THE USE OF A "CAUTION FACTOR" TO INCREASE THE PREDICTIVE VALUE OF THE AMERICAN COUNCIL ON EDUCATION PSYCHOLOGICAL EXAMINATION

by

LAWRENCE WILBERT CARRILLO

A THESIS

submitted to

OREGON STATE COLLEGE

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

June 1949

APPROVED:

Redacted for Privacy

Associate Professor of Education

In Charge of Major

Redacted for Privacy

Dean of School of Education

Redacted for Privacy

Chairman of School Graduate Committee

Redacted for Privacy

Dean of Graduate School

ACKNOWLEDGEMENTS

To Doctor Robert R. Reichart, the writer owes a debt of gratitude. His constructive criticism of the form and content, willingness to give his time, and guidance and suggestions as the experimentation progressed, have aided immeasurably in the preparation of this thesis.

Acknowledgements is made to those individuals on the staff of Oregon State College whose cooperation in making available the necessary data made this study possible, especially Mr. William R. Crooks, Dean Franklin R. Zeran, and Mr. Dallas Norton.

The writer also owes much to his wife, Ruth, for her assistance, and patient encouragement.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

THE USE OF A "CAUTION FACTOR" TO INCREASE THE PREDICTIVE VALUE OF THE AMERICAN COUNCIL ON EDUCATION PSYCHOLOGICAL EXAMINATION

CHAPTER I

GENERAL INTRODUCTION

General Statement of the Problem. For the past two years it has been a portion of the duties of the writer to administer, score, and interpret the American Council on Education Psychological Examination for College Freshmen. This examination is used at approximately four hundred colleges and universities in the United States for the purpose of determining the individual's aptitude for academic endeavor at the college level. Much weight is attached to the results of this test by administrators and personnel workers. However, during practical guidance use, it becomes apparent that in some situations the test results are hardly valid enough to warrant their use for guidance purposes. This is not due to the fact that the test is a poor test; on the contrary, it is probably the best instrument available at the present time for determining college aptitude in technological institutions. The lack of correlation between test results and college grades is more probably due to the complex of factors which enters into the college academic situation.

Some of these factors are, at least at the present,

non-measurable. It would appear, however, that by a refinement of the scoring of the test, at least one factor can be accounted for which may increase the validity of the instrument. This factor may be called the "caution factor", which may be explained in this way. In this test, the recorded raw score is the number of correct responses. Since the test is definitely of the "speed test" variety, this would seem to place a penalty upon the person who proceeds slowly and cautiously, but who answers correctly the largest percentage of the questions attempted. Conversely, the person who does guess on those questions which he does not immediately know covers more territory and would seem to have a better chance to make a higher score. In spite of his higher score this latter type of individual might not be as good a potential student. This could be particularly true in such a curriculum as is found in the School of Engineering, which, because of the nature of the materials studied, demands a high degree of care and accuracy.

Since so few studies have been made concerning the "caution factor", it is necessary in the interest of clarity of presentation to define the term as used here, and to attempt to justify this concept in terms of its fitness and usability for a study of this type.

Definition of "Caution Factor". The term "caution factor", as used in this study, indicates the ratio of the number of correct responses to the number of responses attempted on either a portion of or the whole of the A.C.E. test. The ratio is obtained by dividing the number of correct responses by the total number of items attempted. This method of converting "caution" into a mathematical relationship is apparently highly reliable. One study on the Otis Mental Ability Test, using this method, reports split-half reliabilities corrected by the Spearman-Brown prophecy formula to be .859 for men, .901 for women, and .883 for both combined. (12, p.2).

The "Caution Factor" Concept. In ordinary conversation one speaks of the "cautious" and "rash" individual, designating thus persons of more or less opposite personality types. The person who is spoken of as "cautious" is normally careful and conservative in his actions, exercising discretion, and avoiding risk or danger as much as possible; while the "rash" person would be hasty in thought or action, reckless and impetuous, or at least quicker in decisions, disregarding considerations. There seems no reason to believe that these distinctions are not entirely valid, though differing from each other more in degree than in kind. These concepts are usually thought of

as belonging to the domain of character rather than to that of intelligence. However, according to the study cited above (12, p. 2), there is a high relationship between intelligence and "caution", as measured by the same type of formula as used here.

This formula, in all probability, does not completely evaluate the wrong answers. But, for the most part, the wrong responses are completely ignored in scoring intelligence tests. This is not only true of the A.C.E. test, but many other instruments of the same type. Therefore this would seem to present at least a partial method of incorporating this trait into a measuring instrument of the intelligence test type which is already in use.

Carryover into the Classroom. Another question may be evident at this point: In what circumstances and situations will "caution" assert itself; will "caution" evident on the test for a given individual also be present in the classroom situation for that same individual? It is impossible, of course, to give a complete answer to this query. However, since the classroom presents a "testing" situation which resembles, at least to a large extent, the original testing situation, there should logically be a high probability of the same reaction in both cases. That is, factors such as some lack of knowledge, necessity

for speed in response, the urgency of the situation, and a question to be answered by the person as an individual, are present in both. This would seem to present a sufficient stimulus for the activity of the "caution factor" in both cases.

It is possible and in fact probable that this same factor may be present in other situations. The problem at hand, however, is to determine the measurability and the effects upon academic performance, rather than the exact nature of the factor.

Possible Objections. Since wrong answers made on the test were taken as indicative of the presence or absence of the factor, two possible objections to the method used to determine "caution factor" can be seen at this time. When the reader considers that many of the answers which were scored right must have likewise involved the element of guessing, it is evident that we are measuring only part of an incorrect answer to a question shows that the individual either did not know the right answer or was unable to recall it at the time, but still gave an answer of some sort. And since the chances for a correct "guess" are as large as for one which is incorrect, if any relationship is found at all, the real relationship must actually be greater. This would not seem, in the opinion

of the writer, to detract from the study, but rather to add emphasis to what is found.

It may also be objected that, in many instances, the individual gives answers which he believes to be correct, but which prove to be wrong in the final analysis. This is probably often true, but if the "caution factor" is operating at its maximum, the person would probably refuse to write down any answer of which he was not completely sure, while the less cautious person would hazard a guess. This is another instance of the difference being in degree rather than kind, with the outcome depending on the amount of the factor present, the reactions probably being distributed in a normal curve.

The hypothesis which this study is designed to investigate, then, is: That, by incorporating a "caution factor" into the scoring of the American Council on Education Psychological Examination, the validity of this instrument can be increased for students of engineering.

CHAPTER II

CONDITIONS AND MATERIALS OF THE STUDY

Purpose and Need. Several billions of dollars are expended annually in the United States for the education of coming citizens. Added to these public funds are the contributions of the individual parent in order that his child may remain in school. Some portion of these tremendous expenditures is, in all probability, wasted on students who have no possibility of completing their academic goals.

It is at this point, of course, that a good guidance program is needed. If a small per cent of the moneys spent for education were invested in good counseling facilities, with adequate personnel, much time, money, and wasted effort would eventually be saved. Results are, in this case, the kind that the public does not see, since a counselor does not publicize his results. The individualized results, even if apparent, are difficult to collectivize and are hardly additive in terms of dollars and cents. For this reason it is difficult for the authorities to obtain the necessary funds for a program of real guidance. This is true in spite of the obvious inefficiency of attempting to educate a person fundamentally not educable at the level which he is attempting.

As a portion of the picture we should also consider the fact that real research in guidance and guidance instruments is burdened even further by lack of funds. Few institutions of learning have enough money allocated for this purpose. However, it would appear to be rather obvious that the two main factors affecting the quality of guidance are its personnel and the instruments used. It is well known at the present that our guidance instruments are far from perfect. Therefore, it would seem that if guidance is necessary, (and this probably may be assumed), research on guidance instruments is also necessary in order that improvement can be made.

To state this in another way more specific to the problem at hand: In the analysis of the individual one important phase is the testing program; the testing program can only be as good as its tests taken in the aggregate; any weakness in the chain of tests produces a weakness in the overall usability; and the American Council on Education Psychological Examination is one of these tests. If the individual testing instruments can be investigated, experimented with, and refined to fit specific purposes even better than they now do, the composite results of a testing program can be improved. The overall good done by the guidance program will also be increased.

This paper is a report of one step which, if taken, may improve the quality of predictions based on the A.C.E. test. Many studies of this type are needed on this and other tests before the available instruments will be practicably useful in all situations. Furthermore, these studies should bring to light mistakes, which, if properly understood and eliminated, would make newer tests much better measuring instruments.

Location of the Study. The setting for this experimental study is Oregon State College, a land-grant institution, located at Corvallis, Oregon. The student population of the college has varied from seven to eight thousand in the school years of 1947 and 1948, and would seem to be on the increase because of a large recent increase in the total population of the state.

The schools within the college are the following: Education; Pharmacy; Forestry; Engineering; Home Economics; Business and Technology; Agriculture; Science; and Lower Division (a two-year curriculum). Each of these separate schools have several departments, which, though administered by the school, may be fairly distinctive in themselves. This is mentioned here to bring into focus the conglomerate nature of the student population, which may be drawn from any area within the state and from other states and even other countries, as well.

The Measurement Device Used in the Study. The American Council on Education Psychological Examination, more commonly known as the A.C.E. test, was used as the measuring instrument in this study.

At Oregon State College this test is given to all incoming freshmen during the week before the actual start of classes. It is not the policy of the college to admit or reject students on the basis of this or any other test. However, it is used as a placement device for sectioning in some courses, and is also used as a counseling tool by faculty advisors and counselors in the various schools and departments of the college, and by the counselors at the College Counseling and Testing Bureau.

Preliminary norms are computed each year for each entering freshman class. It has been found in the past that the median score for the group at Oregon State approximates very closely the same score for students all over the United States. This comparison is possible when final national norms are published in the spring of the year. The range of scores at Oregon State appears to be slightly smaller, but this is to be expected with a smaller sample. Decile rankings for any raw score vary little from year to year, since an apparently successful attempt is made to equate the difficulty of the questions in succeeding editions of the test.

Until 1948, the authors of the examination were Dr. L. L. and Thelma Gwinn Thurstone, and the test was published by the American Council on Education, 744 Jackson Place, Washington D. C. . In 1948 the task of constructing, publishing, and standardizing the test was transferred to the Educational Testing Service, Cooperative Test Division, 15 Amsterdam Avenue, New York 23, New York.

Since the year 1940, the examination has been composed of six sub-tests, of which three are linguistic in nature and three quantitative. The grouping of these subtests is done in the following manner:

Quantitative Tests: (The Q-Score)

- Arithmetical Reasoning
- Number Series
- Figure Analogies

Linguistic Tests: (The L-Score)

- Same-Opposite
- Completion
- Verbal Analogies

The linguistic portion of the tests is weighted more heavily than the quantitative portion, due to the fact that there are one hundred twenty possible items in the former, and only eighty in the latter, the time limits for each

section being devised so that more are usually completed in the linguistic sub-sections. The order of taking the sub-tests has been arranged so that linguistic and quantitative tests alternate. This is apparently done because of the fatigue element. The makers of the test do not recommend that the six separate test scores be used for any counseling, but do believe that there is justification for using the two principal sub-scores as well as the total score for this purpose. (32, p. 2)

Another feature of the test is the practice exercise. Each of the six parts has a practice exercise which is to be completed by the student before embarking upon the test proper. These practice exercises are not timed, but there are approximate time limits set, even though all students should be allowed to finish them before the test proper is started.

Each year's edition has only one form. Answers are indicated by blacking in between dotted lines on a separate answer sheet. One answer is to be so indicated out of a possible four or five choices given. The answer sheet may be scored by hand or with the International Business Machines test scoring machine. Since the task for the student is the same whether it is scored by hand or by machine, only one set of norms is necessary. Test booklets may be used several times since no marks are to be made in

the test booklet, a place for scribbling being provided on the answer sheet.

Time limits and order for the respective sections of the test are as follows:

Arithmetical Reasoning	---	10 minutes
Completion	---	5 minutes
Figure Analogies	---	5 minutes
Same-Opposite	---	5 minutes
Number Series	---	8 minutes
Verbal Analogies	---	5 minutes

Mental ages and intelligence quotients do not exist for these tests in the range of college students.

Beginning with the 1940 edition of the test the separate test questions have been subjected to an analysis on the basis of difficulty. The items have then been selected so as to make gross scores in the successive editions directly comparable. This has not worked out absolutely as planned, but the year-to-year difference has been very small.

Norms are published in the spring for four types of colleges, classified according to academic status. The categories are:

Type I.....Four Year Colleges

Type II....Junior Colleges

Type III...Teachers Colleges

Type IV....Technical and Professional Schools

Wider contrasts exist within each type than are found between them.

Subjects. The subjects employed in this study were a total of three hundred freshman students in the School of Engineering at Oregon State College.

The School of Engineering includes the following departments:

General Engineering

Agricultural Engineering

Chemical and Metallurgical Engineering

Civil Engineering

Electrical Engineering

Mechanical Engineering

Industrial Engineering and Industrial Arts.

The enrollment in the entire School of Engineering for the academic year of 1946-1947 was approximately eighteen hundred students, making it the largest school on the campus. By far the larger percentage of these were veterans returning to school after service.

Freshman students were chosen for this study for several reasons. First, most of the students in this school take a course in General Engineering for their freshman year, branching out therefrom into their respective special branches of the field. Those in Civil,

Electrical, Industrial, and Mechanical Engineering all take the same curriculum for their first year, consisting of the following courses and credit hours:

Engineering Problems	(GE 101,102,103)	2
Engineering Drawing	(GE 111,112,113)	2
Elementary Analysis	(Mth 101,102,103)	4
Engineering Physics	(Ph 101,102,103)	3
English Composition	(Eng 111,112,113)	3
Military or Naval Science		2-3
Physical Education		1
Total		17-18

Furthermore, most of these same courses are included in Agricultural Engineering, Chemical and Metallurgical Engineering, and Mining Engineering in the freshman year. The only real differences in the curricula within the School of Engineering for the freshman year are in Industrial Administration, and Industrial Arts Education, and there were very few students included in the study who were majoring in this region, since most of the students taking these majors seem to transfer there later in their college careers. Therefore, the data obtained on these students not only pertains to engineers in general, but for the most part to the specific courses shown above.

Secondly, freshmen were chosen because the first year seems to be the "eliminative" year for most college courses. This is especially so in the engineering curricula. Holcomb and Laslett (20, p.109) quote studies to show that of 5338 freshmen in engineering in 25 large schools of engineering in the United States, 39.1 per cent of the students enrolled as freshmen did not re-enroll as sophomores. This, therefore, is where a testing and guidance program could be most useful. If it were possible to know exactly who would pass and who would fail before the academic work is started, much wasted effort would be saved, and much acute disappointment avoided.

And thirdly, due to the selective factors causing this large percentage of drop-outs during the freshman year, the sampling and the grade curve should be more normally distributed at the freshman level than at any other higher level. This should mean that any statistical procedure which could be applied to the data would be more valid.

