no. 14"

PREDICTION OF SUCCESS IN COLLEGE

A Handbook for Administrators and Investigators Concerned with the Problems of College Admittance or Guidance of College Students

By

DAVID SEGEL

Educational Consultant and Specialist in Tests and Measurements
Office of Education

BULLETIN 1934, No. 15

STATES DEPARTMENT OF THE INTERIOR - - - *Harold L. Ickes*, Secretary
OF EDUCATION - - - - - - - - - - - - - *Bess Goodykoontz*, Acting Commissioner

STATES GOVERNMENT PRINTING OFFICE . . . . WASHINGTON : 1934

the Superintendent of Documents, Washington, D.C. . . . . Price 10 cents

CONTENTS

PART II. STATISTICAL METHODS AND RESULTS OF STUDIES OF COLLEGE PREDICTION

TABLES

FOREWORD

The great growth in college enrollments, the changes in the availability of positions for college graduates, and the development in the science of testing go hand-in-hand in making the problem of the prediction of college success of greater importance than ever before. This bulletin brings the attention of educators upon the work which has been done in this field and the techniques which may be used to advantage in solving these problems. In doing this it brings up to date the practices in this field since the publication of Colvin and MacPhail's Intelligence of Seniors in the High Schools of Massachusetts by the Office of Education in 1924.

The need for the advancement in the methods used for the guidance of entering college students is evident in our modern life. Education must more and more base its practices on the findings of statistical and experimental studies. This bulletin points the way for increasing the accuracy of our knowledge about the possibilities of achievement of our entering college freshmen.

Bess Goodykoontz,
Acting Commissioner.

PREDICTION OF SUCCESS IN COLLEGE

PART I : GENERAL PRINCIPLES OF THE PREDICTION OF COLLEGE SCHOLARSHIP

CHAPTER I : THE PROBLEM

FOR EFFECTIVE guidance of students into college work it is necessary to be able to make some articulation between the student's capacities and interests and the program of studies offered in the college or university. The place of the college in the educational scheme is an important item in discussing the problem of college admission and placement. Many educators consider the college or university to be but an extension of secondary education while others believe that the college or university is for a comparatively few individuals who are particularly able to profit by higher education, and still others contend that college is a place to prepare oneself for a vocation. Whichever one of these particular views is held by a college will influence materially the kind of regulations set up for entrance.

If college work is considered to be an extension of secondary education it may easily be argued that the pathway to further education (i. e. college education) should be made as easy as possible. If college education is of the same sort, essentially, as high-school education, it means that as many individuals as possible should have the benefit of it since universal secondary education is considered by most educators to be desirable. Guidance of pupils from high school into college in this case would have for its purpose the keeping of students in school for the required period of education. The view that college should be a place for a selected few to acquire culture—a place primarily for study and the cultivation of appreciation for the arts and sciences would require a different kind of guidance scheme. Much sifting of student material would be necessary. The view that college is a preparation directly for vocations would also require that serious consideration be given to the kind of student material.

The difficulty level of the curriculums of the different colleges will vary to some extent according to the purposes of the individual institutions. That there is a great variety of ability in student personnel in different colleges has been shown by several studies and particularly by the reports of the American Council on Education Psychological Examination.[1] The level of instruction in individual colleges must therefore differ materially. Colleges having students in the lower ability ranges naturally adapt, to some extent at least, their work to their student body, but as far as the marks obtained in the various schools show, no differences are discernible. These variations in the level of difficulty of the college curriculums indicate a certain lack of standardization of the meaning of the term "college." The problem of Who should go to college? is therefore made complex through the varying objectives and the curriculums of colleges and universities. The students who have been educated in different kinds of institutions, and have different capabilities, achievements, social backgrounds, personality traits, and interests must be considered. This interaction of student material and the college curriculums must be studied carefully.

RELATION TO PRINCIPLE THAT EDUCATION SHOULD BE EXTENDED TO ALL TO THE LIMITS OF THEIR ABILITY

Youth always demands another chance. Insofar as any system of college entrance we now have is concerned, there is some truth in youth's insistence that he may succeed if given a chance in any particular institution of learning. No subjective or objective method of estimating success has as yet been inaugurated which is perfect. It is true that the most valid method for finding out whether or not a student can achieve a certain scholastic standing or be successful in an occupation is for him to try it. If, for instance, a student wants to become an engineer, the most valid method for him to find out if he can become one is first to attempt the training necessary to become an engineer and then to go out and attempt the job. This trial and error method however is costly for both the individual and for society. For the individual it is a costly procedure because he may spend much time in trying out various lines of work. In failing at jobs for which he is actually unfitted, he comes to believe that he is unfitted for any job.

It is costly for society to spend time and effort in educating individuals to do tasks which they, at best, will perform on

[1] Reported in various issues of the *Educational Record.*

a very low level of efficiency. Both the individual and society have a stake in education. Consider, for instance, the students who intend to practice medicine. There is room and need in society for men and women who are able practitioners of medicine. There is hardship, however, on the one hand for the student who finally gets through and is ill-fitted to the profession, and on the other hand, for the public when served by poor practitioners.

Harold Clark and others[2] have studied occupational distribution and trends. They have found that there is no close relationship existing between the number of persons preparing or asking for positions in an occupation and the number of workers needed each year in the occupation. They recommend that some system of following the changes in occupational distribution be made and that information about occupational distribution be made available to college students and others. There is usually no definite relationship between the number preparing for an occupation and the number the occupation can absorb. This fact is important because it shows clearly that our present methods of vocational guidance do not have very clear working objectives. It shows that information regarding the relative availability of employment opportunities is needed so that those who are best fitted for different occupations, insofar as this can be determined, will be directed into them.

The individual student may say that he is willing to take all risks in placing himself in the right curriculum. But since it is society as a whole which is to be benefited by the proper distribution of effort in various fields some believe that the final disposition should be left in its hands. Assume that there are 200 individuals wishing to prepare for a certain vocation; suppose further that through the application of aptitude measures and other means and from a study of the needs of society, the school authorities, representing society, find that only 100 of the individuals should prepare themselves for the vocation. Some error will result in this selection, so that of the 100 selected not all will be of the best 100. Some of the rejected 100 will be those who would have made good in the preparation for the vocation and in the vocation

[2] References and material on studies in occupational distribution may be found in the February 1934 number of *Occupations—The Vocational Guidance Magazine*.

itself. From the standpoint of society this may be a defensible procedure even though detrimental to several individuals because the total result for society is a gain. For privately controlled institutions where the individual or the individual's family pays largely for his education, the question is not so important because the individual is taking a large share of the risk. Also privately controlled institutions may put into effect, within fairly wide limits, any kind of regulation they see fit regarding entrance. In publicly controlled institutions the problem is more acute because there a large part of the cost of education is borne by the State; also it is there that the greatest pressure is found for the admission of students. Indeed it has been so great that the courts have been resorted to at times in order to force admission. Jarvis (61) in reviewing the legal aspects of admission states:

> The admission of applicants in a public educational institution is one thing and the government and control of applicants after they are admitted is quite another thing. The first rests upon well-established rules either prescribed by law or sanctioned by usage, from which right of admission is to be determined.

Regulations regarding the expulsion of students for failing in too many courses seem to be a matter left to the individual institution and are not subject to review by outside authorities. Admission regulations, on the other hand, appear to be subject to appeal to outside authority. Changes in entrance requirements especially in State colleges and universities must be inaugurated with very careful attention to the right of, and with due regard to the reasoning of, various elements of society. Particular care must be taken in using tests which determine the general intellectual level apart from achievement in specific subject matter. This is because of the rejection of testimony regarding such mental levels in some of our courts.

ARTICULATION OF HIGH SCHOOL AND COLLEGE AS A FACTOR IN COLLEGE ADMISSION AND GUIDANCE

The complete articulation of high-school and college work is considered a desirable goal by most educators. By complete articulation we mean the understanding and appreciation of the objectives of both institutions by the personnel in both institutions and the making clear of the relation between those lines of activity which run through both institutions.

This does not mean to imply that either institution will not have activities peculiar to its own level. The presence of various terminal courses in high school would not interfere with this concept of articulation.

Kelly (73) analyzes the tendency toward fusion between the high school and college. He finds the junior college to be largely the medium through which this is taking place. This is a movement which has important guidance implications. Kelly in his conclusions states:

> The outlines above represent five ways in which the line which formerly marked the distinction between high school and college is being obliterated. First, the processes of education are continuous and permit of no sudden break between administrative units. Second, the student-centered curriculum as distinct from the subject-centered curriculum is valid alike for high school and college. Third, college teaching is a profession calling for definite training as does high-school teaching. Fourth, the scientific measurement movement is making rapid progress in the college and bids fair to become the largest influence in breaking down the artificial line between these two institutions. Fifth, and finally, the place of research as a basis for institutional planning is now gaining recognition in both high school and college.

TWO GENERAL ASPECTS OF THE PROBLEMS INVOLVED IN THE GUIDANCE OF ENTERING COLLEGE STUDENTS

In general, high-school and college education is a continuous line of endeavor to bring students to as high a standard of knowledge and cultural appreciation as possible, together with preparation for various vocations at different levels throughout the two segments of the school system with which we are here concerned. From the standpoint of guidance there seem to be two elements in the situation. One question, and one which will usually come first is: What level of schooling or what vocational preparation level will a given student probably attain? The other question is: In which specialized line of activity will the student do best? In the case of a few students the last question will come first. This will happen sometimes when a student shows considerable ability in a fairly specialized line very early in his scholastic career. If this specialized ability can be shown to be a marked variation from the general level of ability of the student, this significant variation is the important item. In such a case the further cultivation of the specialized ability is the important consideration.

As a usual procedure the first problem will be to determine the level of general education which a student will attain. Then later the important choice of specialized study or occupation can be made. This seems to be a logical way of looking at the problem, since general education is a preparation for an infinite number of life's activities. It is economical for the individual and society to thus view education because this plan offers the greatest flexibility for changing educational plans or readjusting of prevocational and vocational preparations. As pointed out before, upon failing in the preparation for one vocation the individual often makes no attempt to prepare himself for another. The tragedy of this lies in the realization that success might have been attained in some occupation on the same level into which a shift could have been accomplished without much loss of the preparation made for the other occupation. The earlier stages of preparation for occupations on the same level are often the same. By shifting from one type of preparation to another there would be only a small loss to the individual and also, consequently, to society.

Therefore what is needed in guidance in the way of objective measurements and ratings is (1) a method of determining which students may attain a certain level of accomplishment in school or occupation, and (2) a method of determining for any level whether or not a student will do better in one line of activity (school or occupational) than in another. At the end of the various stages of secondary and higher education, it should be decided, naturally, whether or not the student should attempt a higher level of education. However, it also becomes increasingly important to decide as higher and higher levels of education are reached which specialized line the student should pursue or what shifting in his educational program should take place. For instance, it is one prediction to estimate that a certain high-school student will probably terminate his education in junior college. It is another prediction to estimate that he will do better in commercial education than in any other line of junior-college endeavor. The former kind of prediction may be called "direct" and the other "differential."

RELATION OF THE SURVIVAL OF STUDENTS IN COLLEGE TO THE PROBLEM OF COLLEGE ENTRANCE

Freeman (45) raises the question of the use of tests for predicting the grade of work which will be attained in college. He believes it is survival in college that is the best criterion of success in college and that general mental tests have not contributed greatly toward this prediction. He states:

> All these investigations of the prediction of college scholarship have yielded interesting and valuable data; but none, it appears, have dealt with the problem which, to the writer, seems most fundamental; namely, how well do the tests of mental ability predict academic survival, regardless of marks? . . . If our results are typical, it must be said that the mental tests are inadequate as selective instruments at the college level, just as any other single criterion is inadequate. Yet they have a contribution to make; and on that basis their employment is justifiable.

However, from the evidence collected to date there seems to be no reason for changing our criterion of college success from scholarship in the college subjects to length of stay in college. In the first place, it can be argued that survival in college is a poor criterion of success. A student may attend school but in the process he may acquire habits of laziness rather than those of work. In the second place, there are a great number of factors which are influential in bringing about elimination from college. The influence of many of these are unpredictable as, for example, factors related to changes in economic standing and the health of the family to which the student belongs. In the third place, the very fact that the correlation between results on general mental ability tests and survival is less than between such tests and scholarship indicates the weakness of survival as a criterion of success since general intelligence may be assumed to be one important factor in college work. Edgerton and Toops (40) have studied this problem of survival in various departments of Ohio State University. They find that whereas the correlation of the Ohio psychological examination using percentiles with point-hour-ratio of marks average .45 for the different departments, the correlation between persistence in college and the Ohio test is only .19.

For these three reasons it seems desirable to measure success in college by the grade of achievement attained while in college rather than by persistence in college. There are problems arising from the fact that students drop out of college.

Therefore the study of factors incident thereto is important in itself. However until the factors related to survival are better understood it is believed that success in college should be based on the level of the scholarship attained.

METHODS WHICH HAVE BEEN USED IN DETERMINING COLLEGE ENTRANCE

The methods used for determining college entrance are numerous. Brammell (12) made a survey of existing practices of college admission. Methods used singly by a fairly large number of colleges and universities were found to be: (*a*) High-school diploma; (*b*) high-school transcript; (*c*) college board examinations; (*d*) examinations by the institutions; (*e*) other examinations (regents, State board, etc.); (*f*) high-school subject certificates; (*g*) rank in high-school class; (*h*) intelligence test; (*i*) principal's recommendations; (*j*) other recommendations; and (*k*) personal interviews. Admission requirements which have been used in conjunction with others by a fairly large number of colleges and universities were: (*a*) Transcript of high-school credits; (*b*) recommendation of principal; (*c*) personal interviews with applicants; (*d*) rank in high-school graduating class; (*e*) recommendation by others than the principal; (*f*) high-school diploma; (*g*) character rating; (*h*) presentation of high-school subject certificate; (*i*) examination devised and administered by the institution; (*j*) college entrance board examination; (*k*) intelligence test; (*l*) college aptitude test; and (*m*) other examination (regents State boards, etc.).

Among specific tests which have been used by higher institutions for college admission and guidance are: American Council on Education Psychological Examination, Thorndike's Intelligence Examination for High-School Graduates, the Scholastic Aptidude Test of the College Entrance Examination Board, and the Iowa Placement Tests. A large number of other tests including achievement tests have been used by colleges and those named represent some of those most widely used. The value of the different tests named and others for purposes of prediction may be ascertained to some extent by a study of the results of studies as given in part II, chapters II, III, and IV.

