at the 1% level.

It can be stated here also that an observation of the correlations for the Remedial Algebra Group for the two year period (Table X) shows that the P. E. I. Test IV Scaled Score - Remedial Algebra Grade correlation is significant at the 1% level, whereas the P. E. I. Composite Scaled Score - Remedial Algebra Grade correlation is significant at the 5% level but not at the 1% level.

Regression Lines for College Algebra Grade based on 227 cases for the two year period are as follows:

$X = 0.67Y + 46.02$

with a Standard Error of Estimate of 7.4

$X = 0.64Z + 47.34$

with a Standard Error of Estimate of 7.6

X - Grade in College Algebra

Y - P. E. I. Scaled Score on Test IV

Z - P. E. I. Composite Scaled Score

Table XIX

Relationship of Chance of Passing College Algebra to P. E. I. Test IV Scaled Score for Two Year Group
N = 227

P. E. I. Test IV Scaled Score	Number of Cases Pass	Fail	% Pass
80-90	3	0	100
70-79	6	0	100
60-69	28	0	100
50-59	71	6	92
40-49	70	9	89
30-39	23	10	70
20-29	1	0	
10-19	0	0	
0-9	0	0	
All groups	202	25	89

It can also be observed that the skewness of the College Algebra grade distributions for the combined three year period was -.57, compared to a value of -.12 for the combined latter two year period.

According to Richardson (11, p. 153-154):

> While it is dangerous to set limits on such indexes, we may say, as a rough measuring stick, numerical values of skewness computed by (2) less than 0.25 may be considered small, numerical values between 0.25 and 0.5 as moderate, and numerical values greater than 0.5 as large.

Formula (2) above was: $Sk = \frac{3(M - Md)}{S.\ D.}$

CHAPTER IV

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

In this study the grades received in College Algebra at the Sault Ste. Marie Residence Center, Michigan College of Mining and Technology, were correlated with each of four test measures, namely the P. E. I. Test IV Scaled Score, the P. E. I. Composite Scaled Score, the A. C. E. (Q-score), and the A. C. E. Composite Score, on a yearly basis, and a three year basis.

When the writer had completed the three year phase of the study, he was impressed by the noticeable difference in the data for the 1946-47 year and the two succeeding years. The author then decided to embark on an additional phase to the study. Hence, the correlations between the two P. E. I. Measures and the College Algebra Grades for the 1947-48 and 1948-49 years were determined.

The Remedial Algebra grades were correlated with each of the four test measures above on a two year basis only. Remedial Algebra was not offered the first year (1946-47), and the limited number of cases (63)

involved hardly justified the computation of correlation coefficients on a yearly basis.

The P. E. I. Test IV was re-given at the end of the second term of the 1948-49 school year. The results were used in two ways, (1) the gains made by various sub-groups were compared, and (2) the re-test scores were correlated with the original test scores to estimate the reliability of the test itself by a self-correlation method. A correlation of .780 was obtained for 118 cases.

Conclusions

(a) The first year, 1946-47, was not a normal year, either for test results or grades.

(b) In this study the Pre-Engineering Inventory measures had higher correlations than the American Council on Education Psychological Examination measures with both College Algebra and Remedial Algebra grades in every instance.

(c) The two Pre-Engineering Inventory measures (Test IV Scaled Score and Composite Scaled Score) correlated about equally with College Algebra grades. If one is only interested in predicting College Algebra grades, it therefore hardly seems necessary to give more

Table XX

Summary of Correlations between
Test Scores and Grades

	1946-47	1947-48	1948-49	3 year	Last 2 years
N	221	96	131	448	227
XY	.479	.724	.655	.528	.667
XZ	.489	.617	.655	.550	.638
XQ	.401	.379	.337	.396	
XT	.410	.435	.229	.376	
X(YQ)				.544	
AY					.511
AZ					.305
AQ					.115
AT					.136
A(YQ)					.528

X - Grade in College Algebra

A - Grade in Remedial Algebra

Y - P. E. I. Scaled Score on Test IV

Z - P. E. I. Composite Scaled Score

Q - A. C. E. (Q-score)

T - A. C. E. Composite Score

than Test IV. In the case of Remedial Algebra grade prediction, the Test IV Scaled Score is noticeably superior to the Composite Scaled Score.