The total number of subjects chosen was three hundred. This number was composed of one hundred each who had taken the 1945, 1946, and 1947 A.C.E. test. It was felt necessary, in the interest of accuracy and usability, to determine whether or not an increased correlation produced by a change in method in one year's test would be

reflected or duplicated on another test.

The first experimental group was composed of one hundred students who were freshmen in engineering in the academic year of 1947-48, and had taken the 1947 A.C.E. test. This group was selected on the basis of grade point averages attained in the first term of this academic year. That is, those were chosen whose grade point average was a 3.00 and above or a 2.00 and below. (This is computed at Oregon State College on the basis of 0.00 for a grade of F, 1.00 for a D, 2.00 for C, 3.00 for B, and 4.00 for an A). These extremes were chosen in order to attempt to increase any differences found, since the procedures used were dependent upon trial and error at first. The supposition is that if the hypothesis of the writer was a workable one the trends would be emphasized, as would the regions in which the workability was highest.

The second group was of two hundred freshman students in the School of Engineering for the academic year of 1946-1947, one hundred of whom had taken the 1945 A.C.E. test, and one hundred the 1946 A.C.E. test. The grades for these students were obtained for the full academic year of three terms. All of the group of two hundred were veterans, and these were used for two main reasons. First, the grades were readily available at the College Counseling

and Testing Bureau for veterans in this academic year. Secondly, it was assumed that the veterans, being older and having had more experience, would have chosen the School of Engineering for more valid reasons. It was hoped that this would exclude the problem of interest to a greater extent than in a group composed of both veterans and non-veterans.

Other than this, there were no selective factors, all two hundred students being taken from the files in alphabetical order--the first two hundred who had taken the A.C.E. test for 1945 or 1946.

There is a definite possibility that the same procedure could be used to advantage in studying records made in other schools on the campus at Oregon State College. This was not possible in this study since the evaluation procedures required such a volume of statistical work. It would seem that whatever would work for general engineering would also be applicable to the entire region of the physical sciences, at least to some extent. This is inherent in the curriculum found in general engineering. Still, this would be impossible to prognosticate definitely without further study involving these different schools.

Sources of Data. The basic facts necessary for the

investigation of the writer's hypothesis were obtained from several sources located on the campus at Oregon State College.

Since the information was readily available for the most part within the files of the College Counseling and Testing Bureau, this source was used for the largest percentage of the data. First of all, it was possible in this department to locate the original answer sheets for the A.C.E. test for any year back to and including 1941. Lists had been compiled of all those who had taken the test in any one year, and these were used to locate the answer sheets. Secondly, this department has in its files for the academic year of 1946-47 a set of cards for each veteran enrolled during that year, showing terms attended, courses taken, year in school, grade point averages for three terms, and other information not pertinent to this study. These cards are separated within the file by schools. This arrangement made it possible to obtain much information without troubling the Registrar's Office.

For the first experimental group, freshman records for the fall term of the academic year 1947-48 were obtained from Mr. W. R. Crooks, Lower Division Head Counselor. Records of individual students were segregated according to school. Grade point averages were given for

the term, providing the basic information necessary for selecting those with averages of 3.00 and above 2.00 and below. On this basis the original test sheets for these students were located and rescored.

The third source of data was the College Registrar's Office. When it was discovered that a few individuals who did not follow the usual pattern were lowering the correlations on the 1946 examination, an attempt was made to trace their educational fates in succeeding years. The Registrar's Office was able to supply some information in these cases not otherwise obtainable.

CHAPTER III

HISTORICAL BACKGROUND

Introductory Remarks. Studies made on intelligence tests in general are legion. The A.C.E. test is one of the most used college aptitude tests, and therefore many of the studies are on this particular test. The selection here is made on the basis of three criteria:

1. Does the study supply general information pertaining to the A.C.E. test and its use in college?
2. Is the study concerned with differential validity of the A.C.E. test, especially as applied to students of engineering?
3. Does the study concern itself with the type of problem as presented by the hypothesis of the writer of this study?

It is to be remembered that there have been many changes from the original version of this test to the form of the present day. This is especially true of any investigations performed on tests published before the year 1941. Since 1941 no changes have been made in number of items, form, or types of sub-test, but before that date a more or less constantly changing test was presented each year. This is not a fault of the test. Rather it was a necessary evil, brought about by a

constant search for types of sub-tests which were more valid in the prediction of college grades. However, this does mean that many of the investigations which were performed on earlier versions of the test have rather doubtful value today. Still, in order to show as completely as possible the background for the present study, a few of these will be included here.

Factors in Student Maladjustment. The use of tests and measurements in the selection and placement of students implies: first, that a selection of college students by some standards other than graduation from high school is either desirable or essential; and second, that tests and examinations are an effective basis for this selection.

In order, then, to determine whether or not a student is able to do college work, we must first know the factors which contribute to success in college or the lack of it. A study attempting to determine this was made by Remmers (37, p.45). He concluded that the order of significant factors in student maladjustment, when grouped under nine general headings, is as follows:

1. Intellectual fitness
2. Emotional adjustment
3. Motivation and interest
4. Educational background
5. Study habits and methods

6. Physical fitness

7. Environmental factors

8. Motor fitness

9. Teaching methods and content

The significant thing here is that the first factor on the list is intellectual fitness. This factor, from the standpoint of a psychometrist, is probably the easiest of all to determine. But other studies should be investigated in order to find just exactly how easily and how well the A.C.E. test accomplishes the measurement of this highly significant factor.

Reliability. Various studies indicate that split-half reliabilities of the total score are from .95 to .97. These figures have apparently remained constant through the years. Stalnaker, (45, p.135), using the 1926 test, found a coefficient of reliability of .95, and Traxler shows reliabilities of .97, with the reliability of the Q-Score being .96 and of the L-Score being .95 on the machine scored edition of 1940 (50, p.53). These figures compare very favorably with such figures given for other intelligence and college aptitude tests.

Retest Gains. Closely related to reliability are studies of test-retest situations using the A.C.E. test. However, real reliability is not measured in the studies cited here due to the large amount of time which elapsed

between the original test and the retest in all cases. The investigations were designed, rather, to test gains in intelligence as measured by this test while in attendance at college and to determine the difference, if any, in predictive value between tests taken by a group while still in high school and upon entrance into college.

Barnes (2) found appreciable gains during the first two years of college. The gains in L-Scores were more marked than in Q-Scores. On the first test the mean of the subjects was nine percentile points above the mean of the national scores; while on the retest the mean of the subjects tested was thirty-four percentiles above the national mean. The coefficient of correlation for initial test-retest was .78.

Hunter (22) retested 276 women during their four-year college course. The retest at the end of the freshman year showed a mean gain of 23 in percentile rank over the initial results; at the termination of the sophomore year the mean gain was 24; at the end of the junior year 26; and at the end of the senior year 31. Correlations between initial tests and retests were, respectively, .81, .85, .84, and .83. Students with lowest initial scores made the most improvement; of the 276 students, only 14 showed losses.

Heilman and Congdon (19) used two groups of 115

students, matched on the basis of age on whom entrance scores on the A.C.E. test were available, for another investigation of this type. They retested these groups of students after two and three years of college work, respectively, in order to determine if those in college who had remained longer made the larger gain on the retest. They concluded, since no reliable difference between the gains was obtained, (though there was a small gain in favor of the group that had remained longest in college) that there was no evidence that the activities of their college influenced the size of the intelligence test scores.

These studies would appear to show, at least on the surface, that intelligence is improved somewhat by college attendance. However, upon analysis, the reasons for this improvement would more probably be: A process of becoming test-wise, causing an improvement in scores on all tests of this type; an improvement in vocabulary causing a gain in L-Score; a lack of the pressure situation evident in freshman examinations at most colleges; an improvement of speed of reading through practice; and more efficient method, that is, better application to the task at hand.

Thompson (47) studied 106 freshmen at Carleton College who had taken the 1935 form of the A.C.E. test as high school seniors and the 1937 form upon entering

college. Correlations between raw scores on the two forms of the A.C.E. test taken thus, eight months apart, were:

Completion	.696
Artificial Language	.784
Analogies	.530
Arithmetic	.707
Opposites	.730
Gross Score	.869

Correlations between the two forms of the test and grade point averages were:

1935	.568
1937	.583

He concludes (47, p. 233):

> The . . . results would seem to indicate that when dealing with large groups for purposes of predicting college success it makes little difference whether the A.C.E. test was given during the last year of high school or upon entrance to college.

A test is said to be reliable when it functions consistently. On the basis of the studies cited here one may conclude that the A.C.E. test is a reliable measuring instrument.

<u>Validity in General Prediction</u>. The second criterion for determining the worth of a measurement device is its validity, i.e., does it measure what it purports to measure. There have been many investigations as regards

the correlation between scores on the A.C.E. test and grades in college. Some of these apply to grades in general and some to grades in specific subjects. The former will be discussed in this section, and the latter in the section following.

Studies of the general validity of the A.C.E. test are fairly consistent in the results found. One study, including more correlations of A.C.E. scores and college grades than any other, summarizes the situation by showing that out of 34 coefficients of correlation reported on various editions, the range was from .27 to .62, with the median coefficient being a .48. For general mental tests and scholarship, in this same study, the correlation was a .44. (41, p.59-61) Obviously, most of the coefficients of correlation found in this study do not give an exact basis for predicting success of college students. The important thing, however, is that a single test of about an hour's duration is at least as good as the average of all marks earned in high school over a period of four years as a measure of probable success in college. And if one is to consider the greater-than-usual variance in high school requirements and grading systems found in the state of Oregon, it may readily be seen that, in spite of some limitations, this measurement device could be quite

useful.

Another general study applicable here is that by Stalnaker (44), in which she found that 49 per cent of an entering class discontinued their college careers before the end of the fourth year. In each semester a larger proportion of withdrawals came from the lower half than from the upper half in intelligence scores.

Schmitz, (42), in a study of entering freshmen at St. Benedicts College at Atchison, Kansas, in the years of 1934-35 and 1935-36, found that the high school scholastic quotient was slightly above the A.C.E. test in predictive value. This study gives the comparative value of several criteria for predicting college success as measured by the individual scholastic quotient, i.e., the total number of quality points earned divided by the total number of hours of work carried. The second best instrument, however, was the A.C.E. test, showing a correlation of .583 between scholastic quotient and A.C.E. score. He also notes (42, p.470):

> A student whose score on the psychological test is below one hundred eight has only one chance in eight of doing average college work or better . . . By far the majority of the individuals rating low on the psychological test have very little hope of success in college.

Correlation coefficients would seem to vary somewhat due to the type of curricula found in the college.

At Oklahoma Agricultural and Mechanical College, Stillwater, Oklahoma, Rigg (39) found slightly different results. Using the A.C.E., correlations obtained between intelligence and four years scholarship for 228 students was .42; between intelligence and achievement (as measured by the Carnegie Achievement Test for sophomores) for 205 students was .56.

On an intensive study of 140 students at Phillips University, Quaid (36) discovered the correlation between A.C.E. scores and first semester average grades to be a .408. Using second semester grades the correlation was .412. He points out (36, p. 373):

> . . . relatively greater accuracy in the upper and lower deciles of a ranked distribution permits the practical use of smaller correlations for prediction than traditional thought has sanctioned.

Attempting to discover the differences between degree, combination, and terminal students at the University of California College of Agriculture at Davis, California, Peterson (35) found correlations which are comparable to other studies. Correlation between general college scholarship and total scores on the A.C.E. test for III degree students was a .47; for 97 combination students was a .30; and for 260 terminal students was a .47.

The results of the investigations shown here would indicate that the A.C.E. test is valid in most situations. Correlations, though not extremely high, are higher for this test than for others of the same type for the most part.

Differential Validity. Investigations have been made of the relationship between the A.C.E. test to specific curricula or to single courses. It is probable that the A.C.E. test was specifically designed with this type of prediction in mind. In 1925, shortly after the first tests were used, Thurstone (48, p.284) wrote:

> If our objective is to predict average scholarship, then, in general, the formal linguistic tests are the most desirable. But if our objective with the psychological test is to ascertain aptitudes for particular subjects, then certainly the discriminating value of the ordinary linguistic forms of intelligence test shows a poor performance.

It is therefore, highly pertinent here to investigate studies having to do with the differential predictive value of this instrument. However, since this study itself pertains only to engineering students, only a few citations will be given showing the relationship of the test to other types of curricula, and the main emphasis will be on those investigations having to do with engineers. First, how valid is this examination when dealing with subjects other than engineering?

The A.C.E. and Minnesota Paper Formboard tests were used by Bryan (5) to find the extent to which measured intelligence contributes to success at the School of Fine and Applied Arts Pratt Institute. For the entire group of approximately one thousand students tested, the mean percentile ranking on the A.C.E. was a 49. But when the group as a whole was taken and A.C.E. test results and achievement in art courses were correlated, the correlation coefficient was a negligible .16. For students in Architecture, the A.C.E. test results and average marks in art was .37, and for these same students average marks in design correlated .22 with A.C.E. results. The corresponding coefficients of correlation for art education students were, respectively, .05 and .02; while for students in design they were, respectively, .12 and .09. The results on the Minnesota Paper Formboard test were no more valuable, the correlations ranging from .17 to .33. Results of this study would seem to indicate that success in these courses in fine and applied arts depends on abilities other than those measured by either of the two tests employed.

Harrison (17) found that prospective teachers at Parks College, over a period of ten years, were slightly but not significantly superior to non-teacher students,

as measured by the A.C.E. test. Seagoe (40) found prospective elementary school teachers to be superior on the A.C.E. test. The mean score of 125 such men and women gave tham a percentile rating of 81. The middle sixty-eight per cent of the group achieved percentile ranks of from 56 to 96. However, these two investigations do not seem to agree altogether with the study by Schneidler and Berdie (43), who tested students in the several colleges and professional schools within a university. They concluded that on the whole there is a "tendency" for college aptitude of the "average student" in colleges of education, agriculture, dentistry, and pharmacy to be lower than in science, technology, business, literature, and arts. There was the usual considerable overlapping, and the differences became much smaller as students were compared from freshman to senior year. It was pointed out that in some of the higher ranking groups, a selective process had already been at work, for their "freshmen" had satisfied a prerequisite of two years of study in a liberal arts college.

Lanigan (25) studied the effectiveness of the Otis, the A.C.E., and the Minnesota Speed of Reading tests for predicting college success. Her study was done in the Liberal Arts College in 1943 at Boston University.

Subject matter fields were grouped under the six areas of English, Social Studies, Language, Mathematics, Science, and Fine Arts. The group tested consisted of 129 women and 34 men.