From a consideration of the various facts in connection with college admission, Brammell states certain problems which arise, three of which are quoted as follows:

(a) The multitude of admission methods and combinations of entrance criteria used by the institutions included in this study reveals in an impressive way the fact that there are no recognized standards in the field of admission requirements. The actual reliability of individual criteria, or various combinations of criteria, and of no criteria at all, needs to be determined. It needs to be demonstrated whether or not traditional entrance standards are worth the trouble they entail.

(b) At present the manner in which personnel work in the higher institutions is carried on and the character of information assembled concerning students are reflections largely of the personal opinions and interests of the directors in charge. The value of certain kinds of information needs to be established and a general personnel program, recognized as good and at the same time flexible, needs to be outlined.

(c) Plans for the improvement of articulation need to be judged in terms of improved student population and their success in higher institutions and not in terms of popular opinion.

The large number of different methods used by colleges and universities is recognized by Brammell to be a sign of weakness rather than strength. He thinks that if the different methods of admission were being experimented with it would be a healthy sign. However, there is very little evidence of planned experimentation.

The problem of admission has to be met by each individual institution according to the kind of student body it wishes to have and its objectives of education. This problem is shown in the great variety of admission requirements, the variety of courses, and the great differences in scores on both achievement tests and intelligence tests in different colleges. One observer (Kurani, 76) writes:

A foreign observer who visits different parts of the country . . . finds institutions refusing 5 out of every 6 applicants, while some are taxing their facilities to the utmost in order to meet the growing pressure. He finds colleges admitting students with a deficiency of 3 out of 15 units and others admitting only those who rank in the upper third of students satisfactorily completing high school; yet all lead to the same goal, the B.A. degree, and, in the eyes of the general public, graduates of both institutions are credited with equal attainment. Variation in quality of work is desirable if it represents a response to a special call arising from a genuine difference in levels of ability or achievement. If the lowering of standards is resorted to as a measure of publicity it cannot be defended, for it defeats the very purpose of the college or school.

It is only through measurement that anything approaching accurate prediction can be made. The use of tests is spreading rapidly throughout colleges and high schools. Especially is there growth in cooperative testing in the high-school area. A survey of the Office of Education[3] of such cooperative testing programs brought to light a very great many such programs. The acceptance of such programs by local school administrators lies partly in the felt need for uniform records which can be used for guidance. Only ~~a few~~ of these programs have been evaluated for their efficiency in predicting scholastic success. For most colleges the principal use of test results will be for advising the individual in regard to whether or not he should enter college and, if he enters, which course he should take.

It should be pointed out that in the testing of students entering college from different high schools there is no problem of accrediting high schools involved. It has been shown by Zook (135) that the testing of graduating high-school seniors does not involve the efficiency of teaching in the different high schools because the initial achievement or ability has not been taken into account.

The purpose of this bulletin is to describe the methods which are being used and suggest the future lines of development in the testing of entering college students. General suggestions regarding the kind and amount of testing are made and the detailed techniques for evaluating the tests used in any particular situation are discussed. Studies in the field are summarized. It is hoped that this bulletin will in some small measure serve as a guide or handbook for those concerned with the administration of the problems of college admission or the guidance of entering college students and also at the same time add something to the scientific study of education.

[3] National and State Cooperative High-School Testing Programs. U.S. Office of Education Bulletin, 1933, no. 9.

CHAPTER II : IMPROVEMENT OF THE PREDICTION OF COLLEGE SCHOLARSHIP THROUGH THE USE OF TESTS AND MEASUREMENTS

THREE TYPES OF TESTS WHICH HAVE BEEN USED IN THE PREDICTION OF COLLEGE SCHOLARSHIP

THE PREDICTION of college scholarship through the use of tests and measurements has been studied in many different colleges. This method of aiding the determination of which students shall enter college has been adopted by many schools. Tests are also used by a few institutions for classification and placement after entrance. A still greater number of schools are using test results for purposes of guidance without the adoption of specific administrative regulations.

There have been at least three general types of testing used at least experimentally for determining admission to college and for the guidance of entering college students. All three of these approaches have shown a definite predictive relationship between the test scores and later success in college. The three types of approach are (1) the testing of general aptitude for college work through general aptitude tests (i.e. intelligence tests); (2) the testing of achievement in specific subject matter studied in high school which is also found in the college curriculum; and (3) the testing of aptitude for the attainment of scholarship in an individual college subject in which the student is assumed to have had no previous training. These three types of testing for the purpose of increasing the accuracy of college guidance will be described by using portions of the tests themselves as illustrations. The tests used as illustrations are only samples of many different tests which could have been used. Results of studies with many different tests are given in part II.

General scholastic aptitude test.—The first type of test—the general scholastic aptitude test—will be illustrated by the 1932 edition of the American Council on Education Psychological Examination. There are five subtests in this examination, each one presumably testing a somewhat different trait or capacity for doing college work. The test items in these subtests have been arranged in order of difficulty. Therefore in making extracts to show the type of material used test items have been taken from various levels of difficulty. In the following examples the directions are used verbatim.

PARTS OF THE AMERICAN COUNCIL ON EDUCATION PSYCHOLOGICAL EXAMINATION—EDITION 1932

Completion

Directions: Think of the most appropriate word to complete each of the sentences. The number in parentheses indicates the number of letters in the most appropriate word. Write the word in the blank space at the right. Do not waste too much time on any one sentence, as you will receive credit for every word correctly given.

1. A (7) is a character dress such as is worn at fancy balls, etc.................................(7)..........
5. An (10) is one who prepares and sells drugs for medicinal use.................................(10)..........
8. A (7) is one who travels to some holy place as a devotee.................................(7)..........
12. By (8) is meant the atrocious killing of a large number of persons.................................(8)..........
17. A (8) is one to whom money is due.................(8)..........
23. (6) is rubbish, especially as results from destruction..(6)..........

Artificial Language

Directions: Read the vocabulary and rules of the artificial language given below. Do not try to memorize the vocabulary or forms, but consult them freely while translating the sentences on the bottom half of this page.

VOCABULARY		RULES
I	—av	1. Plurals are formed by adding "ay". Only nouns and pronouns have plurals.
me	—aver	Example: we—avay.
he	—aj	2. Past time is expressed by placing "lo" before the verb.
him	—ajer	Example: agreed—lofrega.
that	—rin	3. Future time is expressed by placing "sta" before the verb.
is	—nar	Example: will agree—stafrega.
agree	—frega	4. Nouns are formed by substituting the ending "il" for the "a" ending of the verb.
decide	—prista	Example: agreement—fregil.
express	—bans	5. Adjectives are formed by substituting the ending "uc" for the "a" ending of the verb.
impress	—guda	Example: agreeable—freguc.
		6. Adverbs are formed by substituting "ef" for the "a" ending of the verb.
		Example: agreeably—fregef.

Some of the sentences below are correctly translated and some are not. Underline all the errors in the translations as shown in the first four sentences (in the sample the first sentence only) which are marked. Work straight ahead and do not omit any sentences.

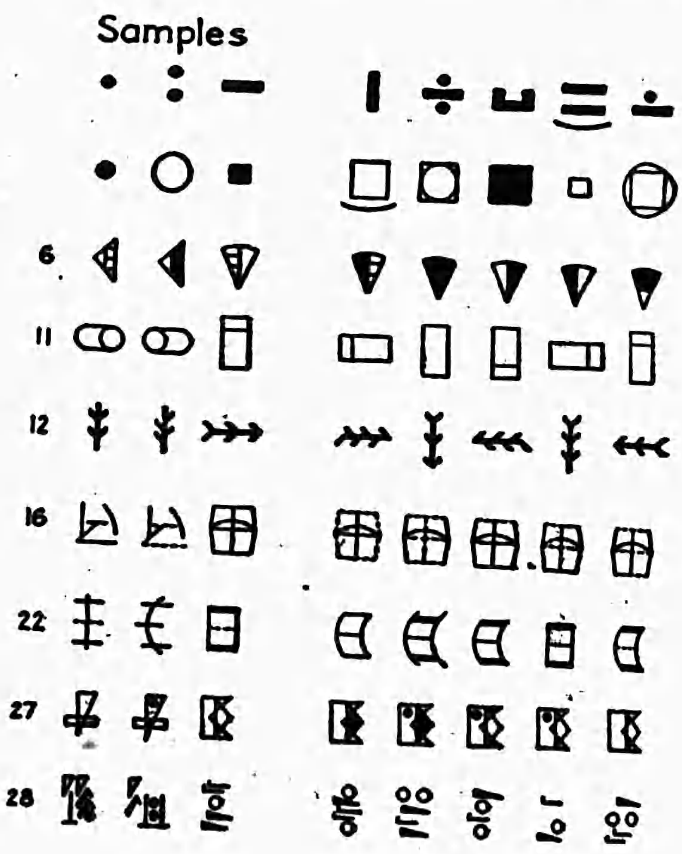
SENTENCES. Do not mark this column.	TRANSLATIONS. Put all your marks in this column.
1. He agrees.	Ajer frega.
5. He is agreeable.	Aj nar fregef.
9. I will impress them.	Av staguda ajay.
12. Pristil bana fregil.	Decision is agreeable.
17. That expression impressed him.	Rin banef loguda aj.

Analogies

DIRECTIONS: This is a test of your ability to see relationships between drawings or figures. The two samples, A and B below, illustrate what you are to do. Look at sample A. Notice that in going from the first to the second figure, there are two dots in the place of one. Then a short line is given. At the right the two short lines are underlined because one line has the same relation to two lines as one dot to two dots.

Now look at sample B. The large white square is underlined. In this case the underlined figure is larger and of the opposite color, as compared with the third figure. Thus, the third figure must be related to the underlined figure in every way that the first figure is related to the second.

In the examples which follow you are to UNDERLINE THAT FIGURE TO WHICH THE THIRD FIGURE BEARS THE SAME RELATION AS THE FIRST FIGURE BEARS TO THE SECOND FIGURE. Underline one and only one figure in each row.



Arithmetic

Directions: Write the answers to as many of these problems as you can in the time allowed.

5. A man owns $\frac{2}{3}$ of a boat and sells $\frac{3}{4}$ of his share for $750. At this rate, find the value of the boat.

   Answer: $____________________

6. A train required 6 hours for a trip of 200 miles, making 18 stops of 5 minutes each, and 1 stop of 30 minutes. What was the average speed of the train when in motion?

   Answer: __________miles per hour.

9. A church was insured for $12,000 in one company at $\frac{7}{8}$ percent, and for $9,500 in another company at 1 percent. Find the total cost of the insurance.

   Answer: $____________________

16. How many times as big an area has a circle with a radius of 9 as a circle with a radius of 3?

   Answer: ___________times as big.

Opposites

Directions: Each group of four words in the thirty lines below contains two words which are either (*a*) the same or nearly the same in meaning, or (*b*) the opposite or nearly the opposite in meaning. Find the two words in each group that are either the same or opposite and write the numbers of these two words in the column at the right, headed "Same" or the column headed "Opposite" as the case may be.

The first group of words, "1. many, 2. ill, 3. few, 4. down," contains two words ("many" and "few") that are opposite in meaning, so that the figures 1 and 3 are entered in the column headed "Opposite." The second and third groups have also been marked correctly.

				Same	*Opposite*
1 many	2 ill	3 few	4 down	......&......	...1.&.3...
1 fair	2 near	3 gone	4 far	......&......	...2.&.4...
1 gay	2 last	3 long	4 happy	...1.&.4...	......&......
1 gaudy	2 priestly	3 husky	4 showy	......&......	......&......
1 simplified	2 destined	3 fated	4 directional	......&......	......&......
1 lax	2 soft	3 strict	4 lazy	......&......	......&......
1 legible	2 illegal	3 ineligible	4 readable	......&......	......&......
1 puzzling	2 oratorical	3 witty	4 enigmatic	......&......	......&......

It will be noted that the five subtests of this examination are quite different in type and content. All have been found to be predictive of general scholarship in college. Three of the tests are linguistic or primarily verbal whereas the other two are quantitative in character. For the prediction of general college scholarship such tests as this have been found to be valuable even without the addition of other tests or prognostic material.

Achievement tests in specific subjects.—The second type of college aptitute test material—the testing of achievement in specific subject matter studied in high school but which is also found in college will be illustrated by the Chemistry Training section of the Iowa Placement Examinations (New Series C. T., form Y). In this test the close relation between the scores of the test and college scholarship arises in part because of the general factor found in linguistic and quantitative material such as was found in general scholastic aptitude tests, but it has this added factor: The material in it and the subject which is to be studied in college actually overlap in content. This overlapping can be accomplished because the high-school course of study in chemistry overlaps in content that of the college course.

The results on this type of test can be used to advantage for purposes of classification or placement. In general the predictive values of these tests for *individual college subjects* are greater than those of the scholastic aptitude tests. If several such tests of individual subjects are used the composite score will predict *general college scholarship* better than the one scholastic aptitude test.

PARTS OF THE IOWA PLACEMENT EXAMINATION–CHEMISTRY TRAINING

PART 3

DIRECTIONS: Examine each statement below and decide whether it is true or false. If the statement is true encircle the *T*; if the statement is false encircle the *F* at the right of each statement. *Do not guess.* You have 7 minutes for part 3.

SAMPLES: Gold is heavier than copper. — T F

Great copper deposits are found in Pennsylvania. — T F

3. When phosphorous burns dense clouds of smoke are formed. — 3. T F
7. Oxygen may be prepared from liquid air by allowing the liquid nitrogen to boil off. — 7. T F
22. Starch is added in the manufacture of baking powders to provide a substance which will form acid when water is added. — 22. T F
25. Hydrofluoric acid cannot be safely stored in glass containers. — 25. T F

PART 4

DIRECTIONS: Solve the following problems, and place the answer to each problem on the line at the right. Do not spend much time on any problem. Use the margins of this page for figuring. You have 16 minutes for part 4.

3. Potassium perchlorate ($KClO_4$) is 46 percent oxygen. If 100 grams of it are heated until only KCl remains, how much KCl will there be? 3. ....

6. Given 50 cc of oxygen at 0° C. The gas at constant pressure is heated to 273° C. How many cc does it then occupy? 6. ....

Aptitude tests in specific subjects.—The third type of testing—that of testing for aptitude in individual subjects not yet studied by the individual—will be illustrated by the Chemistry Aptitude section of the Iowa Placement Examinations (New Series, C.A., form Y). This test assumes that the student has had no chemistry or at the most only such general information as he would pick up because of his own interest outside of any regular class. The question items have been so constructed that they conform as closely as possible to the method of study and content of chemistry without assuming a knowledge of that subject.

PARTS OF THE IOWA PLACEMENT EXAMINATION—CHEMISTRY APTITUDE

PART 1

DIRECTIONS: Solve the following problems. Place the answer to each problem on the dotted line at its right. Do not spend much time on any one problem. Use the margins of this page for figuring. You have 15 minutes for part 1.