(d) In this study it was found that a simple P. E. I. correlation was for practical purposes as good as a multiple P. E. I. - A. C. E. correlation in predicting grade in either College Algebra or Remedial Algebra.

(e) In this study the A. C. E. measures showed no significant correlation with Remedial Algebra grades.

(f) In this study the correlation of .780 found between the test and re-test scaled scores on the Pre-Engineering Inventory Test IV indicates a fairly good reliability for this test.

(g) It is difficult to tell whether it is better for a student to take College Algebra twice or Remedial Algebra, then College Algebra, from the data presented in Table XV. Group III, the "College Algebra twice" group, increased their median percentile more than Group IV, the "Remedial Algebra - College Algebra" group, 44-34. Group III also had a noticeably higher index of achievement, defined here to mean the ratio of the change in group median percentile to the mathematically possible change in group median percentile, .52 compared to .38. However, to draw any conclusions

therefrom is dangerous for two reasons, (1) the limited number of cases considered, and (2) the difference in background training itself between the two groups, Group III having had one and a half units of high school Algebra where Group IV did not.

(h) The modal frequency just over a letter grade division is inherent in the numerical grading system used, and in this study was not due to abnormal conditions of the first year, since it still is noticeably apparent in a frequency table covering the College Algebra grades for the last two year period. This "saw tooth" tendency does not appear to apply in regard to Remedial Algebra grades except at the pass - fail division.

(i) According to this study, in using the latter two years as the base period, about 44% of the variation in College Algebra grades can be explained by the relationship of College Algebra grades to the initial preparedness of the student as measured by the P. E. I. Test IV Scaled Score. Similarly, since the correlation between College Algebra grades and the A. C. E. (Q-score) for this same two year period was .356, about 13% of the variation in College Algebra grades can be explained by the relationship of College Algebra grades to scholastic aptitude as measured by the American

Council on Education Psychological Examination (Q-score). The interpretation of the coefficient of correlation squared as giving the percentage of variance in grades accounted for by the relationship between grades and test scores should not be taken, however, to imply cause and effect. A high correlation proves a close linear mathematical relationship between the variables, but only suggests the probability of a causal relationship between the two variables.

Recommendations

The writer recommends that:

(1) the Pre-Engineering Inventory Test IV continue to be given to incoming students at the Sault Ste. Marie Residence Center during orientation week, and that it be given early in the week.

(2) should the Test IV Scaled Score be <u>below</u> 30 for a student with a unit and a half of high school algebra, he be informed of this fact by his advisor, and that he and his advisor consider the possibility of his taking Remedial Algebra. If this basis for sectioning had been carried out in the past, it would have moved eighty eight students from the College Algebra group into the Remedial Algebra group on the three year period and one student from the College

Algebra group into the Remedial Algebra group during the last two years. The mean grade earned by the eighty eight students was 61.5 and the single student received a College Algebra grade of 70.

(3) should a student's Test IV Scaled Score be 50 or above, he be permitted to take College Algebra by a waiver of requirements. This, if carried out in the past, would have moved five students from the Remedial Algebra group into the College Algebra group. These five students earned a mean Remedial Algebra grade of 87.4.

BIBLIOGRAPHY

1. Bloom, Benjamin, and Allison, Jane M. Developing a College Placement Test Program. Journal of General Education. University of Chicago Press. Vol. III, No. 3. April 1949. Pp. 210-15.