Correlations between the A.C.E. test and the six subject matter fields were: English .325; Social Studies .501; Languages .222; Mathematics .324; Science .442; Fine Arts .364. The large variations in correlation could perhaps be explained by the rather small number of subjects used. In regard to the validity of the A.C.E. over the other instruments used, she writes (25, p.295):

> These findings indicate the American Council Psychological Examination is a more usable instrument for predicting a critical score above which groups succeed and below which many tend to fail or receive low marks. The high achieving group of students were found to have a mean total score of 127.02, while the low achieving students obtained a mean score of 112.36 on this instrument.

Wagner and Strabel (52) studied a total of 661 men and women entering the University of Buffalo between 1925 and 1929. They correlated performance in physical science subjects with various measures, the A.C.E. test faring rather poorly. On 414 men A.C.E. total correlated .18 with average physical science and on 227 women this correlation was a .25. Between introductory college physics and A.C.E. total the correlation was .19 for 176

men and only .03 for 32 women, again being the poorest measure employed. The parts of the A.C.E. test were also correlated with college physics. These correlated no better or no worse on the whole than the total score.

Kent and Schreurs (24) studied the predictive value of four separate factors in freshman English and Mathematics at Northwestern University, using 524 students who entered in 1924-25. One of the factors was the "Mental Alertness Score"--the A.C.E. test. The following table is excerpted from their study. (p. 244)

TABLE I

Mean Mental Alertness Scores for Students in Algebra and Trigonometry Divided According to Grades

	Grade					
240 cases	F	E	D	C	B	A
Algebra Grade	-2.0	-1.0	0.0	1.0	2.0	3.0
Mean Mental Alert-Score	39	41	48	52	50	64
Per cent of class	32	7	15	26	14	6
328 cases						
Trigonometry grade	-2.0	-1.0	0.0	1.0	2.0	3.0
Mean Mental Alert-Score	24	53	55	62	60	71
Per cent of class	24	6	13	28	17	11

It is to be noted here, that even though there would seem to be an unusually high percentage of failing grades, the mean score of the testing instrument advanced as the grades advanced for the most part. The one exception was the ability to differentiate between grades of B and C. This would not be considered to be a bad failing in an advisement situation.

The investigations cited thus far have been increasingly close to the problem at hand. Following are four studies which treat the immediate problem of the present paper, i.e., how well does the A.C.E. test predict college grades for engineering students.

The highest correlation was found by Holcomb and Laslett (20), at Oregon State College in 1932. This is of particular interest, since it was done at the same school as the present study. However, it would seem to be likely that some factors have entered into the situation which were not present at the time of the previous study, since general correlations at the present time do not approach the .555 found by these investigators.

McGehee (28, 29) has conducted two separate investigations on engineering students at North Carolina State College of Agriculture and Engineering. On the first, in 1938, he used 589 freshmen, of whom 299 were in engineering.

He found definite tendencies for higher deciles on the A.C.E. test to receive higher grades, and this is shown by the following table excerpted from his study (p. 224)

TABLE II

Percentage of Grades at Each Decile Group Level of the Total Number of Each Type Grade Given

Grades	A	B	C	D	F
	(Figures shown are all percentages)				
Deciles 8-10	63.44	47.85	30.81	21.94	14.04
Deciles 4-7	26.54	39.47	47.95	47.79	42.61
Deciles 1-3	9.96	13.65	21.24	30.25	43.34

In his second study, in 1943, he used 383 engineering students, and found the correlation of grades with the 1939 edition to be a .48.

At the University of Florida in 1935, Mosier (33) investigated the relationship of the five parts of the A.C.E. test of that date with grades in several curricula, one of which was engineering. Several of his conclusions would seem to be quite applicable here. (33 p.521-522)

1. That there are several abilities measured by the sub-tests in the A.C.E. Psychological Examination which are independent of the ability measured in common by all the tests.

2. That these are required differentially for success in the several curricula offered at the University of Florida.

> 3. On the basis of the abilities required for success, the curricula are roughly separable into two groups, one requiring the abilities measured by completion and opposites tests, in addition to the common ability, the group requiring, in addition to the common ability, the abilities measured by the analogies and arithmetic tests.
>
> 4. On the basis of the content of the curricula in these two groups they may be designated as liberal arts-social sciences, and natural sciences, respectively.

Differentiation of Q- and L-Scores. A few investigators have concerned themselves with the differential predictive values of the quantitative and linguistic sub-scores.

Super (46) used the 1938 edition of the A.C.E. test and several other tests to attempt to determine the relationship between the results and special abilities on a group of 123 high school juniors and seniors. His conclusions in relation to the A.C.E. test are: (p.225)

> Consideration of the guidance implications of these findings suggest that the quantitative scores are not suitable for the differential prediction of achievement in mathematical subjects, but may be of value in diagnosing and treating students with reading difficulties, whereas high linguistic scores are indicative of general academic ability . . . in general mathematics . . . the best predictions are . . . given by total scores.

This study, then, does not seem to uphold the expectations of the makers of the examination. However, in the opinion of the writer, this may be due to

the fact that this study was on a somewhat limited group of high school students. At least, other studies at the college level should be completed before too definite conclusions are drawn on this matter.

Barnes (3, p.581-82) at the University of Illinois, attempted to find the answer to the question: "Does the study of mathematics in college affect ability to score on the quantitative test?" In a test-retest experiment, seventy-five university students who had completed the freshman-sophomore years without having had a course in mathematics were compared to 40 mathematics students who had also completed their first two years of college work with an average of over eight semester hours of mathematics. A slight gain in favor of the mathematics group is found, though it is not statistically significant. It is interesting to note, however, that the mean gross score of the mathematics group was 121.45 and the non-mathematics group was 100.00, showing superior quantitative abilities as measured by this test for those choosing mathematics. This may indicate that the factor of interest is operating here.

Reweighting and Multiple Correlation. It is apparently possible, as is evidenced by several studies, to reweight the separate portions of the A.C.E. test, or to do multiple correlations and construct regression equations,

in order to increase the predictive possibilities of the examination in any particular situation. Two of these studies have been done in connection with engineering students.

At Alabama Polytechnic Institute, Waits (53) investigated the correlation of the sub-tests of the A.C.E. and 212 students in various branches of engineering, finally reweighting these sub-tests and again correlating. He found that the sub-tests and scholastic average correlated in the following degree:

Completion	.3286
Artificial Language	.2599
Analogies	.3726
Arithmetic	.4307
Opposites	.3333
Total Score	.4484

Noteworthy here is the fact that the Arithmetic sub-test shows by far the highest correlation, agreeing with other studies of the same type on engineering students. By a reweighting of the sub-tests, he increased the predictive value of the examination, though not greatly. The correlation was raised between the scholastic average and the total score by about .04. He significantly concludes (53, p.271):

> Enough difference was found . . . to indicate that a test might be constructed which would have differential predictive value for various college groups.

Studies on "Caution Factor". Only two investigations concerned the "caution factor", and neither of those used the A.C.E. test. However, one did use a system for determining caution factor which was exactly the same as used here, and the other used a similar formula. Both were concerned with the application of this factor to intelligence tests.

The study by Crooks (12) was intended to measure unknown personality traits by means of the ratio of the number of omissions to the number of errors made on the Otis Self-Administering Test of Mental Ability. Scores on this test, the Bernreuter Personality Inventory, and the Allpert-Vernon Scale of Values, were used, for 284 women and 283 men at the University of Conneticut in 1937.

The various personality traits measured by the Bernreuter Personality Inventory and the Allport-Vernon Scale of Values, when correlated with the Otis ratios, were not high, but might be considered to be very significant here. For instance, in correlating with the Bernreuter B1-N Scale (the "neurotic" factor), he found correlations of -.113 with the males, -.209 with the females, and -.185 for both. This would indicate that for

those people with a tendency to be neurotic, there was also a tendency to be "rash". Putting this into the academic situation, it is rather obvious that a neurotic person will not do as good work as intelligence tests indicate he should. If this is related to "caution factor", then it follows that the "rash" individual will not do as well as could be expected from inspection of intelligence test results. Further than this, he found positive correlations of .153 for males, .239 for females, and .212 for both with the Bernreuter B2-S Scale (Self-Sufficiency). This might indicate that the self-sufficient person, who has perhaps had success in college work, has also discovered the value of "caution". These generalizations are, of course, hypotheses, and have not been proved in any manner. However, it is at least possible that these are some of the reasons for positive correlations in one case and negative correlations in the other.

The high reliability coefficients found in this study have been mentioned previously (p. 18). Another significant figure found here was a correlation of .709 between intelligence as measured by the Otis test and the "caution factor".

Brown (4) investigated the relationship between a "caution factor" and intelligence test performance, using

the Thorndike Intelligence Examination for High School Graduates. The subjects for the portion of his study that applies to this paper were 375 students of Columbia University, and the caution index on this test was the number of wrong answers as converted to a percentage scale on the basis of the group. For this group he discovered for forty individuals a correlation of .71 was obtained between the caution index and the actual number of times each was placed on probation during the time in which he was in college. It may be objected that the number of cases is too small for any valid conclusions to be drawn on the basis of this. Even so, this might indicate a trend, and certainly should not be altogether overlooked.

Other correlations applicable here which he found were: Between scholastic index and caution index a .27, between scholastic index and intelligence score a .45, and between intelligence score and caution index a .40. (p. 379)

The equation as determined for predicting the probable scholastic achievement of individuals when their intelligence score and caution index are known is as follows: (p. 383)

X_1 equals 2.72 plus .02 X_2 plus .06 X_3

where:

X_1 is the scholastic index on a 15 point scale,

X_2 is the caution index, and

X_3 is the intelligence score or I.Q.

He gives as an illustration of the differences occurring in scholarship between persons with a high and with a low caution index the following case: (p. 384)

> A group of 16 individuals making a low average caution index, and one of 17 individuals making a high average caution index, although all were of approximately the same intelligence as evidenced by their scores on the Thorndike intelligence examination, showed a difference of 8.6 points in favor of the more cautious group as regards their average scholastic performance.

Two of his conclusions which seem to apply here are (p.383, p.384):

> . . . in general, the caution factor seems more evident in persons of high intelligence than in those lower down in the scale . . .

and,

> While the correlations were not high enough to show that the cautious individual invariably excelled in scholarship, they nevertheless served to indicate that the trend [italics] is in that general direction.

<u>Summary of Chapter III</u>. If tests and measurements are to be employed at the college level, they must measure at least some of the factors which are necessary at this level. The A.C.E. test is set up to measure intellectual fitness, one of these factors. According to the studies

cited, it would seem to be a highly reliable test. Validities range greatly, though in general it would seem to be as good a predictive instrument as is available. For specific subjects the validity is not as high, this discrepancy possibly being due to other special factors in many cases which the test is not even attempting to measure. For engineering students correlations are usually approximately as high as correlations for general college aptitude, the test seemingly being the best predictive instrument available for students in this curriculum. By multiple correlation, regression equations, and reweighting, it is possible to increase the predictive value to a small extent. This would indicate that for specific curricula a test might be constructed which would be better than the present test in that one situation. The studies on "caution factor" tend to support the contentions of this thesis, at least to the extent of proving that there is such a thing, that it has some relationship to scholarship, that it is reliable, and that it is connected possibly with both personality and intelligence.

CHAPTER IV

METHODS EMPLOYED IN THE STUDY

In making this study the writer lays no claim to originality of method or of interpretation. The attempt has merely been made to substantiate the hypothesis of a "caution factor" and to discover at least some of the ways in which wrong answers on the American Council on Education Psychological Examination may be evaluated.

The question of the further interpretation of correct answers has been undertaken only insofar as this is connected with the interpretation of those missed. Also, the matter of items not tried could be further investigated, and this was not attempted here. Furthermore, it is probable that other statistical methods might have been used to advantage. The purpose of this paper, however, is not to finish completely the analysis. This is hardly possible for one investigation on a problem of this type. It is hoped that others may become interested enough in this phase of testing research to carry on into these other regions.

The First Experimental Group. The first experimental group, as has been previously explained, was chosen from 1947-48 engineering freshmen. Grade point averages were obtained for the first one hundred on the alphabetical

list for whom an A.C.E. test was available in the 1947 edition. The original answer sheets for these men were then taken from the files at the College Counseling and Testing Bureau. The intention of the experimenter was to use this first group as a determinant, in order to find whether or not the hypothesis was workable to any extent whatsoever, and, if possible, in what regions of the test the workability was greatest.

The first step in the actual analysis was to re-score these answer sheets, determining the number attempted on Page A and Page B, the Q-Score and L-Score sides, respectively. This was done with the aid of the International Business Machines test scoring machine, after first having constructed a key with holes punched in all spots where a possible answer could have been indicated. This gave a dial reading of the total number attempted.

Then, after the scores for correct responses and attempted responses had been found, a list was made, using numbers for each student. These six scores were noted for each: Q-Score, L-Score, Total Score, Q-Attempted, (the number attempted on Page A, the quantitative portion), L-Attempted, (number tried on the linguistic portion), and Total-Attempted, (Q-Attempted plus L-Attempted), as well as grade point average.

First, "caution factors" were computed for quantitative, linguistic, and total scores, by dividing the number of responses scored as correct on each score by the number attempted for each score. These figures were then added to the lists.

Correlations were then run between grade point average and all scores and ratios thus far determined, adapting the Monroe method of simple correlation to the Marchant Calculator (1, pp. 1-8). This method is further discussed in a later section of this chapter.

Then, depending on the correlations found above, combinations were made of the various scores. The Q-Score was multiplied by the caution factor of the Q-Score, the L-Score multiplied by caution factor of L-Score, and the Total Score multiplied by the caution factor of the Total. These scores are known, for the purposes of this study, as the "caution factor Q-Score", "caution factor L-Score", and "caution factor Total Score". The scores found by this method were then correlated with grade point average.

Then, since the scores having to do with the quantitative regions of the test were correlating to a higher degree than others, a combination of the Caution Factor Q-Score plus the L-Score was computed and correlated. The correlation here being slightly better than the correlation for the usual total score, the procedure of

doubling the Caution Factor Q-Score and adding this to the L-Score was then computed and correlated.

Figures for all scores computed for this group are available in Appendix A. Correlation figures will be found in the following chapter.

The Second Experimental Group. Since results for the first group made it appear that there was at least some value obtained by using a "caution factor", another group of engineering students was chosen, this time without any segregation as to grades received. The group selected was of two hundred students, who were freshmen in the academic year 1946-47, one hundred having taken the 1945 A.C.E. test, and the other hundred the 1946 A.C.E. test.

This second experimental group was dealt with in much the same manner as the first, excepting for four changes made in the interests of greater accuracy, facility of operation, and further experimentation. Differences were: (1) The grades were obtained for the full academic year of three terms, and then averaged; (2) The keys used in the test scoring machine had holes punched only for all wrong answers, rather than all possible answers; (3) The figures were entered on 3" by 5" cards, one to each student, rather than on lists; and (4) Other combinations of factors were used near the end of the

experimentation in addition to those which had proved useful in the first experimental group.

Grades were used for the full freshman year in college, and averaged, since it was thought that this would give a more complete picture of the actual usability of the instrument than would a single term's grades. However, it is to be noted that this may perhaps have cut down on the correlation figures, because some of the most ill-adapted students would drop before the three terms had elapsed.