2. Solve for y: $\frac{y}{2}=3x^2$. 2. ....

4. Solve the expression for T_1: $\frac{P_1V_1}{T_1}=\frac{P_2V_2}{T_2}$ 4. ....

9. Solve: $(120)\left(\frac{273}{360}\right)\left(\frac{750}{760}\right)=$ (to 1 decimal place.) 9. ....

10. If 3 pounds of 50-cent tea are mixed with 7 pounds of 60-cent tea, what is the value per pound of the mixture? 10. ....

15. A sample of flour weighing 8 grams, on drying loses 2 grams of water. What percent water was the original sample? 15. ....

Section B

At times an ion may contain more than one kind of element. For example, the formula for nitric acid is $H(NO_3)$. In this case (NO_3) acts as an ion. As one (NO_3) ion combines with one hydrogen ion which has 1+charge, the (NO_3) group must have 1− charge.

Keeping in mind the fact that the number of charges + and − must always balance, correct the formulas below. If the formula is correct as it stands write a large C after it. For example if $H_2(NO_3)$ were given you would correct it to read $H(NO_3)$ because (NO_3) with 1− charge can combine with one and only one H+ ion. If ZnS were given you would mark it C as both ions carry 2 charges and therefore the charges balance.

A table of ions and the number of charges is provided.

Zn^{++} = 2+	Fe^{+++} = 3+	$(SO_4)^{--}$ = 2−	$(PO_4)^{---}$ = 3−
K^{+} = 1+	H^{+} = 1+	$(NO_3)^{-}$ = 1−	$(S)^{--}$ = 2−
		$(OH)^{-}$ = 1−	

SAMPLE:	$H(SO_4)$	----$H_2(SO_4)$---- (corrected formula)
2.	$K_2(SO_4)$	2. ----------
4.	$Fe(NO_3)_2$	4 ----------
7.	$Zn(PO_4)$	7. ----------
8.	$Fe(SO_4)$	8. ----------

POSSIBLE IMPROVEMENTS IN THE USE AND INTERPRETATION OF TEST RESULTS FOR COLLEGE PREDICTION

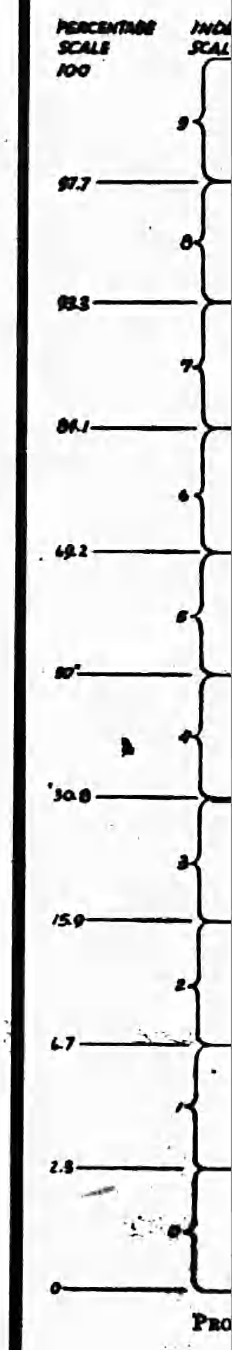
The use of tests for college prediction may be improved in a number of ways. In the learning of college students many different mental elements or processes are involved. It seems evident, therefore, that improvements in the use of tests may come about through an extension of testing to cover these various diverse elements. The bare relationship existing between a series of predictive measures and college success measures do not always tell the whole story. Interpretation of test results differ to some extent according to other circumstances. Also it is believed that there are many aspects of the guidance of entering college students in which testing has not been applied with vigor. It is the purpose in this part of the chapter to discuss various aspects of this problem both as to increasing the efficiency of the testing program itself and the interpretation of the results.

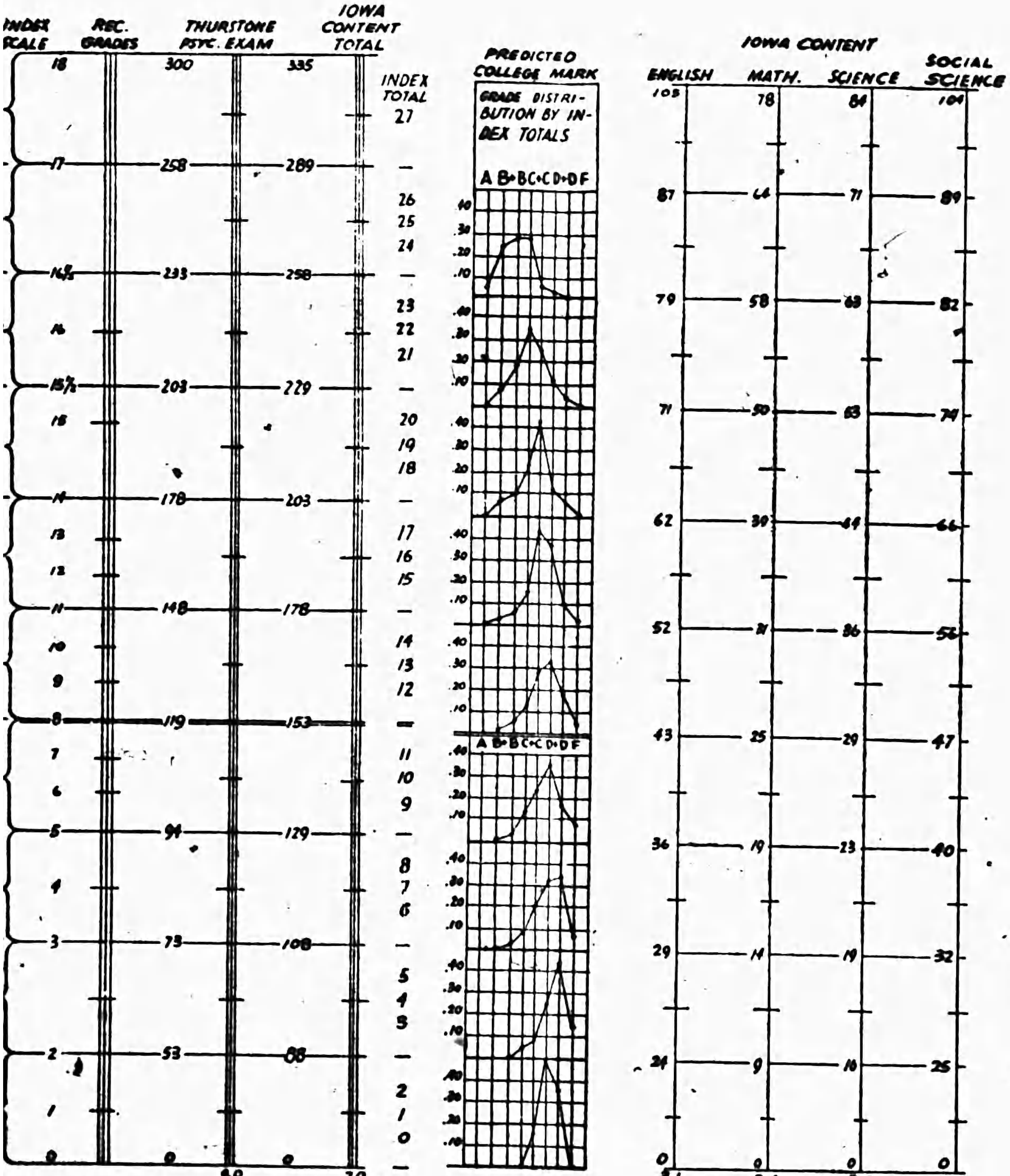
Use of test batteries.—College guidance practices can be improved by the use of test batteries in place of single tests. It may be shown theoretically that the addition of tests to one or more tests already being used will improve the result; and that this improvement will be greatest if the tests added have a low correlation with the tests already being used. The results of using several different tests in the practical situation also shows this to be true. However, much more can be accomplished along this line through choosing the right tests. The search for "independent traits" will eventually show more clearly than ever before the particular types of tests which should be used to cover all the different abilities necessary for the study of college subjects. Of course, for prediction purposes it is not necessary to know exactly the nature of the trait or traits which are being tested as long as the predicting relationship is satisfactory. A knowledge of traits existing in man will make it much easier to construct tests for predictive purposes. This principle is more important than increasing the length of tests. Increasing the length of tests which have intrinsically high reliabilities has limited possibilities. An example of increase in efficiency through picking out the right kind of tests is that of the American Council Examination. Successive editions of this test have had progressively higher correlations with college success (Thurstone, 126). This was brought about by careful selection of tests and not by the lengthening of tests.

The results given in part II, chapter III, show clearly that when several tests or ratings are used a closer predictive relationship is possible than when a single test or rating is used. Also if the tables in that section will be examined carefully it will be seen that the two or more items which when combined give the best predictive values tend to be those items which are most unlike each other. A general mental examination plus some achievement test result or marks in high school or a reading (comprehension) test seem to give good results.

An excellent combination for predicting college scholarship which has been tried out in practice for several years is the use of three factors: (1) scholarship as shown by high-school marks through counting the number of recommended marks;[1] (2) the results on a general mental test; and (3) an objective test on some of the most common elements of high-school achievement. The method may be seen a little clearer by reference to figure I. The double line running through the

[1] The term "recommended marks" is a local term meaning marks of a specified high quality.





Probable Errors of Scores

Probable Errors of Scores

Courtesy of S. L. Brintle, vice-principal, Junior College, Long Beach, Calif.

Figure 1.—Use of test results in the guidance of entering college students

87924—35. (Face p. 18.)

30.8-percent value indicates the level which must be obtained to be admitted to full standing in the regular college course. The total index value for the 3 factors must equal 12 in order to attain this level. Since these charts are used for guidance purposes apart from the strictly administrative purpose of selecting students the marks which various levels of students will make are also given. Also for this purpose the scores on the different parts of the Iowa test are inserted. Through this graphical method of showing results and through the excellence of the method itself, guidance of a superior quality for entering college students is attained. Though the institution (Long Beach Junior College, Long Beach, Calif.) in which the system is used is publicly controlled little friction with the public has resulted.

Research has shown that the addition of a totally different kind of measurement—that of interest rating through questionnaires—adds substantially to the predictive value of achievement and general mental tests.

IMPROVEMENT OF THE INTERPRETATION OF CORRELATION COEFFICIENTS THROUGH AN UNDERSTANDING OF THE EFFECT OF RANGE OF ABILITY

The effect of range on correlation coefficients is likely to be lost sight of in prediction problems because we are not dealing directly with certain grades or ages. Its effect is there just the same. If a preliminary selection of students is made before tests are applied the resulting correlation coefficients suffer from this restriction in range. This is probably the most important reason why predictive values of tests are greater when considering prediction of success in high school than for those used in college.

If tests were given to beginning high-school students it is possible that higher correlations with success in college will be found that is the case where the tests are given to a select group already admitted to college upon the basis of certain scholastic qualifications.

Blair (9) found that correlation between scholarship (in marks) in high school and college decreased as the marks of succeeding years of high school were used.

TABLE I.—*Relationship between first-term college scholarship and certain averages of high-school marks*[1]

High-school marks averaged for—	Correlations	
	Number system	Letter system
First year	.691	.481
Second year	.605	.441
Third year	.601	.406
Fourth year	.584	.453
English I	.575	.516
English II	.522	.421
English III	.543	.346
English IV	.560	.492

[1] After Blair (9).

If this is found upon further research really to be the case it would indicate that guidance work relating to the selection of a college, kind of college work, etc., could begin with profit some time before graduation from high school. It would also indicate that one should not condemn too severely any correlation result which has obtained from a select group of college entrants. If records are made beginning in the early years of the high school, comparisons may be made with a much better perspective. There has not been much evidence collected at this time regarding the positive value of cumulative records but all the evidence and thinking on the subject point inevitably to the conclusion that such records will bring about more accurate guidance than can be obtained by a set of examinations taken at any particular time.

Besides the evidence cited from Blair there is the evidence submitted by Keys who studied the question of value of continuing test results over a period of semesters or years. Keys used records in the elementary school subjects. As is known, the acquisition of knowledge in the elementary school subjects can be measured fairly accurately. For this reason it may be assumed that cumulative records in those subjects will have their most severe test. Keys found that cumulative records increased the prediction accomplishment of work in those subjects up to 3 years. With the result found in this field it seems possible that cumulative records in other lines and in high school may be of considerable value. We need studies in this area to determine the value of cumulative records. When such records are used it can be determined a year or two in advance whether or not a certain college will admit a particular high-school student. This leaves the student and the college both more free to devote their thinking and energies towards the values in the curriculum itself and in the guidance of the individual into the channels in which he will do best.

Use of marks.—The almost universal use of marks as an index to college scholarship is excusable because the aptitude tests in this field were developed before adequate achievement tests were in use. However, through using teachers' marks as an index the predictive value of tests has not received due credit. Even without an improvement in the quality of aptitude testing, the efficiency of prediction will become greater because of the increased use to which college teachers put objective tests in their own subjects and the increased use of comprehensive examinations.

There is evidence at hand to show how the predictive correlation coefficients are increased because of the use of objective tests. Frasier and Heilman (44) report the correlation found between the Thorndike Intelligence Examination and college achievement determined first, subjectively and, second, objectively. The statement of their results showed an average coefficient of .45, with a range from .24 to .57 for the subjective method while for the objective method the average coefficient was .60 and the range from .46 to .69. In the 1933 College Sophomore Testing program (93) 300 sophomores who took the English test had taken an English placement test at the beginning of their freshman year. The correlation between these two tests was .77, while the correlation between the placement test and the 2-year average college was only .47. These two coefficients are not strictly comparable but they are indicative. Tharp (123) found that the Iowa Placement Test for Foreign Language Aptitude with 388 students correlated .47 with semester marks in French 1A and with 379 students .64 with achievement test results in French. The new comprehensive testing program now being organized by the University of Chicago and other schools (Jones, 66) will give ample opportunity for needed research in this field.

ARTICULATION OF HIGH SCHOOL AND COLLEGE

One reason for our difficulty in articulating the work of the high school and college and for the low correlation existing between high-school marks and college marks is that the curriculums of the schools are not very closely allied. An example illustrating this is furnished by Kellogg (72). He finds at McGill University (Canada) that matriculation examinations correlate highly with beginning college work at McGill whereas intelligence tests correlate very moderately. He says this is because the college matriculation examinations are made by the regular college instructors and because there is no break at the college as is found in American universities. The material taught in the lower school is continued in college. He indicates that this conclusion is corroborated when he finds that where matriculation examinations and intelligence examinations are correlated with later college work the correlations with the matriculation examinations are decreased, whereas the correlation with intelligence tests are not appreciably reduced.