2. Conant, James B. The Use of Tests. What the Colleges are Doing. No. 84, Spring 1949. P. 3.

3. Cooprider, Harold Arthur. The Predictive Value of the Stanford Scientific and the Engineering and Physical Science Aptitude Tests, unpublished master's thesis, Oregon State College, 1948. 83 p.

4. Johnson, J. B. Who Should Go to College. University of Minnesota Press, 1930. Pp. 19-20.

5. Jones, George Alfred Arnold. The Prediction of Scholastic Success at Oregon State Agricultural College, unpublished master's thesis, Oregon State College, 1933. 170 p.

6. Krathwohl, W. C. Predictions of Average Class Achievement by Means of Aptitude Tests. The Journal of Engineering Education. Vol. 37, November 1946. Pp. 234-42.

7. Measurement and Guidance Project in Engineering Education. General Policies for Administering the Pre-Engineering Inventory. Memorandum No. 1 to Deans and Examiners. July 15, 1946. 3 p.

8. Measurement and Guidance Project in Engineering Education. The Pre-Engineering Inventory, Summer and Fall 1946, Scaled Scores. Norms Bulletin No. 2, May 1947, 11p.

9. Measurement and Guidance Project in Engineering Education. Scoring and Reporting the Pre-Engineering Inventory Test Results. Memorandum No. 2 to Deans and Examiners. November 15, 1946. 4 p.

10. Odell, Charles W. Predicting the Scholastic Success of College Freshmen. Bureau of Educational Research, College of Education. Published by the University of Illinois. Bulletin No. 37, 1927. 54 pp.

11. Richardson, C. H. An Introduction to Statistical Analysis. Revised Edition. Harcourt, Brace and Company, New York. Pp. 153-54.

12. Seder, Margaret. The Reliability and Validity of the American Council Psychological Examinations. Journal of Educational Research, 1938 Edition. Vol. 34, October 1940. No. 2. Pp. 90-101.

13. Thurston, Dr. L. L. American Council on Education Psychological Examinations, Educational Record. Vol. 8, April 1927. No. 2. Pp. 156-182.

14. Thurston, L. L. and Thurston, Thelma Gwinn. Manual of Instructions, American Council on Education, Psychological Examination for College Freshmen. Copyright 1946 by the American Council on Education. 8 p.

15. Thurston, L. L. and Thurston, Thelma Gwinn. The 1930 Psychological Examinations. Educational Record. Volume 12, April 1931. No. 2. Pp. 160-78.

16. Tolley, William P. American Education and the Testing Movement. The Education Record. January 1948. Pp. 86-96.

17. Tyler, Ralph W. The Road to Better Appraisal. NEA Journal, May 1949. Pp. 336-37.

18. Vaughn, K. W. A Proposed Engineering Aptitude Test for High School Students. The Journal of Engineering Education. Vol. 38, October 1947. Pp. 114-21.

19. Votaw, David F. A Comparison of Test Scores of Entering College Freshmen as Instruments for Predicting Subsequent Scholarship. Journal of Educational Research. Nov. 1946. Pp. 215-18.

20. Wrenn, Gilbert C. Manual of Directions of the Study-Habits Inventory, Revised edition 1941. Stanford University Press. 6 p.

APPENDIX I

CONVERSION OF P. E. I. SCALED SCORES INTO PERCENTILES

from

Norms Bulletin No. 2

The Measurement and Guidance Project
in Engineering Education

Summer and Fall, 1946

Scaled Scores

All Freshman Engineering Students

Total Group

N = 9,994

Scaled Score	I Gen. Verbal	IV Gen. Math.	Composite Score
90		99	99
80		96	95
70		86	87
68		84	84
66		80	80
64		77	77
62		71	73
60		68	69
58		63	64
56		59	58
54		53	53
52		49	48
50	44	42	42
48		38	38
46		32	33
44		28	28
42		24	24
40		20	20
38		16	17
36		13	14
34		10	11
32		9	8
30		6	7
20		1	1
10		0	0