The keys used on the test scoring machine in the first experimental group gave the total number attempted directly as a reading on the dial. However, this same procedure was not followed for the second group. The second method, of punching the keys only for wrong answers, gave a score which could be relied upon to a greater extent, and had only to be added to the correct score in order to give the number attempted. This method also gave, without further manipulation, a wrongs score.

The figures for this group were entered on 3" by 5" cards in order to make for easier arrangement and manipulation during the course of the study.

Toward the end of the experimentation, when it became evident that there was something to be gained by

the application of the caution factor, simpler methods of applying this factor were tried--methods that could be used without too much manipulation of figures.

Distributions were made of all two hundred students and deciles computed for number correct, number of errors, and number attempted. The first of these is the usual figure used at Oregon State College for placement and faculty counseling. The other two deciles could be termed the "error decile" and the "speed decile". These decile rankings were then used as the basis of classifying the students into groups. Those in "group 1" were the students who had received a decile of less than five on all three of the rankings; those in "group 2" had one of the three deciles at five or above; in "group 4" all three deciles were five or above. These groups were then correlated with grade point average, and used in combination with the deciles to correlate.

Right minus wrong scores were then determined for Q- and Total Scores, and these were also correlated with grade point average.

Furthermore, since it had been consistently discovered that the 1946 A.C.E. results did not correlate as highly as the 1945 test results, and this was most probably not due to the test, but due to the group of

subjects, twelve of the most outstanding deviates were picked out, and a further investigation was made into their educational fates at the Registrar's Office. This was done in order to determine, if possible, what factors were preventing these men from following the usual pattern.

Results found with this group will be seen in the following chapter, and raw figures for all computations are located in Appendix B.

The Correlation Method. As before stated, the Monroe method of simple correlation was applied to the Marchant Calculator for the correlations in this study. This system makes it possible to correlate from merely two columns of numbers applying to the same individuals, and obviates the necessity of further computation or distribution of data, excepting for substitution into the formula.

The formula, as given in Monroe's explanatory booklet (1, p.3), is as follows:

$$r = \frac{N\Sigma AX - (\Sigma A \times \Sigma X)}{\sqrt{[N\Sigma A^2 - (\Sigma A)^2] \times [N\Sigma X^2 - (\Sigma X)^2]}}$$

where:

r = Correlation coefficient

Σ = Sigma

ΣA = Sigma A, or the sum of column A

ΣX = Sigma X, or the sum of column X

ΣA^2 = Sigma A squared, or the sum of the squares of individual items of column A

ΣX^2 = Sigma X squared, or the sum of the squares of individual items of column X

ΣAX = Sigma AX, or sums of products of items of column A times column X

$(\Sigma A)^2$ = Sum of the A terms multiplied by itself

$(\Sigma X)^2$ = Sum of the X terms multiplied by itself

N = Number of subjects or observations.

It may readily be seen that this is the Pearson product-moment formula for correlation coefficient, the only difference being in the letters used; Y usually being used in place of X and X in place of A as used here.

Limitations of the Methods Used. The limitations of the findings possible in this study can probably be summarized in the categories given below.

1. The reliability of teacher's grades. Any correlation using grade-point average as one of the correlates is not likely to be higher than the reliability of the grades. Traxler points out (51, p.177) that "the reliability of teacher's marks seldom exceeds .65", though in the study by Quaid (36) a reliability of .78 was found for teacher's marks at Phillips University.

2. If grouping of data is done such as is done here in a part of the study into deciles and further into "groups", this restricts the variability and one

would expect correlations to be lower than if the data remained in its original form for correlation.

3. The fact that for engineering students, in the main, the A.C.E. test is a speed test, rather than a power test. For those for whom the "caution factor" reached 1.00, there was no actual complete measurement of caution.

4. The complex of other factors entering into the college academic situation. Any number of things can destroy the correlation between academic abilities and academic standings, ranging, for example, from the student who spends all his time building a house for his wife and child to one who spends his time in extra-curricular activities of all sorts because of a lack of interest in the actual subject matter presented to him in class time.

CHAPTER V

RESULTS OF THE STUDY

The materials in this chapter have been organized around a series of tables showing the actual numerical and statistical results of this study. These tables are designed to indicate, without further information, the findings on this group of subjects. However, several new concepts and methods of arranging data were introduced, which could not be explained fully in tabular form. Consequently, each table will be accompanied as closely as possible by an explanation and interpretation of the terms used, as well as an interpretation of the results found.

Proof of a Normal Group. The first and most necessary criterion of data which is to be subjected to statistical analysis is that the subjects must be unselected, but still of sufficient number to follow the normal curve. If the subjects in this study, for instance, showed a skewed curve in relation to their grades, then the results found here would not be applicable to other groups of subjects. However, if the distribution of the grades followed a normal pattern of variability, then the group used is a normal group, and any findings would probably apply elsewhere.

TABLE III

DISTRIBUTION OF GRADES FOR 200 ENGINEERING STUDENTS, BY TERMS, SHOWN GRAPHICALLY AND NUMERICALLY FOR THE ACADEMIC YEAR 1946-1947

Grade Point Av. For One Term	Number Receiving Grade Point Average in Any One Term*
0.9 xxx	3
1.0 xxx	3
1.1 x	1
1.2 xxxxx	6
1.3 xxx	4
1.4 xxxx	5
1.5 xxxxx	6
1.6 xxxxxxxxx xxxx	15
1.7 xxxxxxxxx xxx	14
1.8 xxxxxxxxx xxxxxxxxx	20
1.9 xxxxxxxxx xxxxxxxxx xxxxxx	27
2.0 xxxxxxxxx xxxxxxxxx xxxx	25
2.1 xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxxxx	48
2.2 xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxxx	57
2.3 xxxxxxxxx xxxxxxxxx xxxxxxxxx xxx	34
2.4 xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxxx	47
2.5 xxxxxxxxx xxxxxxxxx xxxxxxxxx x	31
2.6 xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxxxxxx xx	42
2.7 xxxxxxxxx xxxxxxxxx xxxx	25
2.8 xxxxxxxxx xxxxxxxxx xxxxxx	27
2.9 xxxxxxxxx xxxxxxxxx xxxxxxxxx xxxxx	36
3.0 xxxxxxxxx xxxxx	16
3.1 xxxxxxxxx xxxxxxxxx xxxxx	26
3.2 xxxxxxxxx xxx	13
3.3 xxxxxxxxx xxxxx	16
3.4 xxxxxxxxx xxxxxxxxx x	21
3.5 xxxxxxxxx xxx	13
3.6 xxxxxxxxx x	11
3.7 xx	2
3.8 xxx	3
3.9 x	1
4.0 xxxxxx	7
Total	600

* Each x equals grade point average received by one student for one term

In order to illustrate the fact that the grades obtained by students used in this study follow very closely the normal curve, and still show the number found at any grade point level, it was necessary to resort to a graphic table: Table III. Referral to this table will show at a glance the bilateral symmetry of the data. There is a slight negative skewness, the crude mode being located at 2.2, and the median at 2.4. That this skewness is rather slight, however, is evident from inspection of the curve shown by the graphical arrangement of the x's in the table. It is also possible from inspection to classify this curve as a mesokurtic or normal curve.

It is therefore apparent, since all correlations were done with grade point average as one of the correlates, that the basis for the results of this study is statistically sound.

Results With the First Experimental Group. The point discussed above does not, of course, apply to the first experimental group. This group was chosen so that there would be no normal curve of grades, as students having grades in the middle regions were arbitrarily eliminated. Such a selective factor was introduced in order to emphasize the trends, since the purpose in using this first group was to discover whether or not anything was to be

gained by the application of a "Caution Factor" to the data. That this procedure was sound is evidenced by the fact that higher correlations are found on this group than on the normal group following, and that trends discovered here tend to be repeated in the analysis of the second group.

Table IV, on the following page, shows the correlations discovered between A.C.E. scores and grade point averages for one term. Some of these scores require further explanation here.

Q-Score, L-Score, and Total Score, are, of course, the usual scores found on this measuring instrument. They are specifically the number correct in the quantitative, linguistic, and combined regions of the test, respectively.

Q-Attempted, L-Attempted, and Total Attempted scores may be defined as the total number of questions marked, either correctly or wrongly, on the quantitative, linguistic, and combined regions of the test, respectively.

The next three scores bear the "caution factor" label, in each of the regions already indicated. The scores for correlation were determined for these by finding the ratio of the number correct divided by the number attempted for each individual. This score is expressed in terms of hundredths, ranging from .00 to 1.00.

TABLE IV

CORRELATIONS BETWEEN 1947 A.C.E. SCORES (INCLUDING CAUTION FACTOR SCORES) AND GRADE POINT AVERAGE FOR ONE TERM WITH ONE HUNDRED ENGINEERING STUDENTS

SCORE	CORRELATION
Q-Score	.515 ± .05*
L-Score	.374 ± .06
Total Score	.487 ± .05
Q-Attempted	.387 ± .06
L-Attempted	.329 ± .06
Total Attempted	.383 ± .06
Caution Factor Quantitative	.342 ± .06
Caution Factor Linguistic	.162 ± .06
Caution Factor Total	.302 ± .06
Caution Factor Q-Score (Q x CFQ)	.535 ± .05
Caution Factor L-Score (L x CFL)	.357 ± .06
Caution Factor Total Score (T x CFT)	.478 ± .05
Caution Factor Q-Score plus L-Score	.492 ± .05
2 Caution Factor Q-Score plus L-Score	.531 ± .05

* Probable errors used throughout.

This computation was completed for quantitative and linguistic regions of the test, separate and combined.

The Caution Factor Q-Score is a combination of other scores defined above, as are the Caution Factor L-Score and the Caution Factor Total Score. These are obtained by multiplying the Q-Score by the Caution Factor Quantitative ratio, the L-Score by the Caution Factor Linguistic ratio, and the Total Score by the Caution Factor Total ratio, respectively.

The last two combination scores are, again, composed of scores mentioned above, and seem self-explanatory as expressed in the table.

Interpretation of Results--First Experimental Group. On this group of engineering students, segregated as it was, higher correlations were found than in the second experimental group. One of the highest of these was on the Q-Score, a correlation of .515 between this score and grade point average for one term being obtained. It is to be noted that this is a much better correlation than is found with the L-Score, and slightly better than that with the Total Score, these being, respectively, .374 and .487. This would seem logical, also, in view of the predominantly quantitative nature of the subjects present in the engineering curriculum.

It should be pointed out here, then, that at least with this group of engineers, the Q-Score was a better measuring instrument than either the L-Score or the Total Score when it came to locating those individuals who would fall at the two extremes of the grade distributions.

The correlation between Q-Attempted and grades was .387, between L-Attempted and grades was .320, and between Total Attempted and grades was .383. This would seem to indicate to the writer that the test is primarily a speed test, due to the fact that mere speed is not usually considered to be an indication of college ability. Since this is one of the original assumptions of the study, these correlations are significant. The question which cannot be answered, however, is: Does this correlation mean that speed has some effect on grades, or does it mean that the test was not powerful enough to test these individuals completely? In any case, it is apparent that, as the test is now constructed, there is a positive relationship between how many questions are attempted and how well the person does in college.

It follows that, in the analysis of the "caution factor" scores, there is the possibility of the variability of the "caution factor" being restricted because of the ease of the test, (shown by the above correlation with speed). However, in spite of this, correlations with

the "caution factor" expressed as a ratio of the number of items correct divided by the number of items attempted were high enough to indicate that something was being measured by this procedure. The correlation with "caution factor" on the quantitative sections of the test was much higher than that on the linguistic section. The latter was only .162, which is within four probable errors of a chance score, and therefore not significant statistically. However, the Caution Factor Quantitative ratio correlated with grades to the extent of .342, and the Caution Factor Total to .302. Both of these are significantly high.

At this point a combination was attempted, since both correlations were positive and significant. It was reasoned that if the Q-Score gives a high correlation with grade point average, and if the Caution Factor Quantitative does the same, a combination of the two should be still higher. This, of course, is true only if the two factors being combined do not counterbalance. There apparently is some such effect, but the correlation of the Caution Factor Q-Score with grades was a .535, which is the highest correlation found in this study. This indicates that this combination gives a slightly better predictability for extremes than does the usual Q-Score.

When the same combination was attempted for the linguistic and total scores, however, the resultant combination showed subtractive tendencies, and the correlations were slightly lower than original scores. This would seem to be due to the fact that the ratio of Caution Factor Linguistic showed such a low correlation. Since this would affect both these scores, apparently chance combination intered into a greater extent than had been anticipated.

Then, since the Caution Factor Q-Score had correlated so highly, this was combined with the L-Score, first on an equal basis, and then doubly weighted. This gives, in effect, a total score for the first, and a weighted total score for the second combination. These figures were then correlated with grades and the result was somewhat encouraging. Correlating the first gave .492 and the second gave .531. Comparing these with the .487 obtained for the usual total score, a fairly definite improvement may be noted.

The work with the first experimental group was terminated at this point. It had been demonstrated that the "caution factor" on the A.C.E. test was related to grades received. Combination scores using this "caution factor" improved predictability to some extent. The

quantitative regions of the test had been found to relate to engineering courses more closely than the linguistic regions. Correlations in one case had been improved .044. Work was terminated due to the fact that the group had served its purpose in pointing out these relationships; and because the figures were becoming unwieldy, necessitating much manipulation of data, and some simpler method was desired.

Results and Interpretations For the Second Experimental Group, Using the 1945 A.C.E. Test. The correlations found for the one hundred students in the second experimental group who had taken the 1945 A.C.E. test are found in Table V. This table is located on the page following. These results are separated from other results in this portion of the study for the purpose of more convenient discussion here.

Consultation of the table reveals that, first of all, the predictability on a normal group is not as high as that on the selected group used in the first experimentation. This was expected, and as before stated, was the primary reason for the selection of the first group. However, in connection with this, the second group, or the normal group, is the one which faculty members see in their classes. It is the one for whom the grades are assigned.

It is the group for whom the test must be able to predict--not a group selected on the basis of previous performance. In view of this, it is somewhat discouraging to note that the correlation of the total score with grades in only .384. It is necessary, then, if possible, to find some method of improving this correlation.

The correlation figures given in Table V, on the following page, support the contention of the writer. That is, by employing a "caution factor" in conjunction with already known scores, the correlations may be raised appreciably. For instance, correlation of grade point average for one year with total score shows a relationship of .384, but when the caution factor of the total score is multiplied by the total score, the resultant figures correlate .472. Improvement is .088, which would seem to be enough to warrant the extra computation involved, at least on questionable cases.

This same improvement, though not in such a large degree, is evident throughout. Wherever, in whatever manner, to whatever score the "caution factor" is employed, an improvement is evidenced.