The junior college is breaking down this stratification of secondary education and as this is done the guidance of students from high school or college will be made easier. It will be possible then to rely more on achievement records made in high school.

According to the way the subjects in the high school are now taught there is little to choose among them for efficiency in preparation for college. Several studies have been made of this question. Bolenbaugh and Proctor (10) after investigating the values of different patterns came to this conclusion: "Colleges can well afford to give high schools more freedom in the matter of courses taken by prospective college students. A good high-school record, regardless of the pattern of subjects taken when considered with a standard intelligence test, such as the Thorndike Examination, is a better basis of selecting candidates for college admission than either the high-school record or the intelligence test alone." Douglass (35) after investigation in the field states: "Compared on the basis of predictive usefulness for psychological scores, high-school marks and principals' ratings on college promise, the pattern of high-school credits is obviously and definitely inferior." Nelson (89) comes to approximately the same conclusion after a study of college-entrance requirements. Sorenson (113) believes that certain subjects are more valid as conditions of college success than others. His study is the only recent study to which the conclusions are not in agreement with those already quoted.

RELATIONSHIP BETWEEN PREDICTIVE ITEMS AND COLLEGE SCHOLARSHIP AS AFFECTED BY THE PURPOSES IN ATTENDING COLLEGE

Among the influences or factors which bear upon the problem of the relationship of aptitude to achievement in the college area is the seriousness with which college students attack their work. Crawford (31) found that the fixity of life purpose was instrumental in changing predictive correlations. By noting the seriousness of the purposes of attending college into five groups, A, B, C, D, and E, he found the following correlations between college marks and an intelligence test given at entrance:

A	.57	D	.40
B	.43	E	.32
C	.36		

He found a similar state of affairs when other ratings of seriousness in pursuing college work were used. These results show how the prognostication from an aptitude test may in part be nullified. The questions in this connection are these: Should not students who have no strong desire to attend college be allowed to enter other lines of endeavor? Should not students entering college be told what is expected of them so that any failure on their part to measure up will be chalked up against them? In an experimental sense students who do not do their best in college should not be encouraged to attend.

Considering these various aspects of this problem it may be concluded that no matter how perfect our instruments for prognosis may be factors outside our control will nullify some of our recommendations. The only other thing which can be done is to rate or test attitudes and estimate studiousness and use these results in connection with the more orthodox tests and measurements. Interest tests may also be added here. Research has shown ratings in interests tests to have some validity. Further research can profitably be carried on in this field.

USE OF TESTS IN DISTINGUISHING THE DIFFERENCES IN THE ABILITIES OF AN INDIVIDUAL

Attention has not been given to differences in abilities in the individual. The high school has tended to cover up differences in performance. Gowen and Gooch (51) through observing that the correlations between marks in different high-school subjects were fairly uniform, came to the conclusion that marks in different high-school subjects were not very indicative of what subjects the individual would do best in college. They think that school (high-school) competition does not bring out great effort in one line but the "demand is that the students shall do well, or at least do passing work, in many different lines of study." The student who does one subject easily lays it aside as soon as possible to spend more time on the subject which is harder for him. Our school system tends rather to emphasize a more generalized factor of pride in accomplishing a more equalized grade of work in many lines.

If this is true it is all the more necessary, by the end of high school, to ascertain by other means those scholastic lines in which the student seems best fitted, if any, or change the emphasis of high-school instruction so that the "gentleman's passing mark" is not the accepted goal for high-school students. Therefore, it seems desirable to use tests to discover special ability bents as well as tests for purposes of discovering general college ability. The method of differential prediction has possibilities in this regard. Especially is this kind of prognosis needed when there is not much distinction in the admission of students. If students have been admitted to college without much knowledge of their capabilities, either general or specific, it becomes the problem of the college to so place these students that they will get the greatest good out of their college course.

EXTENSION OF OBJECTIVE MEASUREMENTS TO NEW AREAS OF STUDENT ACTIVITIES

Although we have discussed the improvement in the prognosis of college success from the addition of tests of traits as independent of one another as possible, we have not discussed to any great extent the possibilities of using other types of evidence than that furnished by test results. This should not be taken to mean that other evidence should not be used. As college instructors become increasingly adept at instilling attitudes, appreciations, etc., and their evaluation, other forms of aptitude evaluation beyond that of tests should be added. It has already been shown in the studies reported that interest questionnaires have value in the prognosis of college success. Tyler (129) mentions various methods of evaluating behavior which may become of value in supplementing one another and objective examinations. Many of these methods are objective in type. He cites the following: (*a*) Observation of behavior; (*b*) laboratory methods; (*c*) written examinations; (*d*) personal interviews; and (*e*) the collection of products made by individual students As research brings out their worth, such methods of evaluating educational projects should be experimented with in searching for instruments of prognosis.

RELATION OF THE PLACE OF THE SCORE IN THE DISTRIBUTION TO ITS USE IN PREDICTION

The correlation coefficient indicating a relationship between a predictive item must be considered in the light of the place of the score in the distribution and the use to which the results are to be put. An extreme score is, for instance, always more indicative of high or low ability than any other score. This seems self-evident but it is an important point in this problem.

Also a predictive instrument with a low predictive power may serve the purposes of some colleges. If the objective of a college is to allow entrance to a large number of students with the expectation that a fair number of them will fail, the predictive instrument need not be so highly predictive. However, if the aim of the college is to pick out only successful students to start with, the finest instrument for measuring aptitude available would not be entirely adequate.

The statistical methods and results given in part II support the principles here set down and describe in detail how prediction of college accomplishment may be set up or investigated.

PART II : STATISTICAL METHODS AND RESULTS OF STUDIES OF COLLEGE PREDICTION

CHAPTER I : STATISTICAL METHODS FOR USE IN THE PREDICTION OF COLLEGE SCHOLARSHIP

GENERAL CONSIDERATIONS

IN PREDICTING success in college we have, on the one hand, to do with items used in making the prediction, and on the other, with some measure of success in college. The items used in making the prediction may be called predictive items. Consider a common practice in college prediction—the use of a general mental test in aiding the determination of admission to college or guidance at the time of entrance to college. In this case the college is concerned with the practical problem of determining college material. What it does in practice is something like this. A general mental test is administered to each entering student. The result of this test is used in aiding in the determination of fitness for entrance. In considering the test results for this purpose some standard meaning of the scores is used. For instance, if a student obtained a score of 68 on the examination, it is referred to one or more critical points which the college has set up for determining entrance. If this college has agreed to eliminate entirely all students below a score of 60 and to take only those with good high-school grades who have test scores of 60 to 70, this student would be admitted upon the basis of his actual high-school marks. In this practical procedure the actual success of the students in college after admission is not considered directly. Therefore, in such a set-up of the prediction of college success one is directly concerned with the predictive items. There may not be any direct check-up to see whether or not the predictions made are correct. However, in this procedure the assumption is made that the tests used are actually predictive measures.

If a college is not accepting the practice in other institutions or the results of research studies made elsewhere, it must go further and investigate for itself the efficiency of any method of college prediction. If in the case of the college situation mentioned the college authorities also had gone ahead after the students had been admitted, and compared their actual success with the predicted success, an evaluation of the method would have been made. This distinction between the practical administrative measure necessary for admitting students and evaluating the general efficiency of prediction should be recognized.

The correlation coefficient.—The basic concept which has made prediction possible and which is the basis for evaluating the efficiency of prediction is that of correlation. By correlation is meant the relationship existing between capacities or traits or various mental and physical functions. For example, in guidance it is very important to know as clearly as possible the relationship between the predictive items and the success criteria. This relationship is indicated best through the correlation coefficient r. The Pearson product-moment r is the best expression of this correlation. Its use is subject to certain limitations. For instance, this r should be used only with linear or fairly linear data and with data which may be considered continuous. Methods of noting whether or not data may be subject to such limitations and the corrections which may be applied will be discussed.

The correlation coefficient ranges in value from +1.00 to −1.00. +1.00 indicates a perfect positive relationship and −.1.00 indicates a perfect negative relationship. Practically all coefficients with educational material will be found to be significantly less than these two extremes.

There are a large number of different formulas for computing r. Symonds in the October 1926 issue of the *Journal of Educational Psychology* presents some 52 different formulas for obtaining r. It is possible to get from this article a very clear conception of the similarity or difference between the various formulas and also statement as to where and when the more important of the formulas first appeared in the literature.

Most correlations are worked from "scattergrams" but some are worked directly from the original distribution of paired facts. The "scattergrams", besides being a part of a method for calculating the correlation coefficient, may also be used through inspection to determine the linearity of the data.

The general form of Pearson's product-moment coefficient correlation for deviations from the true mean is

$$r = \frac{\Sigma x_1 x_2}{N \sigma_1 \sigma_2} \qquad (1)$$

where "r" is the coefficient correlation; x_1, x_2, is the summation of the product of the deviations of each measure from their true means; σ_1 and σ_2 are the standard deviations of the x_1 and x_2 distribution; and N is the number of cases used.

Since

$$\sigma x_1=\sqrt{\frac{\Sigma x_1^2}{N}} \text{ and } \sigma x_2=\sqrt{\frac{\Sigma x_2^2}{N}} \tag{2}$$

The above formula becomes

$$r=\frac{\Sigma x_1 x_2}{\sqrt{\Sigma x_1^2}\sqrt{\Sigma x_2^2}} \tag{3}$$

where ϵx_1^2 and ϵx_2^2 are the summations of the squares of the deviations from the true means of their respective distributions.

The two preceding formulas are used in actual calculations hardly at all, but are basic for other formulas. The formula for the correlation from an assumed mean is

$$r=\frac{\frac{\Sigma x_1'x_2'}{N}-\frac{\Sigma x_1'}{N}\cdot\frac{\Sigma x_2'}{N}}{\sqrt{\frac{\Sigma(x_1')^2}{N}-\left(\frac{\Sigma x_1'}{N}\right)^2}\sqrt{\frac{\Sigma(x_2')^2}{N}-\left(\frac{\Sigma x_2'}{N}\right)^2}} \tag{4}$$

We shall illustrate the calculation of "r" by this formula using a scattergram. This calculation will also illustrate the calculation of the σ of a distribution, i. e.,

$$\sqrt{\frac{\Sigma(x')^2}{N}-\left(\frac{\Sigma x'}{N}\right)^2}$$

when it is being calculated from an arbitrary origin or a guessed mean.

Assume that 100 entering college students were given the American Council on Education Psychological Examination and that at the end of their sophomore year they were given a comprehensive examination covering all phases of the work of the first 2 years of college. For each score on the psychological test there will also be a corresponding composite score on the comprehensive examination. Each of these pairs of scores has been placed in the scattergram in its appropriate cells. For instance, the student who obtained a score of 248 in the psychological examination and a score of 540 on the composite comprehensive achievement test is found in the first cell in the upper right-hand corner. The figures in the centers of the cells show the number of cases in each.

Diagram showing the calculation of the product-moment coefficient of correlation between the scores on the American Council on Education Psychological Examination taken at entrance to college and the scores on a comprehensive achievement examination given at the end of the sophomore year in college with 100 students

PSYCHOLOGICAL EXAMINATION SCORES

Comprehensive achievement test scores	0–119	120–139	140–159	160–179	180–199	200–219	220–239	240 up	f	x'_1	fx'_1	fx'^2_1
–Up 525							(9) 1 9	(12) 1 12	2	+3	6	18
524–450				1	(2) 1 2	(12) 3 4	(24) 4 6	(16) 2 8	11	+2	22	44
–449 375				2	(5) 5 1	(10) 5 2	(9) 3 3		15	+1	15	15
–374 300		1	2	6	10	5	1		25		+43	
–299 225	(3) 1 3		(5) 5 1	8	(–6) 6 –1	(–4) 2 –2			22	–1	22	22
–224 150		(8) 2 4	(10) 5 2	3	(–6) 3 –2		(–6) 1 –6		14	–2	28	56
–149 75	(9) 1 9	(18) 3 6	(3) 1 3	2	(–3) 1 –3				8	–3	24	72
–74 0	(24) 2 12	(8) 1 8							3	–4	12	48
f	4	7	13	22	26	15	10	3	100		–86 +43 —— –43 = $\Sigma x'_1$	275 = $\Sigma x'^2_1$
x'_1	–3	–2	–1		+1	+2	+3	+4				
fx'_1	12	14	13	–39	26	30	30	12	+98 – 39 = +59 = $\Sigma x'_2$			
fx'^2_1	36	28	13		26	60	90	48	301 = $\Sigma x'^2_2$			

The calculation of $\Sigma x'_1$ and $\Sigma x'^2_1$ is shown at the right in the diagram. The frequency for each row is given under the heading of f. The figures in the x'_1 column indicates the number of steps away from the assumed or guessed mean. The assumed mean in the case being taken at the 300–374 step. The fx'_1 are obtained by multiplying the figures in the frequency column f by the figures in the steps column x'_1. Each portion of the fx'_1—the plus portion and the minus portion—is added separately. This gives +43 and –86. These two added give –43, which is therefore the sum of the fx'_1 and is called $\Sigma x'_1$. The third column fx'^2_1 is obtained by squaring each figure in the x'_1 column and multiplying this result by the figures in the f column. The sum of the fx'^2_1 is found to be 275. This is called $\Sigma x'^2_1$. In a similar manner the frequencies in the columns of the scattergram are added and the $\Sigma x_2'$ and $\Sigma x'^2_2$ are obtained. This calculation is shown along the bottom of the scattergram. Therefore the

$$\sigma x'_1 = \sqrt{\frac{\Sigma(x'_1)^2}{N} - \left(\frac{\Sigma x'_1}{N}\right)^2} = \sqrt{\frac{275}{100} - \left(\frac{-43}{100}\right)^2} = 1.60$$

and

$$\sigma x'_2 = \sqrt{\frac{\Sigma(x'_2)^2}{N} - \left(\frac{\Sigma x'_2}{N}\right)^2} = \sqrt{\frac{301}{100} - \left(\frac{+59}{100}\right)^2} = 1.63$$

To obtain $\Sigma x'_1 x'_2$ the only other element in formula (4) which has not been calculated one must turn again to the scattergram itself. $\Sigma x'_1 x'_2$ means the sum of all cell frequencies multiplied by the step values by rows and columns (x'_1 and x'_2). Referring to the scattergram the process of calculation may be shown by taking again the first cell in the upper right-hand corner. The frequency of this cell is 1, the $x'_1 x'_2$ is 12 and $12 \times 1 = 12$, so 12 has been written in the upper left corner of this cell. In a similar manner the figures given in parenthesis in each cell were computed. In the upper right and lower left quadrants of the scattergram these figures will be positive but in the other two quadrants they will be negative. The sum of these figures $(\Sigma x'_1 x'_2) = 162$.