The last two scores on this table are an added feature over those used on the first experimental group. The Group Score is defined as the number of deciles

TABLE V

CORRELATIONS BETWEEN 1945 A.C.E. SCORES (INCLUDING CAUTION FACTOR AND GROUP SCORES) AND GRADE POINT* AVERAGE FOR THE ACADEMIC YEAR OF 1946-1947 (N EQUALS 100)

SCORE	CORRELATION
Q-Score	.310 ± .06*
L-Score	.350 ± .06
Total Score	.384 ± .06
Caution Factor Quantitative	.293 ± .06
Caution Factor Total	.453 ± .05
Caution Factor Q-Score (Q x CFQ)	.350 ± .06
Caution Factor Total Score (T x CFT)	.472 ± .05
Caution Factor Q-Score plus L-Score	.445 ± .05
2 Caution Factor Q-Score plus L-Score	.409 ± .06
3 Caution Factor Q-Score plus L-Score	.400 ± .06
Group	.458 ± .05
Group Times Error Decile	.463 ± .05

*Probable Errors used throughout.

(speed, total score, and error) where the individual is rated at five or above, plus one. For instance, if a student is in the tenth decile on his total score, in the fourth decile for errors, and in the fifth decile for speed, two deciles are five or above, and he is in group three. This gives four possible groups, numbered from one to four, inclusive.

Considering that there are only four groups presented by using this procedure, and that grouping of data into so few categories limits the variability to a great extent, it is agreeably surprising to note the correlation of .458 which is present between scholastic average and Group. This also presents a method which involves only scoring for errors, adding to get speed, and making decile distributions of these factors, since the grouping is a mere matter of inspection. In other words, here would seem to be a method which is relatively simple, but which is more valuable than the usual total score decile method of selection of students.

In order to determine how greatly this group score was being influenced by the error decile, the last combination attempted of this type was that of Group multiplied by the error decile. This would seem to give an extremely high weighting to the "caution factor", since the error decile is assigned so that those who make the

fewest errors receive the highest decile. Furthermore, the error decile is one-fourth of the Group score. However, in spite of the emphasis upon caution inherent in this score, the correlation was .463, the second highest correlation found in this table.

It is evident, then, from inspection of Table V, that any method which was attempted incorporating the "caution factor" is higher in predictive value than the original method of scoring, at least on the 1945 edition of the A.C.E. test.

Results and Interpretations For the Second Experimental Group, Using the 1946 A.C.E. Test. Results with this edition of the Psychological Examination were not as encouraging as those with the 1945 edition. This would seem to be due to the fact that the L-Score correlates to a very low degree with grades. (Refer to Table VI, next page). This factor causes low correlations to be present throughout, whenever this score is a part of the combination. However, there are indications that this lack of correlation was due to a few individuals, and these individuals were subjected to further study. Information located will be found in the last section of this chapter.

Noteworthy findings with this group of one hundred engineering students are the following: (1) the Q-Score

TABLE VI

CORRELATIONS BETWEEN 1946 A.C.E. SCORES (INCLUDING CAUTION FACTOR AND GROUP SCORES) AND GRADE POINT AVERAGE FOR THE ACADEMIC YEAR OF 1946-1947 (N EQUALS 100)

SCORE	CORRELATION
Q-Score	.345 ± .06*
L-Score	.130 ± .07
Total Score	.262 ± .06
Caution Factor Quantitative	.293 ± .06
Caution Factor Total	.001 ± .07
Caution Factor Q-Score (Q x CFQ)	.395 ± .06
Caution Factor Total Score (T x CFT)	.313 ± .06
Caution Factor Q-Score plus L-Score	.295 ± .06
2 Caution Factor Q-Score plus L-Score	.344 ± .06
3 Caution Factor Q-Score plus L-Score	.365 ± .06
Group	.282 ± .06
Group Times Error Decile	.241 ± .06

*Probable Errors used throughout.

seems to be the most consistent measuring stick thus far; (2) the Caution Factor Quantitative ratio was found to correlate to exactly the same degree with grades on this edition as it did on the 1945 edition-- .293; (3) correlations were all improved in regard to any type of score when the "caution factor" was incorporated, excepting in the Group scores. This latter exception is probably due to the almost exactly chance correlation of the Caution Factor Total ratio, placing all the load of improvement on the quantitative section. Apparently, in this case, the error factor was overweighted.

Correlations were again increased, however, by a fairly substantial margin. The original Q-Score correlation was .345, but the Caution Factor Q-Score correlated with grades to the extent of .395, an increase of .050. The same increase is evident from the Total score at .262 to the Caution Factor Total score at .313, an improvement of .051.

It may be concluded, therefore, that even though the instrument was rather low in predictive value for this group at the outset, correlations were improved once more by application of a "caution factor".

TABLE VII

CORRELATIONS BETWEEN 1945 AND 1946 A.C.E. SCORES COMBINED (INCLUDING CAUTION FACTOR AND GROUP SCORES) AND GRADE POINT AVERAGE FOR THE ACADEMIC YEAR OF 1946-47 (N EQUALS 200)

SCORE	CORRELATION
Q-Score	.327 ± .04*
L-Score	.236 ± .04
Total Score	.323 ± .04
Caution Factor Quantitative	.293 ± .04
Caution Factor Total	.233 ± .04
Caution Factor Q-Score	.372 ± .04
Caution Factor Total Score	.392 ± .04
Caution Factor Q-Score plus L-Score	.370 ± .04
2 Caution Factor Q-Score plus L-Score	.382 ± .04
3 Caution Factor Q-Score plus L-Score	.382 ± .04
Group	.365 ± .04
Group Times Error Decile	.352 ± .04

*Probable Errors used throughout.

Results and Interpretations For the Second Experimental Group When 1945 and 1946 Subjects Were Combined. Correlations found for the combination of the two groups are organized in Table VII, on the preceding page. Scores follow the same order as in the previous two tables.

Combination of the two student groups produces greater consistency of correlation, and therefore reduces the probable error of the correlation, there being a total of two hundred students altogether. Probable errors were carried to two places in all cases throughout the study, this being deemed sufficient to show the reliability of the coefficient.

The same potentialities of the test which were evident in the last two tables seem also to be brought out in Table VII. First of all, the Q-Score gives a slightly higher correlation than either the L-Score or the Total Score. Secondly, the "caution factor" increases the correlations no matter how applied. Again, the Caution Factor Q-Score and the Caution Factor Total Score show improvements over the Q-Score and the Total Score, the increases in this case being, respectively, .045 and .069. Also, the simpler procedure of the Group score shows an increase of .042 over the Total score. Consistency of this sort would indicate that, even though

predictability is not raised to a statistically significant degree (since these are just over one probable error), something is being measured here which is a valuable and contributing factor to student scholastic adjustment. The increase of the Caution Factor Total Score over the Total Score would show 87 chances in 100 of being statistically significant. That is, there are 87 chances in 100 that this difference is actually greater than zero. Therefore, though this improvement is not established completely, the chances are definitely in its favor.

Correlations for Scores Involving Less Computation. At this point in the study, after it had been ascertained that the use of a "caution factor" was valuable in predicting college grades, it became evident that the manipulation of figures in order to obtain the various scores was becoming too great. The only score used thus far which involved little manipulation of other scores was the Group score. Since such overly ambitious mathematical combinations might not even be attempted in guidance work, due to the work involved, this detracts from the usability of the concept of "caution factor". Therefore several other methods were attempted which were simpler in construction, in the hope that one would be found which approached the more complicated methods.

TABLE VIII

CORRELATIONS BETWEEN A.C.E. RIGHT MINUS WRONG SCORES AND GRADE POINT AVERAGE FOR THE ACADEMIC YEAR OF 1946-1947

SCORE--NO.	CORRELATION FOR A.C.E. OF YEAR 1945	1946	1945-46
Number of Subjects	100	100	200
Quantitative Right Minus Wrong	.337 ± .06*	.386 ± .06	.361 ± .04
Total Right Minus Wrong	.488 ± .05	.313 ± .06	.400 ± .04

TABLE IX

CORRELATIONS BETWEEN A.C.E. DECILE SCORES AND GRADE POINT AVERAGE FOR THE ACADEMIC YEAR OF 1946-1947

SCORE--NO.	CORRELATION FOR A.C.E. OF YEAR 1945	1946	1945-46
Number of Subjects	100	100	200
Total Score Decile plus Error Decile	.496 ± .05*	.300 ± .06	.398 ± .04
Error Decile	.306 ± .06	.174 ± .06	.235 ± .04

*Probable Errors used throughout.

Correlations for these other scores are found in Tables VIII and IX, on the preceding page. Table VIII shows that Total Right Minus Wrong Score correlates .400 with college grades. Adding the Total Score decile to the Error decile produced a correlation of .398 with college scholastic average. These correlations are slightly higher than any others found for the complete group of two hundred students in engineering. Furthermore, the latter combination of the Total Score decile and the Error Decile gives a quick inspection method of evaluation which has 86 chances in 100 of being significantly better for predictive purposes than the Total score.

It should be noted that the Error decile by itself (Table IX) correlates with grade point average to the degree of .235. This one figure, then, does not tell the whole story, but should be used in combination with the Total score decile. Apparently, however, the two factors are about evenly weighted insofar as predictive value is concerned.

It may be concluded, then, that these more easily arrived at scores are as high in predictive value as the complicated scores. This would then appear to give a method which can be used in a practical counseling situation to improve the value of the testing instrument.

Further Analysis of Data. In order to illustrate the ability of the various measures used in this study to separate students according to grade point average, other tables were constructed. Comparisons are first shown between Total Score decile and the number of students in various grade point average groupings. Similar comparisons are made for Error decile and Speed decile. In these tables the high deciles are arranged at the top, and the progression is down toward the lower deciles. That is, for Total Score, the deciles indicating the highest score are 10 and 9, and these are at the top of the table. For the Error deciles the same is true, excepting that in this case these top deciles mean that fewer responses were scored as wrong for the students grouped into the decile. The meaning of the deciles is, however, the same, the most desirable decile is 10, and the least desirable 1.

The first two tables in this section, Tables X and XI, show the comparison of students in grade point groupings with the Total Score decile. Table X shows the segregation above and below a 2.5 grade point average, while Table XI divides the average into three groups.

These two tables show that the Total Score decile is valuable in this type of differentiation, though Table XI makes it evident that not many engineering students are

TABLE X

COMPARISON OF NUMBER OF STUDENTS ABOVE AND BELOW A 2.5 GRADE POINT AVERAGE WITH TOTAL SCORE DECILE ON THE A.C.E. TEST
(N equals 200)

Total Score Decile	Below 2.5	Above 2.5
9-10	30	42
3-8	71	49
1-2	8	--

TABLE XI

COMPARISON OF NUMBER OF STUDENTS IN THREE GRADE POINT GROUPS WITH TOTAL SCORE DECILE ON THE A.C.E. TEST

(N equals 200)

Total Score Decile	0.9-2.1	2.2-2.8	2.9-4.0
9-10	13	36	25
3-8	39	58	21
1-2	5	3	--

present in the lowest two deciles, eight being shown out of the total of two hundred. This would indicate that perhaps it might be of value to construct a separate decile distribution for engineers alone on this score.

In Table XII the same type of separation was obtained for the Error decile. This compares favorably with the Total Score decile, apparently, for purposes of segregation. If anything, this score grouping would seem to give a better balanced distribution than the other.

Table XIII carries the grouping to a higher point, the first three and the last three deciles being grouped, instead of merely two. In spite of this higher grouping, the pattern of segregation is still apparent. This shows that the Error decile holds its predictive value throughout the distribution, and not merely at the extremes.

These last two tables make it apparent that, in spite of the fact that an Error decile has never been used previously on the A.C.E. test, the "caution factor" concept is a workable idea when expressed in this form.

The next comparison shown concerns the Total Attempted score. This is expressed as a decile, called the Speed decile. Table XIV, gives the number of students above and below a 2.5 average in the decile groupings, as

TABLE XII

COMPARISON OF NUMBER OF STUDENTS IN THREE GRADE POINT GROUPS WITH ERROR DECILE ON THE A.C.E. TEST
(N equals 200)

Error Decile	0.9-2.1	2.2-2.8	2.9-4.0
9-10	7	22	8
3-8	36	59	33
1-2	15	18	2

TABLE XIII

COMPARISON OF NUMBER OF STUDENTS IN THREE GRADE POINT GROUPS WITH ERROR DECILE ON THE A.C.E. TEST
(N equals 200)

Error Decile	0.9-2.1	2.2-2.8	2.9-4.0
8-10	9	42	15
4-7	25	34	21
1-3	24	23	7

separated on the basis of the Speed decile. This decile showed little value in separation of the students. This should have been expected, since, if the Error decile and the Total Score decile are both valuable as predictive tools, but scored in opposite directions, a combination of the two would cancel the worth of both. This would seem to be exactly the effect noted in Table XIV.

TABLE XIV

COMPARISON OF NUMBER OF STUDENTS ABOVE AND BELOW A 2.5 GRADE POINT AVERAGE WITH SPEED DECILE ON THE A.C.E. TEST (N equals 200)

Speed Decile	Above 2.5	Below 2.5
9-10	18	22
3-8	57	65
1-2	16	24

The four tables following are concerned with the Group scores, since this seemed to be one of the most consistent and simple measures employed in the study. The Group score, it will be remembered, is based on the individuals standing in Total Score, Error and Speed deciles. The number of these deciles for which the individual is rated at five or above, plus one, is the Group score.

TABLE XV

COMPARISON OF NUMBER OF STUDENTS ABOVE AND BELOW A 2.5 GRADE POINT AVERAGE WITH GROUP ON THE A.C.E.
(N equals 200)

Group	Below 2.5	Above 2.5
4	25	37
3	56	44
2	15	9
1	13	1

TABLE XVI

PERCENT OF STUDENTS IN EACH GROUP RECEIVING ABOVE OR BELOW A 2.5 GRADE POINT AVERAGE FOR THE YEAR 1946-47
(N equals 200)

Group	Percent Below 2.5 Av.	Percent Above 2.5 Av.
4	40	60
3	56	44
2	62	28
1	93	7

Table XV shows the comparative number of students above and below a 2.5 average in relation to the Group wherein they were classified in the study. The results within this table are very consistent. This consistency is shown more graphically by Table XVI, where percentages are given within each Group for those above and below a 2.5. Since the highest group possible is 4, the tables are started with this group.

Tables XVII and XVIII give a comparison of the number of students in three and five grade point groupings with the four Group scores. This latter breakdown is quite complete for this type of comparison, since grade point averages are grouped into five separate categories. In spite of this, however, the relation between Group score and grade point average follows through to a great extent.