The completed calculation will therefore be

$$r = \frac{\frac{\Sigma x'_1 x'_2}{N} - \frac{\Sigma x'_1}{N} \cdot \frac{\Sigma x'_2}{N}}{\sigma x'_1 \sigma x'_2} = \frac{\frac{+162}{100} - \left(\frac{-43}{100}\right)\left(\frac{+59}{100}\right)}{1.60x \text{ or } 1.63} = \frac{1.87}{2.61} = .716$$

This correlation coefficient .716 indicates the predictive relationship existing between the scores on the American Council Examination and the standing in the comprehensive achievement test over the first 2 years of college. This index of relationship may be used in the actual calculation of predicted scores for students of about the same caliber and surrounding circumstances as the 100 upon which the index was calculated.

For further directions of this method of obtaining the product-moment correlation coefficient using a scattergram, the texts on statistics given in the references at the end of the chapter should be consulted.

From the scattergram the linearity of the data can be to some degree distinguished. In the scattergram here shown the data are fairly linear. The data in this case are grouped around a nearly straight line as shown in figure 2.

Scattergrams which depart materially from such an arrangement indicate that the correlation coefficient calculated

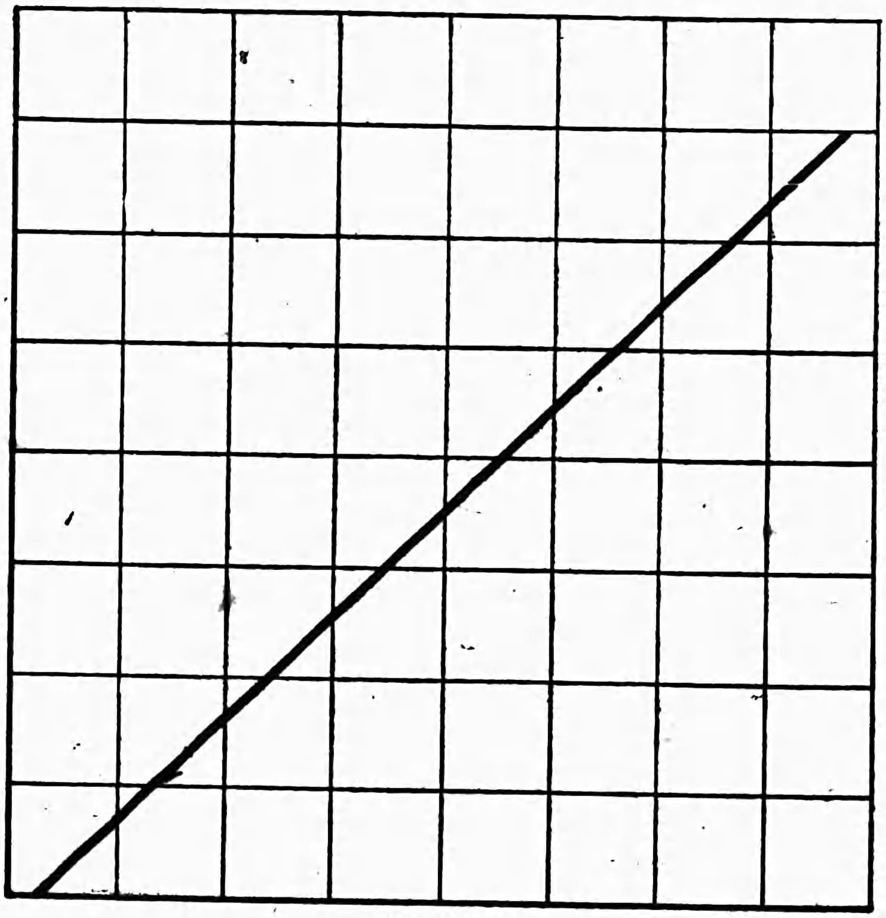


FIGURE 2.

according to the plan given here will not lead to very accurate results without making adjustments.

Standard error and probable error of the coefficient of correlation.—The formula for the standard error of the coefficient of correlation is:

$$\sigma r_{12} = \frac{1 - r^2_{12}}{\sqrt{N}} \qquad (5)$$

or when N is small

$$\sigma r_{12} = \frac{1 - r^2_{12}}{\sqrt{N-1}} \tag{6}$$

The probable error of the coefficient of correlation will therefore be:

$$P.E._{r_{12}} = .6745 \frac{1 - r^2_{12}}{\sqrt{N}} \tag{7}$$

If, for example, a correlation coefficient of .58 is obtained between two measures using 90 cases the P.E. will be equal to $.6745 \frac{1-(.58)^2}{\sqrt{90}} = .05$ and the coefficient of correlation may be written $.58 \pm .05$. This indicates that the chances are even that if another correlation coefficient is calculated upon 90 cases drawn from the same population as the first was drawn that the coefficient of correlation will not vary more than five points lower or higher than .58. (For further discussion of this point see pp. 40–41 and table II, p. 38.) Unless the number of cases is quite large the standard error or probable error should be calculated so that the stability of the correlation coefficient in this regard can be seen.

Correction for fineness of grouping.—If the test results or ratings for either the predictive item or the success criteria are not given in units sufficiently fine then the calculation according to the method just described will not reveal the correct relationship. Whether or not a correction should be applied will depend upon the use to which the result is to be put. There has been much confusion on the part of research workers regarding the use of these corrections. There are, in general, two types of corrections which may be made to correlation coefficients. These are the corrections for (1) the inadequate grouping used and (2) the errors of measurements.

Correction for inadequate grouping.—Test data can usually be put into enough classes so that corrections are unnecessary. School marks and ratings of various types, however, are often given in only a few categories. Correcting for grouping in such cases will raise the value of the correlation coefficient. However, whether or not corrections should be made depends upon the use to which the results are to be put. For example, let us consider a predictive item such as the American Council on Education psychological examination which is being used to estimate what marks will be made in first-semester college English. The American Council test scores may be groups in many different ways. The marks made in English do not have this possibility of grouping in various ways since they are given simply as A, B, C, D, and E where E is the failing mark. If the correlation between the predictive-item groups in intervals of 10 units and the English marks grouped simply as A, B, C, D, and E, was found to be .620, a correction for grouping in the case of the predictive item would be a sound procedure if it is thought that the correlation would be raised thereby. If the number of classes is quite large—20 or more—the grouping error must necessarily be negligible. It is only as the groups are fewer than 10 that this grouping error is usually large enough to be taken into account. Such a grouping error could be taken account of by Sheppard's correction

$$\sigma_c = \sqrt{\sigma^2 - \frac{i^2}{12}} \tag{8}$$

where σ_c is the corrected standard deviation; σ is the standard deviation as originally found; and i is the number of steps in the interval.

In the case of the marks in this sample no correction would be justifiable even though the marks are grouped into only five classes. This is because the marks are to be predicted on the basis of the grouping, i.e., A, B, C, D, and E. Any correction which will predict on the basis of a finer grouping would be contrary to the fact that the prediction must be made in the coarser grouping.

If, however, instead of predicting the grade which the student is most likely to make, it is desired to predict his ability in college English, then it is justifiable to correct for grouping. In this case the correction for grouping cannot be made with Sheppard's correction because the steps in each of the intervals are not the same. In this case the correction for broad categories as developed by Kelley [1] which involves the finding of r first through assuming the distributions are normal by taking the mean of each class and making the calculation of r through the use of the Kelley-Wood tables.

[1] Kelley, T. L. Statistical method. New York, Macmillan Co., 1924. pp. 167-171.

After r has been calculated in this manner it may be corrected by the formula

$$r_{z1z2}=\frac{r_{x1x2}}{\sigma_{x1}\sigma_{x2}} \tag{9}$$

Correction for attenuation.—We have discussed corrections on account of inadequate grouping in the last section. Inadequate grouping nullifies or conceals the accuracy of measurement because of the inadequacy of the units used. A more general correction for the inadequacies of measurement which will cover the ones mentioned, together with other inadequacies of the measuring instruments, is the correction for attenuation

$$r_{\infty\infty}=\frac{r_{12}}{\sqrt{r_{1I}}\sqrt{r_{2II}}} \tag{10}$$

where $r_{\infty\infty}$ is the coefficient of correlation corrected for attenuation, r_{12} is the correlation between any two measures and r_{1I} and r_{2II} are correlation coefficients known as reliability coefficients. As can be seen from the formula the amount of correction depends upon the size of these reliability coefficients. In general, the higher the reliability coefficient the better the test and also other things being equal, it is better to get predictive correlation coefficients as high as possible without correcting for attenuation. For these reasons the reliability coefficient will be discussed before taking up the appropriate situations in predictions when the correction for attenuation may justifiably be made.

This reliability coefficient indicates the stability of a test or scale in getting equivalent results from the application of it to an individual a second time. Strictly, the reliability coefficient for any single test means the correlation of that test with another test having the same mean, the same standard deviation, and which is made up from material similar to that in the first test. Practically, where there are several forms of a test this means the correlation between any two of these forms.

The reliability coefficient may be obtained from the application of one test alone by the formula

$$r_{1I}=\frac{2r_{\frac{1}{2}\frac{I}{II}}}{1+r_{\frac{1}{2}\frac{I}{II}}} \tag{11}$$

where r_{II} is the reliability coefficient and $r_{\frac{1}{2}\frac{I}{II}}$ is the correlation between two halves of a test where the halves satisfy the assumptions above expressed as to equal means, standard deviations, and similar material. This splitting of a test into two halves usually may be done satisfactorily by taking the odd items as one-half and the even items as one-half.

The reliability coefficient obtained by the split halves method is particularly useful when only one form of the test is available, or when the necessary assumptions in regard to two forms of the test cannot be relied upon.

The practice of using the correlation between two administrations of the same test should not be encouraged. The coefficient using this method is too high if the time between tests is short, because of the practice effect and memory factor and too low if the time between tests is considerable because of changes in functional reactions of the students themselves.

The Spearman-Brown formula for correlation between A forms of a test and A other similar forms (the generalized form for finding the reliability coefficient) is of use in this connection. It is

$$r_{AA}=\frac{Ar_{II}}{1+(A-1)r_{II}} \qquad (12)$$

where r_{AA} is the reliability coefficient for a test as long as one with A number of forms; A is the number of forms; and r_{II} is the reliability coefficient between any two forms. With this formula the added value in reliability in lengthening a test or in making it shorter may be ascertained. By solving this formula for A we obtain the following

$$A=\frac{r_{AA}(1-r_{II})}{r_{II}(1-r_{AA})} \qquad (13)$$

where A is the number of forms required r_{II} is the reliability coefficient between any two forms and r_{AA} is the reliability coefficient desired. This formula enables one to find the length of a test given the reliability desired.

In our discussion of the correction for attenuation let us assume again the situation where the American Council Psychological Examination is used to predict certain measures of success in college. Suppose the following correlations have been found:

Psychological examination and first-year French	.54
Psychological examination and first-year average marks	.58

and that the reliabilities have been found to be—

Psychological examination	.94
First-year French	.44
First-year average marks	.53

First, should any correction be made on account of the lack of perfect reliability on account of the psychological examination? Since the instrument is being used as is, no correction for attenuation should be made on its account. It is impossible to correct for a better instrument than is being used in making the prediction. This same reasoning would be valid regarding any test or rating being used for predictive purposes.

On the other hand, should any correction on account of attenuation in first-year college French marks or in average first-year college marks, be made? It depends on the purpose for which the predictions are used. If the prediction is to be the actual mark in French which this student will get, or the actual average mark made by this student in his first-year college work, then no correction is to be made since this actual mark is as unreliable as is shown by the reliability coefficient. However, if the purpose of the prediction is to be the achievement of the students in first-year French apart from the marks given by any one instructor, or the achievement of this student in first-year college work apart from the marks given by any particular set of instructors, then the correction for attenuation in the criterion should be applied. This is because the prediction of success is being made which is not bound by the reliability of the marks of any one instructor or set of instructors.

Therefore a correction for attenuation in the criterion may sometimes be desired in practical prediction problems. In this case the correction for attenuation formula (10) becomes

$$r_{1\infty}=\frac{r_{12}}{\sqrt{r_{2II}}} \qquad (14)$$

where $r_{1\infty}$ is the correlation corrected for attenuation in the criterion, r_{12} is the correlation before correction, and r_{2II} is the reliability coefficient of the criterion. The prediction for first-year French for the unreliability of any one teacher's marks would therefore be

$$r=\frac{.54}{\sqrt{.44}}=\frac{.54}{.66}=.82$$

As will be noted this increase in validity is considerable. Similarly the correction for attenuation for the first-year average mark would be

$$r=\frac{.58}{\sqrt{.53}}=\frac{.58}{.73}=.80$$

When data are not linear in nature or if the data are not continuous the calculation of the correlation is not obtained through the method described. In some cases curvilinear data can be changed to linear data for the calculation of the coefficient of correlation as here described. This does not always solve the problem, however, because the regression equation would have to be adjusted. Some of the indices used in dealing with discontinuous data may be equated to the Pearson coefficient of correlation but others cannot be so equated. Descriptions and formulas for calculating the correlation for curvilinear or discontinuous data may be found in the references given at the end of this chapter.

Barr (4) lists several errors which arise in the interpretation of correlation coefficients. He discusses in particular, three: (1) Errors in estimating the amount of relationship between traits represented by coefficients of correlation of different sizes; (2) errors growing out of the assumption of causal relationship from obtained coefficients of correlation; and (3) errors due to the acceptance of coefficients of correlation at face value. It cannot be emphasized too much that coefficients of correlation are not percentages and that in and of themselves do not indicate causal relationships.

For a clearer understanding of the meaning of the forecasting efficiency of the correlation coefficient it is advisable to use the coefficient of alienation

$$K=\sqrt{1-r^2} \tag{15}$$

or the formula for finding the efficiency of forecasting in percent which is

$$E=1-K \text{ or } 1-\sqrt{1-r^2} \tag{16}$$

The efficiency in percentages for various correlation coefficients are given in the following table:

TABLE II.—*Showing the relation between the correlation coefficient to the percent of forecasting efficiency*

Coefficient	Percent	Coefficient	Percent
10	½	50	13
20	2	60	20
30	5	70	29
40	8	80	40

General principles for choosing items for tests and lengthening of tests.—Lengthening any test will increase its reliability and validity providing these are not zero to begin with. As far as reliability is concerned by lengthening the test indefinitely a reliability of 1.00 is approached. The reliability of tests to be used in prediction is only one element in the situation. Lengthening of a test indefinitely does not necessarily bring its validity up toward 1.00. The validity of a test has an upper limit of

$$\frac{r_{12}}{\sqrt{r_{II}}} \tag{17}$$

r_{12} is the validity coefficient and r_{II} is the reliability coefficient. By lengthening any one particular type of test the validity is very strictly limited because in any one test the reliability of its different forms would be greater than the validity. This would always make

$$\frac{r_{12}}{\sqrt{r_{II}}}$$

considerably less than 1.00.

The most important thing in aptitude testing in this connection is probably to get as many different tests each of which have high correlations with the criterion, but low correlations with each other. It is the uniqueness in predicting scholastic success that tests are to be judged by. Hull (60) has shown the change in multiple correlation coefficients resulting when intercorrelations are assumed. This can be done by the formula

$$R = r'\sqrt{\frac{N}{1+r''(N-1)}} \tag{18}$$

when r' is the correlation between any of the tests and the criterion, r'' is the correlation between all the tests, assuming all those intercorrelations equal, N is the number of tests, and R is the correlation with the criterion of the combined battery when weighted in the best possible way by means of the multiple-regression equation. We will illustrate the effects on the validity coefficient of increasing the number of tests for two sets of correlations—one when r', the original validity coefficient, and r'', the reliability coefficient, are .60 and .70, respectively, and the other when these coefficients are .50 and .20, respectively.

TABLE III.—*Showing the validities obtained from varying numbers of tests when $r' = .60$ and $r'' = .70$*

Number of tests	Validity coefficients	Number of tests	Validity coefficients
2	.650	10	.702
3	.671	20	.710
4	.682	40	.713
5	.688		
6	.693		

TABLE IV.—*Showing validities obtained from varying numbers of tests when $r' = .50$ and $r'' = .20$*

Number of tests	Validity coefficients	Number of tests	Validity coefficients
2	.646	6	.866
3	.732	10	.945
4	.791	20	1.000
5	.835		

Tables III and IV show very definitely that it is tests with low correlations with each other and high validity which should be sought for if more tests are to be added to those already in use in forecasting.

Although the correlation coefficient indicates the relationship existing between a predictive item and a success criteria it does not give directly the probable limits of error of a predicted success score. Methods for determining predicted scores and their errors will be the subject considered in the remaining part of this section.

Prediction of success from one predictive item.—We shall outline the methods of estimating success from a predictive item or items and give in each case the standard error and probable error formula. This standard error or probable error gives the variation in the estimated success score. The use of these errors will be illustrated in connection with the first case.

CASE I

Where a score on one test ($\overline{X}_1$) is to be estimated from the score on a similar test (X_I) i.e. $\overline{X}_I = X_1$ the standard error of such an estimation is

$$\sigma\sqrt{2-2r_{1I}} \qquad (19)$$

where σ may be

$$\frac{\sigma_1 + \sigma_I}{2}$$

or simply σ_1, σ_1 being the standard deviation of X scores and σ_I is the standard deviation of the X_I scores and r_{1I} is the reliability coefficient—the correlation coefficient found between two forms of the test.

Example: Given

$$X_1 = 80$$
$$\sigma_1 = 22$$
$$\sigma_I = 24$$
$$r_{1I} = .85$$

Then $\overline{X}_I = 80$

and the standard error

$$\frac{\sigma_1 + \sigma_I}{2}\sqrt{2-2r_{1I}} = \frac{22+24}{2}\sqrt{2-2(.85)} = 12.6$$

The probable error is equal to $.6745 \times 12.6 = 8.5$ or $\overline{X}_I = 80 \pm 8.5$.

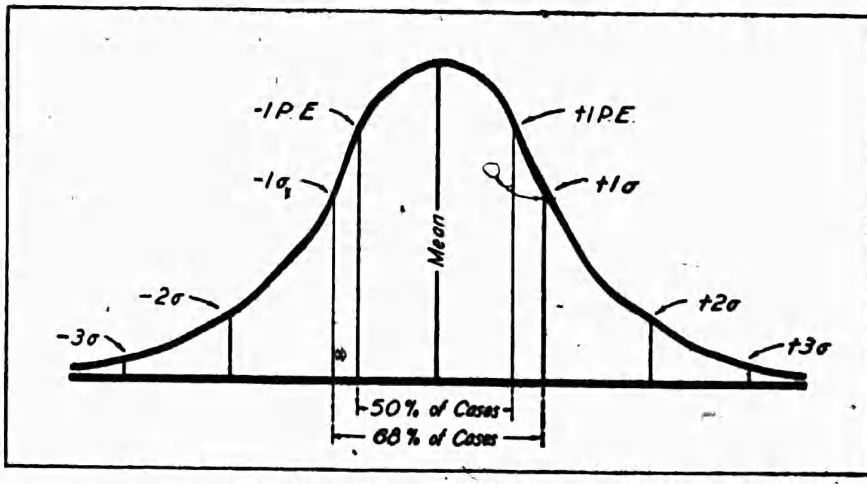


FIGURE 3.—Normal distribution curve.

The ± 8.5 indicates the amount of variation which may be expected from the estimated score 80. The scores resulting from measuring an ability in an individual a large number of times will approximate the normal distribution curve.

The standard deviation (S.D. or σ) is a measure which includes 68 percent of the probability of variation in a score. The standard deviation is used as the unit of variation because it is the most stable of any unit which may be used in describing variations. The probable error (or P.E.) is .6745 times σ and includes 50 percent of the variations when the distribution is normal. The percentage of cases found for various units of σ and P.E. are given in tables V and VI.

TABLE V.—*Percentage of chance of score being within certain probable error limits*

Probable error units	Percent of cases	Probable error units	Percent of cases
±.5	26	±3.0	96
±1.0	50	±3.5	98
±1.5	69	±4.0	99.3
±2.0	82	±5.0	99.96
±2.5	91		

TABLE VI.—*Percentage of chance of a score being within certain limits of the standard error* (σ)

σ units	Percent of cases	σ units	Percent of cases
±.5	38	±2.0	95
±1.0	69	±2.5	99
±1.5	87	±3.0	99.7

Now in the case of the predicted score 80 the chance that another predicted score would still be 80 would be as follows according to table V:

Score range	Percent of chance
80±8.5 (1 P.E.) or 71.5 to 88.5	50
80±17 (2 P.E.) or 63 to 97	82
80±25.5 (3 P.E.) or 54.5 to 105.5	96
80±34 (4 P.E.) or 46 to 114	99.3

CASE II

Where a true score on a test ($\overline{X}_a$) is to be estimated from a single score on a test (X_1)

$$\overline{X}_a = X_1$$

Then the standard error called

$$\sigma_{1.a} = \sigma_1\sqrt{1 - r_{1I}} \qquad (20)$$

Example: Given

$$X_1 = 80$$

$$\sigma_1 = 23$$

$$r_{1I} = .85$$

Then $\overline{X}_a = 80$

and the standard error is

$$\sigma_1 = \sigma_1\sqrt{1-r_{1I}} = 23\sqrt{1-.85} = 8.9$$

The probable error $= .6745 \times 8.9 = 6.0$
Therefore $\bar{X}_a = 80 \pm 6.0$

CASE III

Where a score on one test $(\bar{X}_I)$ is to be estimated from a regressed score on another form of the test $(r_{1I}X_1)$

$$\bar{X}_I = r_{1I}X_1 + (1-r_{1I})M_1 \qquad (21)$$

$$\sigma_{1.I} = \sigma_1\sqrt{1-r^2_{1I}} \qquad (22)$$

Example: Given

$$X_1 = 80$$

$$\sigma_1 = 23$$

$$r_{1I} = .85$$

$$M_1 = 64$$

Substitute in

$$\bar{X}_I = r_{1I}X_1 + (1-r_{1I})M_1$$

We get

$$\bar{X}_I = .85(80) + (1-.85)64 \text{ or } \bar{X}_I = 77.6$$

and substituting in

$$\sigma_{1.I} = \sigma_1\sqrt{1-r^2_{1I}}$$

we get the standard error or

$$\sigma_{1.I} = 23\sqrt{1-(.85)^2} = 12.1$$

The probable error $= .6745 \times 12.1 = 8.2$
Therefore $\bar{X}_1 = 77.6 \pm 8.2$

CASE IV

Where a true score $(\bar{X}_a)$ on a test is to be estimated from a single score on a test X_1

$$\bar{X}_a = r_{1I}X_1 + (1-r_{1I})M_1 \qquad (21a)^3$$

$$\sigma_{a.1} = \sigma_1\sqrt{r_{1I}-r^2_{1I}} \qquad (23)$$

Example: Given

$$X_1 = .80$$

$$\sigma_1 = 23$$

$$r_{1I} = .85$$

$$M = 64$$

[3] The formula for finding the estimated true score from a single score is the same as the one to find one test score given the regressed score on another form of the test.

Substituting in

We get

$$\bar{X}_e = r_{1I} X_1 + (1 - h_{1I}) M_1$$

$$\bar{X} = .85(80) + (1 - 88)64$$

$$\bar{X}_e = 77.6$$

and substituting in

$$\sigma_{e.1} = \bar{\sigma}_1 \sqrt{r_{1I} - r^2_{1I}}$$

we get the standard error or

$$\sigma_{e.1} = 23\sqrt{.85 - (.85)^2} = 8.2$$

and the probable error $= .6745 \times 8.2 = 5.5$

Therefore $\bar{X}_e = 77.6 \pm 5.5$

Any of the first four cases may be used when the prediction is to be based on immediate evidence and without direct correlation with a predicted criterion. For instance, the scores on an algebra test may be used as evidence of mathematical ability without attempting to get a relationship between such algebra scores and later success in mathematics. It is assumed in such cases that algebraic ability is related to later achievement in college algebra, trigonometry, etc. Of course, such use of test results are permissible only where previous investigations have shown substantial relationships. The standard and probable errors in these four cases are therefore variations of measurement only. This limitation in the use of these formulas for purposes of prediction should be recognized.

As to which case is to be used will depend upon the degree of accuracy desired and the time and effort at one's disposal. It depends also on the kind of prediction desired. If the question is: What will this pupil do on another test of similar character? it is one thing and quite different from: What is the student's true ability as measured by this instrument? To answer the first question cases I or III are to be used, whereas to answer the second question, cases II or IV are to be used.

It is to be noted that the formula for obtaining the scores is the same for cases III and IV. The standard errors and consequently also the probable errors for the two are, however, different.

In case V and others following, the formulas for use when standard scores are used are also given. Standard scores are designated as

$$Z_1,\ Z_2,\ Z_3,\ \text{etc.}\ Z_1=\frac{X_1-M_1}{1}-Z_2=\frac{X_2-M_2}{2}$$

etc., where in each case X is the new score, M the mean of the distribution and σ the standard deviation.

Case V

Where a score (X_1) on one test is to be used as a prediction of success on some other measure or criterion ($\overline{X}_o$) not similar in score value to the first named.

$$\overline{X}_o=b_{o1}X_1+(M_o-b_{o1}M_1) \qquad (24)$$

$$\sigma_{o\cdot1}=\sigma_o\sqrt{1-r^2_{o1}} \qquad (25)$$

Where $\overline{X}_o$ is the predicted score, b_{o1} is the regression coefficient and is in this case equal $\frac{\sigma_o}{\sigma_1}$ r_{o1} and r_{o1} is the correlation between X_o and X_1 scores.

Example: Given

$$X_1=45$$
$$r_{o1}=.76$$
$$M_o=50$$
$$M_1=58$$
$$\sigma_o=22$$
$$\sigma_1=28$$

Substituting in $\overline{X}_o=b_{o1}X_1+(M_o-b_{o1}M_1)$
We get

$$\overline{X}_o=\frac{22}{28}(.76)(45)+\left[50-\left(\frac{22}{28}\right)(.76)(58)\right]$$

$$\therefore\ \overline{X}_o=42.1$$

Substituting for the standard error $\sigma_{o\cdot1}=\sigma_o\sqrt{1-r^2_{o1}}$
We get

$$\sigma_{o\cdot1}=22\sqrt{1-(.76)^2}$$

or

$$\sigma_{o\cdot1}=14.3$$

The probable error is equal to $.6745\times14.3=9.6$. When standard scores are used formulas (24) and (25) become

$$\overline{Z}_o=r_{o1}Z_1$$

$$\sigma_{o1}=\sqrt{1-r^2_{o1}}$$

CASE VI

Where a score (X_1) on one test is to be used as a prediction of success on a true measure or criterion ($\bar{X}_\alpha$) not similar in score values to the first named.

$$\bar{X}_\alpha = b_{o1}X_1 + (M_o - b_{o1}M_1) \qquad (24a)^{[3]}$$

$$\sigma_{\alpha \cdot 1} = \sigma_o\sqrt{r_{o0} - r^2_{o1}} \qquad (26)$$

Where the symbols have the same meaning as in case V except that $\bar{X}_\alpha$ represents a true measure of the criterion and r_{o0} is the reliability coefficient of the criterion.

Example: Given

$$X_1 = 45$$
$$r_{o1} = .76$$
$$r_{o0} = .84$$
$$M_o = 50$$
$$M_1 = 58$$
$$\sigma_o = 22$$
$$\sigma_1 = 28$$

Substituting in $\bar{X}_\alpha = b_{o1}X_1 + (M_o - b_{o1}M_1)$

We get $\bar{X}_\alpha = 42.1$ just as in case V since the right part of this formula is the same as that in the other corresponding formula.

Substituting in

$$\sigma_{\alpha \cdot 1} = \sigma_o\sqrt{r_{o0} - r^2_{o1}}$$

We get

$$\sigma_{\alpha \cdot 1} = 22\sqrt{.84 - (.76)^2}$$
$$\sigma_{\alpha \cdot 1} = 11.3$$

The probable error $= .6745 \times 11.3 = 7.6$.

When standard scores are used formulas (24a) and (26) become

$$\bar{z}_\alpha = r_{o1}z_1$$
$$\sigma_{\alpha \cdot 1} = \sqrt{r_{o0} - r^2_{o1}}$$

Cases V and VI are to be used when one measure is used to estimate success in some line of work. This usually means the prediction of success occurring at some future time although there is nothing in the statistical techniques which assumes a time relationship. However, the formula allows for the success criterion to be expressed in a different type of measure than the predictive item. This is a natural condition in prediction problems since success in many cases, if not most cases, would have to be measured by a different test than the prognostic one. For instance, in predicting success in college mathematics a prognostic test would probably be used

[3] The formula for the prediction of the true measure from one test score not similar in score value to the criterion measure is the same as the one where the simple prediction of the measure has been made.

which involved items of mathematical knowledge and reasoning which may have been acquired on the high-school or previous levels of instruction. In testing success at the end of a year of college work in mathematics a much more difficult test than the original test must be constructed and at the same time it must conform to the course of instruction given in the institution. Therefore it may be seen that the prognostic test and the test which is measuring success will often if not usually be different.