A weakness is also evident here, however. For Group 3 there are, in the first place, too many of the total number, as shown by Table XVIII. In the second place, this particular Group does not predict grade point average performance to as great an extent as it should. In view of this, it is possible that the Speed decile should not have been included in the formation of the Group. This is probable because of the lack of relationship between the

TABLE XVII

COMPARISON OF NUMBER OF STUDENTS IN THREE GRADE POINT GROUPINGS WITH GROUP ON THE A.C.E. TEST
(N equals 200)

Group	0.9-2.1	2.2-2.8	2.9-4.0
4	8	28	26
3	31	52	17
2	7	14	3
1	10	4	--

TABLE XVIII

COMPARISON OF STUDENTS IN FIVE GRADE POINT GROUPINGS WITH GROUP ON THE A.C.E. TEST
(N equals 200)

Group	0.9-1.9	2.0-2.2	2.3-2.6	2.7-3.0	3.0-4.0
4	2	9	17	17	17
3	17	36	24	19	14
2	3	7	11	--	3
1	5	6	3	--	--

Speed decile and grades (Table XIV). By eliminating this, however, we would have only Groups which ranged from 1 to 3, inclusive, being based on the Total score decile and the Error decile. This would seem to make the groups too few in number for any real use in segregation of students.

In constructing the Group score from these two deciles, it might be possible to arrange a total of six Groups. This could be done by arbitrarily dividing at 3, 7, and 10, instead of the previous arrangement of the break at 5 and above. This would give six groups, a larger number than has been used here, and groups which did not incorporate the Speed decile. Groups for this study were postulated on correlation, the Speed decile being used because of correlations which were higher than the "caution factor" correlations on the first experimental group. This relationship did hold true for a normal group, however. Therefore, an arrangement of six Groups, using only the Total Score and Error deciles would seem to be a much better possibility from the data found here. This cannot, however, be supported by any specific correlation figures, since this combination was considered feasible at too late a date for information to be included in this study.

At any rate, Table XVIII shows that the Group scores, even as they now are, can be used as a segregative device

with a high degree of validity, for Groups 1, 2, and 4. In Group 1, nearly all students received grades that averaged below a 2.2. This means that anyone in this Group will probably have difficulty with the engineering curriculum. In Group 2, chances of success are much better. In Group 4, nearly all students received above a 2.2 grade point average.

These tables of comparison are included merely to add emphasis to the correlation figures discovered. They show that some of the measures used here are quite consistent in their ability to segregate students in engineering according to probable grade point average. They also would seem to add to the usability of the A.C.E. test for counseling purposes, since they offer other methods of improving validity of the examination.

Follow-Up of Deviates. The following pages are devoted to brief case studies regarding the educational fates of twelve men in the second experimental group. These men were some of those who, in the 1946 test group, deviated to such an extent from the usual pattern as to lower the correlations to a considerable degree. The case studies are not complete as it was only possible to get information relative to their education, and very little else. However, they are included here since they seem to throw further light on the investigation of the usability

of the A.C.E. test and the "caution factor" for predictive purposes. Case number, as used here, is the number of this man, alphabetically, in the 1946 group of 100 students. This same number is used in Appendix B for his corresponding numerical data. The students were chosen to represent as many various types of deviation as possible.

Case No. 5. This student graduated from a large high school, his high school decile being 7. On the A.C.E. test, Total score decile was 10, Error decile 8, and Speed decile 10. His English test decile was 9. In high school he had taken the Otis test in 1941, his I.Q. being shown as 126. He received a Bausch and Lomb honorary science award his senior year in high school. However, his grade point average for the first year at Oregon State College was only a 2.16.

Other measuring devices, therefore, tend to agree with the A.C.E. results. It would appear from all indications that this student should be able to do good work, or at least better than he did the first year at Oregon State.

This veteran had been a Radio Technician in the Navy. His father is an engineer . . These two facts probably explain his choice of the engineering curriculum upon entrance into college.

He completed a second year in engineering at Oregon State, with grade points of 1.65, 2.21, and 2.23 for the three terms of that year in order. During this time he had received one grade of F in Differential and Integral Calculua, which he repeated with a C. He had received D's in Elementary Analysis (Math 102) and in Engineering Physics (Ph 102). In contrast, he received A's in American National Government, Literature Survey, and Economic Geography. These courses seem to group themselves as quantitative and linguistic courses, showing him to be greatly superior in the second type.

At this point he transferred to the School of Business and Technology, with a cumulative grade point average of 2.10, and received a 2.71 his first term and a 2.35 his second term.

It would seem then, in summary, as though this student was in a curriculum where his interest was at least lagging to some extent. A guess might be hazarded that he did not, even with the change he made, go far enough over to the linguistic side. This is merely a possibility on the basis of the grades given above. Still, in order to salvage at least some of his college time, a further change was probably not possible. But, at least, he would seem to be out of danger now, insofar as grade point average is concerned. This means that the

A.C.E. test was a good predictive instrument in his case, in all probability, but that he was taking a course to which he was not specifically adapted.

Case No. 24. In this instance the student had taken his high school work in a combination high school and junior college, but had transferred to Oregon State College after service, (as a freshman). His high school decile ranking was a 9, his English test decile 5, and his A.C.E. Total score decile 10, with an Error decile of 3. Considering this combination it was suspected that he might drop to some extent in grade point average, because of his lack of "caution".

He continued at Oregon State in Electrical Engineering. His first year's grade point averages were 3.53, 3.43, and 3.27; second year averages were 3.22, 2.47, and 2.00; and in his third year two terms had been completed with 3.13 and 3.20 averages. His cumulative grade point average, then, had dropped from a 3.41 to a 2.90, mainly because of his second year's work.

Since all other scores point to above average work, it is at least plausible to explain the 2.47 and 2.00 terms as due to the lack of activity of the "caution factor". Of course, there are many other possible explanations, this being only one of them, any of which could have caused this lowering of the scholastic average.

Many other cases with this particular combination would be necessary before any conclusion could be reached.

Case No. 26. This student showed a ranking of 9 in Total score decile, and 9 in Error decile. This should, it would seem, indicate good possibilities as a student. However, his first year he received 2.44, 1.86, and 2.20 grade point averages, the first of these being in the Lower Division curricula, giving him a 2.16 average for the term.

His English test decile again indicated good potentialities, being an 8. High school decile, however, was another story, being a 4. This, too, was the reason for the first term in Lower Division, as it was necessary for him to make up a deficiency in mathematics before he could enroll in the School of Engineering. He was a Private First Class in a combat engineers battalion while in the service.

Second year grade points were: 2.14, 2.35, and 3.07. This latter score came after a transfer to Industrial Administration from his original curriculum of Mechanical Engineering. His third year shows a continuation of the good work, being a 3.17 and 2.71 for the two terms available.

This would make it likely that the student was well

adapted to the school of Engineering, as the tests indicated, but had chosen the wrong place of emphasis within that field.

Case No. 33. This student received only a 1.73 grade point average for his freshman year in engineering, in spite of having deciles of 10, 9, and 8, respectively, from the A.C.E., the English test, and high school. His Error decile, however, was only 4. He had received an F in Math 103, and D's in Engineering Problems, Engineering Drawing, Engineering Physics, and English Composition, the latter possibly because of spending extra time on the other studies.

At this point he had an advisement from the College Counseling and Testing Bureau, and it was suggested that he transfer to Business and Technology. He did this, and has received no grades below a C since the transfer. His second year grade points were: 2.62, 3.38, and 2.81; and his third year 3.25 and 2.50. He is apparently realizing his potentialities in this curriculum, whereas he did not at all realize them in engineering. There is a possibility, then, that the "caution factor" might perhaps be more applicable to an engineering curriculum than a business curriculum.

Case No. 42. This student completed only four terms of school, with grade points of 2.31, 1.79, 0.93,

and 1.40, the last term being in Business and Technology. He was suspended with a 1.62 cumulative grade point average. However, on the A.C.E. he was in the 10th decile, and his Error decile was also a 10. English test decile was 6 and high school decile 6.

He stated at entrance that his interest was in Radio Engineering, and when he transferred three terms later, he stated that he ". . . thought I was interested in engineering but found I wasn't".

Since he did not return to college after the suspension, no further information could be located. However, it would seem probable that he was looking for a more "practical" and not as "theoretical" a course as he found here.

Case No. 46. A.C.E. decile was a 10, and Error decile 9. High school decile was 6 and English decile 9. On the basis of the first two it was expected that this student would improve from his 2.31 average for the first year. However, the opposite is what actually happened. His second year grade point averages were: 1.62, 1.88, and 1.92; and his third year 1.93 and 1.57. He apparently has remained in school this long due to an oversight. On the basis of the test results here there is no explanation for these scholastic averages.

Case No. 53. This veteran received an 8th decile on the A.C.E. Total score, and 8th decile on Errors. This would seem to mean that grades should be near a 3.00, whereas they were actually a 2.21 for his freshman year. However, upon consulting the records, it was found that his high school decile was 4 and English test decile a 6, putting him into the average range here. Furthermore, after this first year he transferred to Education as a major field, stating he was not interested in engineering, and received 2.56, 2.50, and 2.23 for the next three terms. He then transferred to another college, where he could obtain the particular emphasis he desired.

Test results apparently give a good picture of his abilities when he finally finds a curriculum that is suitable.

Case No. 64. Total score decile was 10 and Error decile 2, but grade point average only 2.29 after three terms. He spent two more terms in engineering and received 2.56 and .179. At that point he decided that he was out of place, and transferred to the school of Science with a major in Geology. His cumulative grade point average at the time of transfer was 2.25. The next three terms show that he apparently was out of place, and is now in a curriculum in which his abilities are being put to work, for these terms show averages of 3.22, 3.18,

and 3.06.

Again, in this case, then, the test results gave a very good picture of his potentialities, the trouble being that he was in a curriculum which was not satisfying to him as an individual.

Case No. 69. This is apparently another case in which the advisement given at the College Counseling and Testing Bureau caused a decision which produced favorable results for the student.

Results shown on the usual tests given were the following: English decile 8; high school decile 9; A.C.E. Total decile 10, and Error decile 10. He shows extremely high potentialities, but had received average grades in his first year of engineering, ending with an average of 2.44.

After advisement, he transferred to Business and Technology, where his grade points were: 2.88, 3.20, 3.50, 2.69, and 3.28. Three out of the five terms he received above a B average, and in one term he was on the honor roll.

Test results, then, apparently hold good here also, the veteran merely not realizing his possibilities because of other factors, which were pointed out in the advisement, and a change made. Since the change, he has been doing rather well.

Case No. 77. Here is a case where A.C.E., Total, English, and high school deciles were all 10, but where the Error decile was only a 3, indication a lack of the "caution factor", but certainly not of intelligence or study method. His grades for the second and third years remain this high or higher. Therefore, it might be concluded that his lack of caution on the test is not carrying over into the academic situation. However, two facts were available in the records which might serve to indicate the operation of "rashness". First, his service time was spent as a pilot, not normally considered to be an occupation for an overly cautious person. Secondly, he had at one time received an "E" grade in one course because of failure to take the final examination. He petitioned to remove the grade, and received an "A". He stated that he missed the final due to a " . . . misunderstanding of the final examination schedule". This perhaps might indicate that he was not being too careful at all times, and point out one way in which the activity of the "caution factor" may influence course grades.

Case No. 84. This student received scores giving him a ranking in the 9th decile for Total score, and 10th decile for Errors, indicating that he should do good work in college. To add to this, his English test decile was a 10 and high school decile an 8. However, his grade

point average for the first year in engineering was only a 2.11, and his fourth term in that school produced an average of 1.94. He thereupon transferred to the school of Education, where his grades showed an immediate rise. He received 2.40, 2.38, 3.06, and 3.00 in the four terms completed here. He stated at the time of transfer that he took engineering because " . . . everyone advised that engineering was 'the course' to take". He also stated that he dropped it because of " . . . greater interest in more liberal fields".

This seems to state, again, the case for the test. Good abilities were not being used to the full, because of a lack of motivation and interest in the curriculum being taken. A transfer to the type of courses he wanted procuced a sudden rise in grades.

Case No. 100. This student dropped his college work after the first three terms. Good abilities are shown on all tests, as evidenced by 10th and 9th deciles on the A.C.E. for score and errors, and 10th decile on the English test. However, high school records showed only a 3rd decile, and this may perhaps be the explanation here. That is, the potentialities were there, but the person had never learned to use them to any advantage in school work. This is, of course, one factor which the results on the tests cannot indicate.

Summary of Case Studies. Individual cases cited here were chosen on the basis of their not fitting into the usual pattern. That is, usually they had received high Total deciles and Error deciles on the A.C.E., but were not achieving high grades; or possibly they were overachieving on the basis of these two deciles. However, upon closer analysis, it is evident that they do fit into the pattern for the most part. Taken in this way they add to, rather than subtract from, the results found by correlation. These men did not fit the pattern for their first year in school for various reasons. However, when they got straightened out, the scores on the A.C.E. test were just as indicative as for others.

In a few cases, the A.C.E. test did not give a complete enough analysis. This was, of course, expected. This merely points up the known fact that any single testing instrument cannot be sufficient in all cases.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Summary. This study is an attempt to determine the effect of "caution" upon performance of students in the School of Engineering at Oregon State College. "Caution", for the purposes of this study, is measured by a rescoring of the American Council on Education Psychological Examination. The term "caution factor" indicates the ratio of the number of correct responses to the number of responses attempted on either a portion of or the whole of the A.C.E. test. This ratio is obtained by dividing the number of correct responses by the total number of items attempted.

The need for a study of this type is apparent when one considers that engineering students are being segregated and counseled on the basis of a method of scoring the test which has been shown to have a rather low validity for this particular curriculum. Correlations found for engineering students in this particular study range from .262 to .384 between total score and grade point average. Other earlier studies show correlations between A.C.E. scores and grades for engineering students to be .42, .448, .48, and .555 at different institutions, using different editions of the test (39, 53, 29, 20).

The deviation between results of the present study and these results may be due to the recent influx of students into the engineering curriculum. However, some studies show very little relationship between courses necessary in engineering and the A.C.E.. One study, for instance, gives correlations of .03, .19 and .25 between A.C.E. score and grades in physical science (52). It is apparent, then, that differential validity tends to be somewhat lower than is desirable, even though general validity is found to compare very favorably with other tests of this type (7, 41). This means that further study is necessary in specific regions of academic work in order that the test may be used to the fullest advantage in these regions.

The region which, to the writer, seemed to present the greatest challenge and chance for improvement was the curriculum of engineering. Thus arose the hypothesis this study was designed to investigate: That, by incorporating a "caution factor" into the scoring of the A.C.E. test, the validity of this instrument can be increased for students of engineering.

Subjects to test this hypothesis were in two experimental groups, totaling three hundred, all of whom were freshmen students in engineering. The first one hundred were chosen on the basis of separation toward the

extremes in grade point average. The second two hundred were a normal, unselected group for the School of Engineering. Results are based on findings obtained from these two groups.

Results pertaining to the "caution factor" concept would seem to have implications in guidance of students in engineering. Though improvements noted in validity through the use of this new method of scoring on the A.C.E. test are not completely significant statistically, enough consistency and tendency toward significance is noted so that the method may be assumed to have value. Let us suppose, for example, that a student who is thinking of taking engineering in the coming academic year comes to us for advice. He has taken the A.C.E. test, and no other test. He wishes to know if it is possible for him to complete the curriculum in engineering.