CASE VII

Where scores on several tests, X_1, X_2, X_3, etc., are to be used as prediction of success ($\overline{X}_o$) measured by some criterion different from that represented by the scores X_1, X_2, X_3, etc., themselves.

The generalized formula (multiple regression equation) is

$$\overline{X}_o = b_{o1.23\ldots n}X_1 + b_{o2.13\ldots n}X_2 + \ldots\ldots + b_{on.123\ldots(n-1)}X_n + C. \qquad (27)$$

where

$$C = M_o - b_{o1.23\ldots n}M_1 - b_{o2.13\ldots n}M_2 - \ldots\ldots - b_{on.12\ldots(n-1)}M_n$$

where

$$b_{o1.23\ldots n} = \beta_{o1.23\ldots n}\frac{\sigma_o}{\sigma_1}$$

$$b_{o2.13\ldots n} = \beta_{o2.13\ldots n}\frac{\sigma_o}{\sigma_2} \text{ etc. to}$$

$$b_{on.12\ldots(n-1)} = \beta_{on.12\ldots(n-1)}\frac{\sigma_o}{\sigma_n} \qquad (28)$$

and where

$$\beta_{o1.23\ldots n} = \frac{\beta_{o1.3\ldots n} - \beta_{o2.3\ldots n}\beta_{21.3\ldots n}}{1 - \beta_{12.3\ldots n}\beta_{21.3\ldots n}}$$

$$\beta_{o2.13\ldots n} = \frac{\beta_{o2.3\ldots n} - \beta_{o1.3\ldots n}\beta_{12.3\ldots n}}{1 - \beta_{21.3\ldots n}\beta_{12.3\ldots n}}$$

etc. to

$$\beta_{on.12\ldots(n-1)} = \frac{\beta_{on.2\ldots(n-1)} - \beta_{o1.2\ldots(n-1)}\beta_{1n.2\ldots(n-1)}}{1 - \beta_{n1.2\ldots(n-1)}\beta_{1n.2\ldots(n-1)}}. \qquad (29)$$

These β's will finally be reduced to β's, of the order $\beta_{o1.2}$, $\beta_{on.2}$, etc., for which the formulas will be

$$\beta_{o1.2} = \frac{r_{o1} - r_{o2}r_{12}}{1 - r^2_{12}}$$

$$\beta_{on.2} = \frac{r_{on} - r_{o2}r_{n2}}{1 - r_{n2}^2} \text{ etc.} \qquad (30)$$

The general formula using standard scores is

$$\overline{Z}_o = \beta_{o1.23\ldots n}Z_1 + \beta_{o2.13\ldots n}Z_2 + \ldots\ldots - \beta_{on.12\ldots(n-1)}Z_n. \qquad (31)$$

In case it is not thought necessary to obtain β's the general formula is

$$\bar{X}_o = b_{o1.23\ldots n}X_1 + b_{o2.13\ldots n}X_2 + \ldots + b_{on.12\ldots(n-1)}X_n + C \qquad (32)$$

in which the b's are as follows:

$$b_{o1.23\ldots n} = r_{o1.23\ldots n}\frac{\sigma_{o.123\ldots n}}{\sigma_{1.023\ldots n}}$$

$$b_{o2.13\ldots n} = r_{o2.13\ldots n}\frac{\sigma_{o.123\ldots n}}{\sigma_{2.013\ldots n}} \text{ etc. to}$$

$$b_{on.12\ldots(n-1)} = r_{on.12\ldots(n-1)}\frac{\sigma_{o.123\ldots n}}{\sigma_{n.o12\ldots(n-1)}} \qquad (33)$$

where the general formula for $r_{o1.23\ldots n}$ is

$$r_{o1.23\ldots n} = \frac{r_{o1.23\ldots(n-1)} - r_{on.23\ldots(n-1)}r_{1n.23\ldots(n-1)}}{\sqrt{1-r^2_{on.23\ldots(n-1)}}\sqrt{1-r^2_{1n.23\ldots(n-1)}}} \qquad (34)$$

and sigmas of the form

$$\sigma_{o.123\ldots n} = \sigma_o\sqrt{1-r^2_{o1}}\sqrt{1-r^2_{o2.1}}$$
$$\sqrt{1-r_{o3.12}}\cdot\ldots\ldots\cdot\sqrt{1-r^2_{on.12\ldots(n-1)}} \qquad (35)$$

The subscripts for all formulas needed may be obtained from those given here by analogy. The main thing to keep in mind is that the arrangement in the right hand of these equations is always related in the same manner to the subscripts in the left-hand member.

The standard error of estimate is:

$$\sigma est X_o = \sigma_{o.123\ldots n} \qquad (36)$$

$\sigma_{o.123\ldots n}$ is given above (formula 35).

In regard to the order of the integers in the subscripts the order before the period must always be preserved but the order of the integers after the period is immaterial. For instance $r_{12.345}$ and $r_{12.435}$ are the same, but $r_{12.345}$ and $r_{21.345}$ are different. However, where there are only two integers in the subscript and no period, such as r_{12} the arrangement has no particular meaning, i.e., $r_{21} = r_{12}$.

The multiple correlation coefficient may be found by the formula

$$R_{o.123\ldots n} = \sqrt{1 = \frac{\sigma^2_{o.123\ldots n}}{\sigma^2_o}} \qquad (37)$$

All the formulas required for a four variable problem, i.e., one with three predictive items, will be given next using the formulas without β's. At the same time a numerical example will be worked through. For this purpose the following data are given:

$$
\begin{array}{llll}
\sigma_o = 6 & M_o = 11 & r_{o1} = .60 & r_{12} = .51 \quad r_{23} = .41 \\
\sigma_1 = 10 & M_1 = 28 & r_{o2} = .70 & r_{13} = .44 \\
\sigma_2 = 13 & M_2 = 35 & r_{o3} = .64 & \\
\sigma_3 = 18 & M_3 = 43 & &
\end{array}
$$

The predictive equation for these predictive items following the general formula (32) is

$$\bar{X}_o = b_{o1.23}X_1 + b_{o2.13}X_2 + b_{o3.12}X_3 + C$$

where $C = M_o - b_{o1.23}M_1 - b_{o2.13}M_2 - b_{o3.12}M_3$

Following the general formula (33) for the regression coefficient we get

$$b_{o1.23} = r_{o1.23}\frac{\sigma_{o.123}}{\sigma_{1.023}}$$

$$b_{o2.13} = r_{o2.13}\frac{\sigma_{o.123}}{\sigma_{2.013}}$$

$$\text{and } b_{o3.12} = r_{o3.12}\frac{\sigma_{o.123}}{\sigma_{3.012}}$$

The partial r's based on the general formula (34) are

$$r_{o1.23} = \frac{r_{o1.2} - r_{o3.2}r_{13.2}}{\sqrt{1-r^2_{o3.2}}\sqrt{1-r^2_{13.2}}}$$

$$r_{o2.13} = \frac{r_{o2.1} - r_{o3.1}r_{23.1}}{\sqrt{1-r^2_{o3.1}}\sqrt{1-r^2_{23.1}}}$$

$$r_{o3.12} = \frac{r_{o3.1} - r_{o2.1}r_{23.1}}{\sqrt{1-r^2_{o2.1}}\sqrt{1-r^2_{23.1}}}$$

and first order partials as follows must therefore be found. The numerical values for them can be found from the data given.

Calculation

$$r_{o1.2} = \frac{r_{o1} - r_{o2} - r_{12}}{\sqrt{1-r^2_{o2}}\sqrt{1-r^2_{12}}} \qquad \frac{.60-(.70)(.51)}{\sqrt{1-(70)^2}\sqrt{1-(.51)^2}} = .396$$

$$r_{o3.2} = \frac{r_{o3} - r_{o2}r_{23}}{\sqrt{1-r^2_{o2}}\sqrt{1-r^2_{23}}} \qquad \frac{.64-(.70)(.41)}{\sqrt{1-(70)^2}\sqrt{1-(41)^2}} = .542$$

$$r_{13.2} = \frac{r_{13} - r_{12}r_{23}}{\sqrt{1-r^2_{12}}\sqrt{1-r^2_{23}}} \qquad \frac{.44-(.51)(.41)}{\sqrt{1-(51)^2}\sqrt{1-(.41)^2}} = .295$$

$$r_{o2.1} = \frac{r_{o2} - r_{o1}r_{12}}{\sqrt{1-r^2_{o1}}\sqrt{1-r^2_{12}}} \qquad \frac{.70-(.60)(.51)}{\sqrt{1-(60)^2}\sqrt{1-(.51)^2}} = .573$$

$$r_{o3.1}=\frac{r_{o3}-r_{o1}r_{13}}{\sqrt{1-r^2_{o1}}\sqrt{1-r^2_{13}}}\qquad \frac{.64-(.60)(.44)}{\sqrt{1-(60)^2}\sqrt{1-(.44)^2}}=.524$$

$$r_{23.1}=\frac{r_{23}-r_{12}r_{13}}{\sqrt{1-r^2_{12}}\sqrt{1-r^2_{13}}}\qquad \frac{.41-(.51)(.44)}{\sqrt{1-(51)^2}\sqrt{1-(44)^2}}=.241$$

Therefore

$$r_{o1.23}=\frac{r_{o1.2}-r_{o3.2}r_{13.2}}{\sqrt{1-r^2_{o3.2}}\sqrt{1-r^2_{13.2}}}=\frac{.396-(.542)(.295)}{\sqrt{1-(542)^2}\sqrt{1-(.295)^2}}=.294$$

$$r_{o2.13}=\frac{r_{o2.1}-r_{o3.1}r_{23.1}}{\sqrt{1-r^2_{o3.1}}\sqrt{1-r^2_{23.1}}}=\frac{.573-(.524)(.241)}{\sqrt{1-(524)^2}\sqrt{1-(.241)^2}}=.541$$

$$r_{o3.12}=\frac{r_{o3.1}-r_{o2.1}r_{23.1}}{\sqrt{1-r^2_{o2.1}}\sqrt{1-r^2_{23.1}}}=\frac{.524-(.573)\ (.241)}{\sqrt{1-(.573)^2}\sqrt{1-(.241)^2}}=.485$$

The sigmas needed are

$$\sigma_{o.123}=\sigma_o\sqrt{1-r^2_{o1}}\sqrt{1-r^2_{o2.1}}\sqrt{1-r^2_{o3.12}}$$

$$6\sqrt{1-(.60)^2}\sqrt{1-(.573)^2}\sqrt{1-(.485)^2}=3.444$$

$$\sigma_{1.o23}=\sigma_1\sqrt{1-r^2_{12}}\sqrt{1-r^2_{13.2}}\sqrt{1-r^2_{o1.23}}$$

$$10\sqrt{1-(.51)^2}\sqrt{1-(.295)^2}\sqrt{1-(.294)^2}=7.860$$

$$\sigma_{2.o13}=\sigma_2\sqrt{1-r^2_{12}}\sqrt{1-r^2_{23.1}}\sqrt{1-r^2_{o2.13}}$$

$$13\sqrt{1-(.51)^2}\sqrt{1-(.241)^2}\sqrt{1-(.541)^2}=9.129$$

$$\sigma_{3.o12}=\sigma_3\sqrt{1-r^2_{23}}\sqrt{1-r^2_{13.2}}\sqrt{1-r^2_{o3.12}}$$

$$18\sqrt{1-(.41)^2}\sqrt{1-(.295)^2}\sqrt{1-(.485)^2}=13.731$$

Therefore the regression coefficients (b values) become as follows:

$$b_{o1.23}=r_{o1.23}\frac{\sigma_{o.123}}{\sigma_{1.o23}}=.294\times\frac{3.444}{7.866}=.129$$

$$b_{o2.13}=r_{o2.13}\frac{\sigma_{o.123}}{\sigma_{2.o13}}=.541\times\frac{3.444}{9.129}=.204$$

$$b_{o3.12}=r_{o3.12}\frac{\sigma_{o.123}}{\sigma_{3.o12}}=.485\times\frac{3.444}{13.732}=.122.$$

Therefore the equation

$$X_o=b_{o1.23}X_1+b_{o2.13}X_2+b_{o3.12}X_3+M_o-b_{o1.23}M_1-b_{o2.13}M_2-b_{o3.12}M_3$$

becomes

$$X_o=.129X_1+.204X_2+.122X_3+11-(.129)(28)-(.204)(35)-(.122)(43)$$

or

$\bar{X}_o = .129X_1 + .204X_2 + .122X_3 - 5.$

The probable error of estimate is

$$.6745\sigma Est = 6745 \times \sigma_{o.123} = .6745 \times 3.444 = 2.323.$$

The actual use of this equation involves the substitution of scores on the three predictive tests for any one individual and solving for $\bar{X}_o$. Suppose we have scores $X_1 = 35$, $X_2 = 48$, and $X_3 = 60$. Substituting these in the equation

$$\bar{X}_o = .129\ X_1 + .204\ X_2 + .122\ X_3 - 5$$

we obtain

$$\bar{X}_o = .129(35) + .204\ (48) + .122\ (60) - 5$$

$$\bar{X}_o = 16.627$$

or with its probable error

$$\bar{X}_o = 16.627 \pm 2.323.$$

The formula for the multiple correlation coefficient, $R_{o.123}$ according to the general formula (37) would be

$$R_{o.123} = \sqrt{1 - \frac{\sigma^2_{o.123}}{\sigma_o^2}}$$

$$R_{o.123} = \sqrt{1 - \frac{(3.444)^2}{(6)^2}} = .819.$$

CASE VIII

Formulas for the differential prediction using several predictive items have been recently developed (Segel (109)). These formulas should be used when as accurate determination as possible is desired regarding the comparative chance of success between pairs of subjects. In this case, the first order correlation coefficient which is called a differential correlation coefficient, is the correlation of the scores on a predictive item with the difference between two criterion scores. Such coefficients may be calculated in either of two ways. They may be calculated directly, i. e., the correlation coefficients between the predictive items and the difference between two criterion scores may be calculated directly. Such differential coefficients may be designated as $r_{(a-b)x}$,

$r_{(a-b)y}$, $r_{(a-b)z}$, or $r_{(a-b)x_1}$, $r_{(a-b)x_2}$, etc., where a–b indicates differences between two success criteria and x, y, z, x_1, x_2, etc., are scores on various predictive items. Differential correlation coefficients may also be obtained indirectly from correlation coefficients such as r_{ax}, r_{bx}, etc. This may be done by the application of the formula,

$$r_{(a-b)x}=\frac{r_{ax}\sigma_a-r_{bx}\sigma_b}{\sqrt{\sigma_a^2+\sigma_b^2-2r_{ab}\sigma_a\sigma_b}} \tag{38}$$

as developed by Lee and Segel (79).