First, let us examine his total score on the test. If this score is in the 9th or 10th decile, he would have approximately 82 chances out of 100, according to results found in this investigation, to achieve a grade point average of 2.20 or better. (This average is set as the lower limit here, because it will ordinarily indicate he is not in danger of "flunking out"). If the score is 3rd to 8th decile, his chances are 67 in 100 of receiving 2.20 or better. If his decile was 1 or 2, he has only 37

chances in 100, or in other words, the chances are against him, this being especially so if he is in the first decile. In this study, all those who were in the first decile for total score received below 2.20 grade point average.

Now we should rescore the test for errors, and find the error decile for this student. Taken singly, according to the results found in this study, chances are as follows: In the 9th and 10th decile, 81 in 100 of making a grade point average of 2.20 or more; in the 3rd to 8th deciles, inclusive, 72 out of 100. It may be seen from this that the error decile, taken by itself, is not quite as good a separating device as the total score. However, if these two scores are both considered as part of the picture, giving different facets of the student's potentialities, some gain in predictive value of this one test is evident. Furthermore, the second score may give some insight into the character of the student which could prove of value.

Statistically, the Total Right Minus Wrong Score, shows almost nine chances out of ten of being a better predictive measure than does the usual total score. The Caution Factor Q-Score and the Caution Factor Total Score are very nearly this much better than their corresponding scores without the caution factor. It would seem, therefore, that a counselor stands a better chance of being

correct in his diagnosis if the caution factor scores are used.

Conclusions Reached on the Basis of the Results. An analysis of the correlations found between various A.C.E. test scores and scholastic average makes it possible to reach the following conclusions.

1. The score found for the linguistic regions of the A.C.E. test does not usually correlate highly with scholastic averages attained in an engineering curriculum.

2. Of the present scores obtained on the A.C.E. test, the Q-Score would appear to be the most valid for students in engineering. The margin of difference between the Q-Score and the Total Score is, however, slight.

3. The A.C.E. test shows better predictive ability at the extremes of scholastic average than at the means for students in engineering. Such prediction may be increased by the application of "caution factor" to the Q-Score, L-Score, or Total Score.

4. In some individual cases interest, motivation, background, and other factors, may interfere with validity of the A.C.E. test. However, if these difficulties are alleviated, the student is usually able to perform at the level which the test indicated.

5. Scoring of the A.C.E. test for "caution factor", and combining this factor with Q-Score or Total Score, increases the value of the test as an instrument for prediction of college grades in the engineering curriculum.

6. Of the various scores used for the A.C.E. test, those which give the highest correlation with grades received by engineering students in this study are: The Total Right Minus Total Wrong score; the Group score; the Total Speed decile plus Error decile score; the Caution Factor Q-Score; and the Caution Factor Total Score. These scores all correlate to a somewhat higher degree with scholastic average than do the usual Q-, L-, and Total Scores.

Limitations of the Study. The limitations of the results found in this study can probably be summarized in the following categories:

1. Further statistical interpretation of the data available could possibly have brought to light more information about these groups of students. However, the simple correlation employed would seem to the writer to be sufficient for the proof of the hypothesis.

2. The question of items "skipped" would seem to

hold many possibilities. The items "skipped" may themselves indicate items not attempted. This may represent a different type of "caution factor". This possibility was not investigated here, the writer assuming that since the test was a speed test, methods as applied here would be more valuable.

3. The complex of factors inherent in the college academic situation is a limitation of this study. Any number of things may destroy the correlation between academic abilities and academic standings. The method of treatment of the data, however, was an attempt to eliminate this limitation from the comparison between usual scores and "caution factor" scores. That is, each group was its own control group, due to the fact that the same scholastic averages were used in both cases, but for the "caution factor" scores a new measured quantity was added.

Suggestions for Further Investigation. Possibilities for further study are evident in:

1. The analysis of items "skipped" in relation to "caution factor".

2. Group scores computed on a six point scale, using only Total decile and Error decile, might

perhaps give higher correlations than the Group score as computed here.

3. Further study of the relation of "caution factor" with other curricula besides engineering might prove to be extremely valuable.

4. The construction of a power test of the Q-Score type, incorporating caution as part of the scoring method.

5. A "job analysis" of the requirements for academic success in a given curriculum would reveal the types of aptitude and knowledge needed there. Then a test could be constructed which would be more specifically fitted to the job it is expected to perform.

BIBLIOGRAPHY

1. Accounting Service Department, Monroe Calculating Machine Co. inc. Monroe application to simple correlation. Orange, N.J. 1937. 8pp.

2. Barnes, Melvin W. Gains in the A.C.E. psychological examination during freshman-sophomore years. Sch. and Soc. 57:250-52; Feb. 27 1943.

3. Barnes, Melvin W. The relationship of the study of mathematics to Q-Scores on the A.C.E. psychological examination. Sch. Sci. and Math. 43:581-2m, June 1943.

4. Brown, William M. A study of the "caution factor" and its importance in intelligence test performance. Amer. J. Psychol. 35:388-86, 1925.

5. Bryan, Alice I. Grades, intelligence, and personality of art school freshmen. J. Ed. Psychol. 33:50-64, Jan. 1942.

6. Buros, Oscar K., editor, Ninteen forty mental measurements yearbook. Highland Park, N. J., Mental Measurements Yearbook, 1941, 674 pp.

7. Buros, Oscar K., editor, The third mental measurements yearbook. New Brunswick, N. J. Rutgers University Press. 1949. 1047 pp.

8. Butsch, R.L.C. Improving the prediction of academic success through differential weighting. J. Ed. Psychol. 30:401-20, Sept. 1939.

9. Canady, Herman G. A study of sex diffenences in intelligence test scores among 1306 negro college freshmen. J. of Negro Ed. 12:167-72, April 1943.

10. Crawford, Albert B., and Burnham, Paul S., Forecasting college achievement, Part 1, New Haven, Yale Univ. Press. 1946. 194pp.

11. Crawford, J. E. Better basis for guidance. Ind. Arts, and Voc. Ed. 34:145-7, April 1945.

12. Crooks, William R. A preliminary report on the measurement of certain unknown personality traits by means of the ratio of the number of omissions to the number of errors made on the Otis Self-Administering Test of Mental Ability. Unpublished paper, University of Connecticut. 1937.

13. Cureton, Edward E. Note on the validity of the American Council on Education psychological examination. J. Appl. Psychol. 23:306-7, April 1939.

14. Freeman, Frank S. Predicting academic survival. J. Ed. Res. 23:113-23, Feb. 1931.

15. Garrett, Henry E. Statistics in psychology and education. Second edition, New York, Longmans, Green and Co. 1946. 493 pp.

16. Gray, Willian S., editor. Tests and measurements in higher education. Chicago, University of Chicago Press. 1936. 237 pp.

17. Harrison, Mary R. Intelligence scores of prospective teachers in a liberal arts college. Sch. and Soc. 56:416-20, Oct. 31 1942.

18. Hart, B. and Spearman, C., General ability: its existence and nature. Brit. J. Psychol. 5: 51-79, 1912.

19. Heilman, J. D., and Congdon, Nora A. The influence of college activities on intelligence test scores. Ed. Adm. and Sup. 28:448-56, Sept. 1942.

20. Holcomb, G. W., and Laslett, H. R. A prognostic study of engineering aptitude. J. Appl. Psychol. 16:107-17, April 1932.

21. Hunsicker, Lilian. A comparison of scores on two college freshman intelligence tests. J. Ed. Res. 26:666-7, May 1933.

22. Hunter, Elwood C. Changes in scores of college students on the American Council psychological examination at yearly intervals during the college course. J. Ed. Res. 36:284-91, Dec. 1942.

23. Jones, George A.A., and Laslett, H.R. The prediction of scholastic success in college. J. Ed. Res. 29:266-71, December 1935.

24. Kent, R.A., and Schreurs, Esther. Predictive value of four specified factors for freshman english and mathematics. Sch. and Soc. 27:242-6, Feb. 25 1928.

25. Lanigan, Mary A. Effectiveness of the Otis, the A.C.E., and the Minnesota Speed of Reading tests for predicting success in college. J. Ed. Res. 41:289-96, Dec. 1947.

26. Laslett, Leota L. The educational fates of college entrants in the two lowest American Council on Education test deciles. Unpublished master's thesis, Oregon State College, June 1938.

27. Livesay, Thayne M. Intelligence of high school seniors in Hawaii. J. Ed. Psychol. 32:377-80, May 1941.

28. Mc Gehee, William, Freshman grades and the American Council psychological examination. Sch. and Soc. 47:222-4, Feb. 12 1938.

29. Mc Gehee, William, The prediction of differential achievement in a technological college. J. Appl. Psychol. 27:88-92, Feb. 1943.

30. MacPhail, Andrew H. Q and L Scores on the A.C.E. psychological examination. Sch. and Soc. 56: 248-51, Sept. 19 1942.

31. MacPhail, Andrew H. and Foster, Laurence S. New data for placement procedures. J. Chem. Ed. 18:235, May 1941.

32. Manual of Instructions, American Council on Education Psychological Examination, 1948 edition, New York, Cooperative Test Division, Educational Testing Service, 1948. p4.

33. Mosier, Charles I. Group factors in college curricula. J. Ed. Psychol. 26:513-22, Oct. 1935,

34. National Education Association of the United States. Review of educational research. Psychological tests and their uses. Washington D.C., Vol. 1, No. 1, Feb. 1944.

35. Peterson, B. H. Can you predict who'll be able to do college work? Sch. Executive. 61:26-7, Nov. 1941.

36. Quaid, T.D.D., A study in the prediction of college freshman marks. J. Exp. Ed. 6:350-75, March 1938.

37. Remmers, Hermann H., A diagnostic and remedial study of potentially and actually failing students at Purdue University, No. 12 Lafayette, Indiana, Purdue University press, May 1928.

38. Remmers, Hermann H., and Gage, Norbert L. Educational measurement and evaluation, New York, Harper and Brothers, 1943, 560 pp.

39. Rigg, Melvin G. The relation of college achievement tests to grades and to intelligence. J. Ed. Psychol. 30:397-400, Aug. 1939.

40. Seagoe, May V. Standardized tests in the pre-training selection of teachers. J. Ed. Res. 36:678-93, May 1943.

41. Segel, David. Prediction of Success in College. United States Office of Education bulletin No. 15, 1934.

42. Schmitz, Sylvester B. Predicting success in college: a study of various criteria. J. Ed. Psychol. 28:465-73, Sept. 1937.

43. Schneidler, Gwendolen G., and Berdie, Ralph F. Educational hierarchies and scholastic survival. F. Ed. Psychol. 33:199-208, March 1942.

44. Stalmaker, Elizabeth M. A four year study of the freshman class of 1935 at the West Virginia University. J. Ed. Res. 36:100-18, Oct. 1942.

45. Stalnaker, John M. American Council psychological examination for 1926 at Purdue University. Sch. and Soc. 27:86-8, Jan. 21 1928.

46. Super, Donald E. The A.C.E. Psychological Examination and special abilities. J. Psychol. 9:221-6, April 1940.

47. Thomson, William A. Note on retest results on the A.C.E. psychological examination for college students. J. Ed. Psychol. 31:229-33, March 1940

48. Thurstone, L. L. Psychological tests for college freshmen. Ed. Rec. 6:69-83, 282-94, April, Oct. 1925.

49. Thurstone, L. L. Attitudes can be measured. Amer. J. Soc. 33:529-534, 1928.

50. Traxler, Arthur E. Techniques of guidance. New York and London, Harper and Brothers. 1945. 394pp.

51. Traxler, Arthur E. Some comments on "the prediction of differential achievement in a technological college". J. Appl. Psychol. 27:176-9, April 1943.

52. Wagner, Mazie Earle, and Strabel, Eunice. Predicting success in the college physical sciences. Science Ed. 19:4-9, Feb. 1935.

53. Waits, J. Virgil. The differential predictive value of the psychological examination of the American Council on Education. J. Exp. Ed. 1:264-71, March 1933.

54. Webb, E. Character and intelligence: an attempt at an exact study of character. Brit. J. Psychol. Mon. Suppl. 1, No. 3, 1915.

APPENDIX A

RAW DATA FOR THE FIRST EXPERIMENTAL GROUP

APPENDIX A

EXPLANATION OF COLUMN HEADINGS USED

All scores designated apply to the 1947 edition of the American Council on Education Psychological Examination, as used in this study for the first experimental group.

N. . . . . Number given the student in engineering for the purposes of this experiment.

G.P.A. . . Grade point average for the fall term of the academic year 1947-48.

TSd. . . . Total Score decile.

Q-S. . . . Q-Score; the number correct on the quantitative portions of the test.

Q-A. . . . Q-Attempted; the total number attempted on the quantitative portions of the test.

L-S. . . . L-Score; the number correct on the linguistic portions of the test.

L-A. . . . L-Attempted; the total number attempted on the linguistic portions of the test.

T-S. . . . Total Score; the total of linguistic and quantitative scores.

T-A. . . . Total Attempted; the total number attempted on both quantitative and linguistic portions.

RAW DATA FOR THE FIRST EXPERIMENTAL GROUP

N	G.P.A.	TSd	Q-S	Q-A	L-S	L-A	T-S	T-A
1	1.50	8	51	58	65	74	116	132
2	1.87	10	60	64	79	98	139	162
3	3.50	10	67	67	92	97	159	164
4	1.50	3	39	43	49	102	88	145
5	1.80	10	55	56	86	91	141	147
6	3.35	2	53	56	15	25	68	81
7	1.88	6	47	53	60	62	107	115
8	3.53	10	61	67	91	107	152	174
9	1.41	5	46	48	52	66	98	114
10	1.93	4	37	43	58	70	95	113
11	1.76	3	32	36	56	70	88	106
12	3.07	10	55	59	82	94	137	153
13	0.71	7	50	52	63	69	113	121
14	4.00	7	47	52	64	76	111	128
15	1.75	10	65	70	91	103	156	173
16	1.50	9	49	59	80	86	129	145
17	1.76	5	49	52	52	66	101	118
18	3.12	9	56	64	72	89	128	153
19	3.86	10	64	65	80	86	144	151
20	1.00	5	39	48	64	78	103	126
21	1.94	9	54	63	71	83	125	146

N	G.P.A.	TSd	Q-S	Q-A	L-S	L-A	T-S	T-A
22	3.47	6	50	50	57	66	107	116
23	1.36	9	54	59	69	76	123	135
24	1.33	2	32	39	43	57	75	96
25	3.57	10	65	67	70	91	135	158
26	1.47	5	48	52	50	58	98	110
27	1.83	7	47	72	65	98	112	170
28	1.25	9	48	54	77	86	125	140
29	3.06	8	55	62	65	99	120	161
30	3.29	10	56	60	95	105	151	165
31	3.65	10	59	63	76	89	135	152
32	1.36	5	42	51	59	65	101	116
33	1.33	5	34	37	66	76	100	113
34	3.00	9	53	55	71	81	124	136
35	1.40	9	53	59	75	86	128	145
36	1.27	10	53	60	88	96	141	156
37	0.71	7	47	51	63	70	110	121
38	1.93	7	39	46	73	83	112	129
39	3.28	9	57	63	65	79	122	142
40	1.27	4	49	61	42	57	91	118
41	1.45	6	44	51	61	75	105	126
42	1.65	10	56	65	79	93	135	158
43	3.00	9	51	58	71	82	122	140

N	G.P.A.	TSd	Q-S	Q-A	L-S	L-A	T-S	T-A
44	3.00	10	57	68	79	96	136	164
45	3.05	8	54	62	66	79	120	141
46	3.47	10	64	72	76	118	140	190
47	1.28	9	55	57	74	77	129	134
48	3.83	10	46	52	86	103	134	155
49	3.23	10	71	80	99	120	170	200
50	3.12	9	45	47	79	94	124	141
51	1.07	4	39	53	56	74	95	127
52	3.80	7	53	53	58	61	111	114
53	1.24	5	40	50	58	77	98	127
54	3.12	9	67	74	63	80	130	154
55	3.69	10	73	75	96	110	169	185
56	1.36	3	39	43	50	53	89	96
57	1.39	10	53	59	70	80	123	139
58	1.64	8	43	47	76	90	119	137
59	1.31	9	60	70	67	86	127	156
60	3.31	10	57	60	88	93	145	153
61	0.42	3	49	54	37	53	86	107
62	3.40	6	46	52	60	68	106	120
63	1.14	6	38	46	67	78	105	124
64	1.71	8	48	50	68	88	116	138
65	1.73	7	39	48	73	92	112	140

N	G.P.A.	TSd	Q-S	Q-A	L-S	L-A	T-S	T-A
66	4.00	10	57	65	101	114	158	179
67	1.65	9	50	54	75	86	125	140
68	1.18	4	42	56	55	84	97	140
69	1.73	7	45	59	64	80	109	139
70	3.20	10	52	54	90	94	142	148
71	1.59	6	48	63	58	79	106	142
72	3.53	7	39	44	74	76	113	120
73	1.12	7	45	52	64	78	109	130
74	3.00	10	50	55	91	95	141	150
75	3.23	9	55	60	74	80	129	140
76	3.00	9	58	65	68	88	126	153
77	3.50	10	67	68	78	87	145	155
78	1.81	5	39	40	64	69	103	109
79	3.47	6	41	45	63	64	104	109
80	1.40	10	58	63	103	109	161	171
81	3.78	8	52	61	66	79	118	140
82	3.00	10	65	68	71	89	136	157
83	3.76	10	67	71	109	115	176	186
84	3.53	10	65	66	91	99	156	165
85	1.79	9	48	55	77	103	125	158
86	1.76	10	47	55	86	96	133	151
87	1.20	6	47	49	60	61	107	110

N	G.P.A.	TSd	Q-S	Q-A	L-S	L-A	T-S	T-A
88	3.13	10	63	67	81	90	144	157
89	3.00	7	50	56	60	92	110	148
90	1.00	9	53	56	76	89	129	145
91	3.06	7	49	51	65	73	114	124
92	1.76	8	42	48	74	93	116	141
93	1.93	1	35	42	38	45	73	87
94	1.12	6	43	52	62	75	105	127
95	1.71	7	51	55	62	65	113	120
96	3.25	6	41	47	66	71	107	118
97	3.33	10	58	60	73	80	131	140
98	1.73	5	49	73	49	89	98	162
99	1.29	6	43	49	61	78	104	127
100	3.00	10	66	73	107	115	173	188

APPENDIX B

RAW DATA FOR THE SECOND EXPERIMENTAL GROUP

APPENDIX B

EXPLANATION OF COLUMN HEADINGS USED IN APPENDIX B

All scores designated apply to the American Council on Education Psychological Examination. In Part A of this appendix, scores apply to the 1945 edition of this test, and in Part B scores apply to the 1946 edition.

N. . . . . Number given the student in engineering for the purposes of this experiment, in either Part A or Part B of this appendix.

G.P.A. . . Grade point average for the three terms of the academic year 1946-47.

Q-S. . . . Q-Score; the number correct on the quantitative portions of the test.

L-S. . . . L-Score; the number correct on the linguistic portions of the test.

T-S. . . . Total Score; the total of the linguistic and quantitative scores.

Q-A. . . . Quantitative Attempted; the total number attempted on the quantitative sub-tests.

L-A. . . . Linguistic Attempted; the total number attempted on the linguistic sub-tests.

T-A. . . . Total Attempted; the total number attempted on both quantitative and linguistic sub-tests.

TSd. . . . Total Score decile.

PART A

RAW DATA FOR THE SECOND EXPERIMENTAL GROUP

1945 A.C.E.

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
1	3.06	52	71	123	55	81	136	9
2	2.14	48	51	99	50	64	114	6
3	1.77	51	61	112	58	92	150	6
4	2.47	33	62	95	33	72	105	3
5	2.00	40	54	94	53	69	122	3
6	2.06	60	71	131	65	79	144	10
7	2.74	43	81	124	44	91	135	9
8	3.05	65	87	152	70	95	165	10
9	2.31	51	66	117	51	83	134	7
10	2.24	43	73	115	46	89	135	7
11	1.70	50	76	126	62	115	177	8
12	1.55	42	55	97	47	69	116	4
13	1.97	39	72	111	45	92	137	8
14	2.37	44	51	95	52	78	130	3
15	2.16	41	79	120	51	112	163	7
16	2.72	52	61	113	54	77	131	8
17	2.23	40	52	92	44	61	105	3
18	1.71	23	48	71	36	68	104	1

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
19	2.27	41	53	94	42	64	106	3
20	2.95	48	65	113	49	70	119	6
21	2.32	45	43	88	48	75	123	3
22	2.64	49	53	102	51	56	107	4
23	1.75	53	64	117	62	93	155	8
24	2.23	50	71	121	56	87	143	8
25	3.46	58	77	135	60	84	144	9
26	2.63	34	46	80	43	68	111	3
27	2.70	55	68	123	61	77	138	8
28	2.04	48	60	108	50	79	129	6
29	2.03	55	62	117	62	79	141	8
30	2.39	60	105	165	64	119	183	10
31	3.22	51	51	102	52	58	110	4
32	2.42	50	58	108	58	84	142	6
33	1.57	41	62	103	45	73	118	6
34	2.21	42	53	95	55	73	128	3
35	1.54	34	50	84	47	70	117	2
36	1.78	33	75	108	45	100	145	6
37	1.89	57	58	115	68	90	158	6
38	2.38	48	77	125	50	83	133	8
39	2.45	47	56	103	61	82	143	5
40	2.36	53	68	121	55	85	140	8

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
41	2.34	42	65	107	45	75	120	5
42	2.66	57	90	147	61	97	158	10
43	3.09	45	74	119	54	81	135	7
44	1.89	41	51	92	45	68	113	3
45	2.21	36	60	96	41	69	110	5
46	2.47	43	66	109	48	82	130	6
47	2.57	49	56	105	51	60	111	7
48	2.51	42	69	111	49	75	124	6
49	2.40	40	73	113	48	85	133	6
50	3.28	28	67	95	32	74	106	3
51	2.16	37	72	109	38	83	121	7
52	3.46	58	75	133	62	85	147	9
53	2.93	71	84	155	72	100	172	10
54	3.37	53	65	118	55	71	126	7
55	2.41	46	63	109	54	72	126	7
56	2.06	45	81	126	48	84	132	9
57	3.32	50	93	143	80	119	199	10
58	3.50	47	77	124	48	90	138	8
59	2.02	45	76	121	48	86	134	8
60	1.76	49	65	114	57	82	139	7
61	2.15	36	67	103	43	86	129	5
62	3.00	54	76	130	60	87	147	10

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
63	1.82	40	57	97	45	66	111	4
64	2.47	48	83	131	51	90	141	9
65	2.67	64	84	148	67	96	163	10
66	3.20	50	78	128	50	92	142	8
67	3.17	49	79	128	58	87	145	9
68	1.99	51	71	122	56	86	142	8
69	3.45	58	94	152	68	107	175	10
70	2.20	48	47	95	52	56	108	3
71	3.21	43	68	111	48	83	131	6
72	3.19	45	55	100	50	65	115	4
73	3.37	48	75	123	49	92	141	8
74	2.07	30	58	88	43	82	126	3
75	3.10	49	72	121	65	98	163	9
76	2.21	69	94	163	70	111	181	10
77	3.47	52	78	130	54	90	144	9
78	3.15	55	70	125	57	80	137	8
79	2.32	63	66	129	64	77	141	9
80	2.07	46	63	109	49	74	123	6
81	3.35	51	69	120	56	91	147	7
82	3.39	53	75	128	60	98	158	8
83	2.72	48	73	131	62	86	149	9
84	2.01	59	63	122	67	97	164	8

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
85	2.47	36	69	105	47	78	125	5
86	3.67	46	71	117	47	81	128	7
87	1.94	49	79	128	53	104	157	8
88	3.91	57	64	121	59	70	129	8
89	2.05	38	32	70	39	54	93	1
90	3.00	57	74	131	60	86	146	10
91	3.23	57	88	145	66	101	167	10
92	1.49	51	57	108	60	71	131	7
93	1.96	52	68	120	65	88	153	7
94	2.29	40	55	95	51	63	114	5
95	3.68	52	70	122	55	79	134	9
96	2.69	50	65	115	53	74	127	7
97	2.38	51	83	134	56	98	154	9
98	2.65	52	54	106	56	68	124	5
99	2.41	32	49	81	41	66	107	2
100	3.17	41	65	106	42	71	113	7

PART B

RAW DATA FOR THE SECOND EXPERIMENTAL GROUP

1946 A.C.E.

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
1	1.96	51	54	105	52	60	112	7
2	3.93	55	59	114	55	61	116	8
3	1.83	42	51	93	50	61	111	5
4	3.17	50	77	127	52	89	141	9
5	2.16	57	90	147	60	99	159	10
6	2.60	43	61	104	46	77	123	6
7	2.08	25	57	82	30	72	102	3
8	2.66	42	60	102	44	64	108	6
9	3.44	67	84	151	70	92	162	10
10	2.68	42	76	118	44	80	124	9
11	2.07	34	49	83	54	92	146	3
12	2.47	57	44	101	59	54	114	6
13	2.85	38	72	110	46	83	129	7
14	2.10	40	68	108	42	78	120	7
15	2.56	45	79	124	47	89	136	9
16	2.70	29	79	108	40	93	133	7
17	2.52	44	61	105	50	85	135	7
18	2.42	47	77	124	50	85	135	9

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
19	2.04	37	50	87	38	65	103	4
20	2.16	51	54	105	55	58	113	7
21	2.83	54	68	122	54	73	127	9
22	2.70	50	84	134	62	91	153	10
23	1.85	36	52	88	41	68	109	4
24	3.41	65	78	143	67	97	164	10
25	1.82	52	52	104	57	63	120	6
26	2.16	41	77	118	45	83	128	9
27	2.28	58	50	108	66	61	127	7
28	2.73	33	77	110	36	86	122	7
29	2.64	49	80	129	51	87	138	10
30	2.67	45	53	98	46	57	103	5
31	2.48	32	51	83	34	61	95	3
32	1.84	65	82	147	79	113	192	10
33	1.73	59	102	161	64	116	180	10
34	2.31	44	86	130	51	94	145	10
35	2.84	57	79	136	59	83	142	10
36	2.33	32	58	90	33	70	103	4
37	2.29	42	74	116	45	82	127	8
38	2.77	45	78	123	46	82	128	9
39	1.48	34	47	81	42	56	98	3
40	3.52	57	66	123	59	89	148	9

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
41	2.52	58	78	136	61	108	169	10
42	1.67	46	83	129	50	85	135	10
43	3.21	52	69	121	53	71	124	9
44	2.51	57	86	143	69	105	174	10
45	2.68	46	51	97	50	58	108	5
46	2.31	57	93	150	58	99	157	10
47	2.69	43	51	94	44	65	109	5
48	2.86	51	66	117	53	77	130	8
49	2.40	51	93	144	57	99	156	10
50	1.90	34	56	90	35	69	104	4
51	1.64	46	63	109	51	82	133	7
52	2.09	42	63	105	46	79	125	7
53	2.21	39	74	113	43	82	125	8
54	2.02	37	84	121	45	95	140	9
55	3.00	43	82	125	47	94	141	9
56	2.59	44	71	115	47	75	122	8
57	3.11	50	72	122	59	80	139	9
58	2.24	49	71	120	53	95	148	9
59	2.81	50	63	113	58	88	146	8
60	2.28	46	63	109	48	75	123	7
61	2.60	37	57	94	42	69	113	5
62	2.28	56	82	138	60	95	155	10

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
63	2.48	36	71	107	37	73	110	7
64	2.29	45	91	136	49	97	143	10
65	2.02	43	75	118	47	85	132	9
66	2.46	40	68	108	44	77	121	7
67	2.04	54	76	130	61	87	148	10
68	2.95	45	70	115	52	76	128	8
69	2.44	60	80	140	62	84	146	10
70	2.81	43	80	123	43	85	128	9
71	3.02	56	58	114	60	74	134	8
72	2.44	44	80	124	45	86	131	9
73	2.26	31	78	109	34	87	121	7
74	2.40	62	82	144	65	92	157	10
75	2.09	27	40	67	36	56	92	1
76	1.94	48	66	114	49	78	127	8
77	3.17	59	95	154	65	112	177	10
78	2.76	49	50	99	50	61	111	6
79	2.43	50	76	126	56	82	138	9
80	2.95	68	69	137	71	82	153	10
81	2.09	40	57	97	44	61	105	5
82	1.64	44	74	118	48	89	137	9
83	3.90	59	78	137	59	80	139	10
84	2.11	40	79	119	42	83	125	9

N	G.P.A.	Q-S	L-S	T-S	Q-A	L-A	T-A	TSd
85	2.14	17	58	75	20	62	82	2
86	2.81	40	89	129	47	101	148	10
87	1.76	47	62	109	51	73	124	7
88	2.25	51	48	99	56	57	113	6
89	2.06	42	72	114	50	92	142	8
90	2.73	54	74	128	59	91	150	9
91	2.15	48	73	121	50	86	136	9
92	2.33	51	84	135	51	99	150	10
93	2.37	32	47	79	33	55	88	2
94	2.24	20	54	74	34	77	111	2
95	3.16	57	74	131	59	83	142	10
96	2.95	44	58	102	48	65	113	6
97	2.21	62	85	147	68	103	171	10
98	2.37	42	59	101	46	70	116	6
99	2.25	46	57	103	55	82	137	6
100	1.85	47	93	140	52	97	149	10