The use of this formula is recommended where these direct prediction coefficients r_{ax}, r_{bx}, etc., are to be calculated, in any case, or when data have been published in the literature in the form of r_{ax}, r_{bx}, σ_a, σ_b, etc., and the original scores are not available.

Let us consider now the regression equation in the case where differential correlation coefficients $r_{(a-b)x}$, $r_{(a-b)y}$, etc., have been calculated directly and the means and standard deviations and the differences are available.

The general regression equation adapted from formula (32) is

$$\bar{X}_d=b_{d1.23\ldots n}X_1+b_{d2.13\ldots n}X_2+\ldots\ldots+b_{dn.12\ldots(n-1)}X_n+C$$

where

$$C=M_d-b_{d1.23\ldots n}M_1-b_{d2.13\ldots n}M_2-\ldots\ldots b_{dn.12\ldots(n-1)}M_n \tag{39}$$

where d=differences such as $(a-b)$ and takes the place in all parts of formula (32) for the subscript o and

$$\sigma_{d.i}=b_{d1.23\ldots n}\sigma_1\sqrt{1-r_{II}}+b_{d2.13\ldots n}\sigma_2\sqrt{1-r_{2II}}$$
$$+\ldots\ldots+b_{dn.12\ldots(n-1)}\sigma_n\sqrt{1-r_{nN}} \tag{40}$$

where $\sigma_{d.i}$ is the standard error of a difference in an individual; $b_{d1.23\ldots n}$, $b_{d2.13\ldots n}$, etc., have the same meaning as in formula (39), and σ_1, σ_2, etc., are the standard deviations of the scores in the different predictive items.

The multiple correlation coefficient may be obtained by formula (37). In case the differential correlation coefficients $r_{(a-b)x}$, $r_{(a-b)y}$, etc., have not been calculated directly from the differences but indirectly through formula (38) from r_{ax}, r_{bx}, etc., we must obtain $\sigma_{(a-b)}$ and $M_{(a-b)}$ before the substitution can be made in formula (39). The formulas are

$$\sigma_{(a-b)} = \sqrt{\sigma_a^2 + \sigma_b^2 - 2r_{ab}\sigma_a\sigma_b} \tag{41}$$

and

$$M_{(a-b)} = M_a - M_b \tag{42}$$

Formula (41) is but a part of formula (38) and should be noted when the values for formula 38 are being calculated.

In some cases of placement of entering college students it may be desired by some to determine the present differences of ability existing in the individual student and assume from such existing differences what particular line of scholastic endeavor the student should enter upon. To do this most accurately the methods of differential diagnosis should be used. These methods have been gathered together by Segel (107). In general these methods are most valuable to be used after the student has been admitted to college. However, there is nothing to prevent the use of differential diagnosis in predictive problems if a predictive meaning can be assigned to the diagnosis.

Use of critical point.—As has been shown in the statistical procedures described in this chapter the prediction of a success score gives us a figure which is correct statistically within certain limits, assuming that the prediction is taking place with students of the same general kind upon which the original equation was determined. In some universities there is a rule that a certain grade standard must be maintained during the college course in order to insure graduation. For instance, in one university using the grading system A, B, C, D, and E, where A is excellent, B is good, C is fair, D is barely passing, and E is failure, the rule was that an average of C should be maintained throughout the college course and that an average of at least C must be made in order to insure graduation. For this school then the basis for considering whether a student will be a success or not means whether or not he can maintain an average grade of C. Assume it is found that freshmen generally make poorer marks than seniors and that it is found that freshmen making an average grade of D.8 or 2.8 (where A is 5, B is 4, C is 3, D is 2, and E is 1) will balance up during the course of the 4 years of college to C.0 or 3.0. This university wants to allow entrance to students who have some degree of promise in making this grade of 2.8 or above. If a student's estimated grade on first-year work is exactly 2.8 does that make the student a good risk? Actually this means with a large number of students that half of them will attain this average of 2.8 but one-half of them will not. In order to be more certain that an individual student will attain the acceptable average the estimated predicted average must be higher. That is, the institution accepting students upon the basis of predictive measures of students must themselves set up a critical score which represents a certain percentage of students who will fail to make a standard. Of course, which students of several attaining a certain predicted average mark will fail cannot be foretold. By accepting students with lower predicted marks the institution will get more students as to total number who will maintain an average grade of 2.8 as freshmen, but it will also receive more students who will fail to attain this standard. We will describe how an institution can settle as to the percentage of passing and failing students it will admit for various critical scores which it might settle upon.

Let us assume that the university considers an average mark of 2.8 for the freshman year as the critical point above which the student is considered successful. Assume further the following data to have been found on a fairly large number of students in this institution:

Correlation between intelligence test K and first-year college marks $(r_{o1}) = .60$.
Correlation between general achievement test M and first-year college marks $(r_{o2}) = .63$.
Correlation between test K and test M $(r_{12}) = .41$.
Mean of first-year marks $(M_o) = 3.7$.
Mean of scores in test K $(M_1) = 155$.
Mean of scores in test M $(M_2) = 168$.
Standard deviation of first-year marks $(\sigma_o) = .7$.
Standard deviation of scores on test K $(\sigma_1) = 22$.
Standard deviation of scores on test M $(\sigma_2) = 36$.

The regression equation for the prediction of marks which will be actually obtained in this school calculated from this data is $\overline{X}_o = +.013\,X_1 + .009\,X_2 + .173$ and the standard error would be $\sigma_{o.12} = .476$ and the probable error (P.E.) of the score $= .321$.

This means that any score obtained by the use of the equation must be judged in the light of this standard error or probable error. If it is decided that 2.8 is the grade which must be reached the college must still decide upon what estimated grade basis it will admit its students. Suppose various students seeking admission make the following scores on the tests *K* and *M*:

	Test *K*	Test *M*
Student A	160	172
Student B	98	126
Student C	75	93

Substituting these values in the equation

$$\bar{X}_e = +.013\ X_1 + .009\ X_2 + .173$$

where X_1 are scores on test K and X_2 are scores on test M, the predicted scores are as follows:

$$\bar{X}_e = +.013\ (160)\ +.009\ (172)\ +.173 = 3.8$$
$$\bar{X}_e = +.013\ (98)\ +.009\ (126)\ +.173 = 2.6$$
$$\bar{X}_e = +.013\ (75)\ +.009\ (93)\ +.173 = 2.0$$

Now what are the chances in an individual case that the student will be successful in college if admitted? Consider the first predicted mark. The predicted mark of 3.8 is (3.8—2.8) or 1.0 points above the critical point. Dividing this 1.0 by the probable error .321 we obtain a quotient of 3.11. This means that the predicted mark of 3.8 is 3.11 probable errors above the critical point.

By referring this to table V (p. 41) we see that ±.3 probable error units represents 96 percent of the cases. Since this is at a point +3 probable error units the percent of cases below it would be represented by 98 percent. At the critical point there is a chance of success or failure of 50 percent. Therefore at +3 probable error units the chance of success would be 98 percent and the chance of failure 2 percent.

For the determination of the chances of success for many predicted marks it is best to construct a table from which the chances of success for various predicted marks can be read off. Table VII is such a table where 2.8 is considered the critical mark. Only the first column of this table would be changed in case a different critical mark is chosen. The relationship between the probable errors and the chances of failing or of being successful given in percentages is stable for a normal distribution. By adding or subtracting appropriate marks values from this critical point corresponding to the probable error values set down in the second column the table is completed. In table VII for each +.1 probable error a value of .0321 was added, since that is equal to .1 of the probable error as given in this example. Opposite +.1 the marks value of 2.80±.0321 or 2.83 is therefore given. Opposite +2 the marks value of 2.80 + 2 (.0321) or 2.86 is therefore likewise given. In a similar manner the other marks values in the table were calculated.

Returning now to the predicted success scores of the two other students, B and C, student B has a predicted mark of 2.6. According to table VII this is −.6 probable error units below the critical point and indicates a chance of 66 out of 100 of failing and a chance of 34 out of 100 of being successful. Student C has a predicted mark of 1.8. This is about 2½ P. E.'s below the critical mark and indicates that the chance of failure is 95 or 96 percent.

TABLE VII.—*Chances of failure and success for various predicted scores using 2.80 as the critical point expressed in percentages*

Mark	Probable error rating	Percent failing	Percent successful	Mark	Probable error rating	Percent failing	Percent successful
1	2	3	4	1	2	3	4
4.08	+4.0	0.4	99.6	2.77	−0.1	52.7	47.3
3.92	+3.5	.9	99.1	2.74	− .2	55.4	44.6
3.76	+3.0	2.2	97.8	2.70	− .3	58.0	42.0
3.60	+2.5	4.6	95.4	2.67	− .4	60.6	39.4
3.44	+2.0	8.9	91.1	2.64	− .5	63.2	36.8
3.41	+1.9	10.0	90.0	2.61	− .6	65.7	34.3
3.38	+1.8	11.2	88.8	2.58	− .7	68.2	31.8
3.35	+1.7	12.6	87.4	2.54	− .8	70.5	29.5
3.31	+1.6	14.0	86.0	2.51	− .9	72.8	27.2
3.28	+1.5	15.6	84.4	2.48	−1.0	75.0	25.0
3.25	+1.4	17.3	82.7	2.45	−1.1	77.1	22.9
3.22	+1.3	19.0	81.0	2.41	−1.2	79.1	20.9
3.19	+1.2	20.9	79.1	2.38	−1.3	81.0	19.0
3.15	+1.1	22.9	77.1	2.35	−1.4	82.7	17.3
3.12	+1.0	25.0	75.0	2.32	−1.5	84.4	15.6
3.09	+ .9	27.2	72.8	2.29	−1.6	86.0	14.0
3.06	+ .8	29.5	70.5	2.25	−1.7	87.4	12.6
3.02	+ .7	31.8	68.2	2.22	−1.8	88.8	11.2
2.99	+ .6	34.3	65.7	2.19	−1.9	90.0	10.0
2.96	+ .5	36.8	63.2	2.16	−2.0	91.1	8.9
2.93	+ .4	39.4	60.6	2.00	−2.5	95.4	4.6
2.90	+ .3	42.0	58.0	1.84	−3.0	97.8	2.2
2.86	+ .2	44.6	55.4	1.68	−3.5	99.1	.9
2.83	+ .1	47.3	52.7	1.52	−4.0	99.6	.4
2.80	.0	50.0	50.0				

Where scholastic success is indicated by scores on certain examinations, tables like no. VII may still be built up for the purpose indicated if the critical examination score corresponding to our critical mark is known or can be ascertained.

Thus this university which conceives of 2.8 as being the critical point must still decide what predicted mark shall be the point at which entrance shall be denied. If the university does not like to have many failing students enter, it must put its point of admission somewhere above the critical point. It can if it wishes put this point so high that few students will fail. Or again, if numbers of students are wanted rather than quality, the point of admission may be placed below the critical point. Of course the values given in table VII may be used in guidance apart from strict admission. A university such as a State university may be compelled to admit all high-school graduates. In this case, predicted marks may still be used to advantage in the guidance of students before or at the time of admission.

The notation used in the formulas in this section follow in general that of Dunlap and Kurtz (38) who have attempted to establish a uniform system of notation. For details of statistical method and other formulas the following references may be consulted:

Brown and Thomson (20)
Dunlap and Kurtz (38)
Ezekiel (43)
Garrett (47)
Gilliland and Jordan (49)
Hines (57)
Holzinger (58)
Kelley (70) (71)
Otis (96)
Rietz (102)
Rugg (104)
Thurstone (125)

CHAPTER II: DIRECT RELATIONSHIPS FOUND BETWEEN SINGLE PREDICTIVE ITEMS AND VARIOUS CRITERIA OF COLLEGE SUCCESS

THE earliest studies of the prognosis of college success used the psychological tests developed by Wundt, Cattell, and others and were in general represented as to type by the cancelation test. Not much relationship was found between the results on these tests and college success. The World War gave an impetus to the practical study of college prediction through the construction of Army Alpha. Many studies of the value of Army Alpha in predicting college success were made. Other general mental ability tests for the direct purpose of testing students about to enter college were developed. During this early use and development of general mental ability tests at the college level, and at other levels, there was a great growth in achievement testing with new-type short-answer tests in the elementary and lower secondary level. Later this sort of material was also introduced in the upper secondary level, both to predict college success in general and success in specific college subjects. Thus at first the objective was the prediction of college success in general while later the placement and classification was added as an objective in testing entering college freshmen. A still later development—actually a refinement in statistical method—was the differential prediction method which makes it possible to show more clearly than before the comparative probability of success in different subjects.

The studies in the field will be summarized in the following sections. First will be presented the relationships found between single prediction items and some success criterion. Second, the relationships resulting from combinations of items will be presented. Third, the descriptions of some recent studies in differential prediction will be given.

In presenting these relationships the common expression for the relationship will be the correlation coefficient. In practically all cases of the zero order correlation coefficients it will be the correlation obtained through the Pearson product moment method as described in the section preceding.

By the fact that the correlation coefficient is used to show this prognostic relationship, it is not implied that this is the only method of showing the relationship. It is, however, the one most commonly used and one which is the best expression mathematically to show the relationship. The correlation coefficient is often misinterpreted, largely because of the false assumption by many that the figure is a percent. A correlation coefficient of .50 is not 50 percent and does not represent a prognostic capacity of 50 percent. Quotations will be given from only a few of the available references on this subject. Many studies of college prediction do not give the results in such a form as to be usable and many references are of a general character and do not contain evidence. It is thought that the studies from which the figures are taken represent the more important and valid work in the field. Cowley's Personnel Bibliographical Index (28) gives valuable leads to material in this field up to the year 1932.

In the tabulated material giving a summary of the correlation coefficients calculated between single prediction items and a criterion, the criterion is, in all cases except where noted differently, scholarship in college as measured by the marks given by instructors for work performed in regular course work. It is not specified in the briefly summarized material whether or not the first quarter, first half, first year, or the first 2 years' marks are used as a criterion. In all cases, however, it represents the marks which the students received in college from the time they entered. The beginning point is set but not the end point. Where correlations with several successive semesters were found in a study only the first one was given. These correlation coefficients represent, therefore, the relationship of predictive items with beginning scholarship in college. As might be expected the relationship between a predictive item and the first semester's scholarship is closer than its relationship with the second semester. There is evidence to show that the relationship between a predictive item diminishes for successive semesters or quarters in college. Anderson and Spencer (2) found evidence to this effect with the Yale Classification Test, a modified Army Alpha (table VIII).

TABLE VIII.—*Correlation coefficients showing the relationship between Army Alpha scores and marks in the first and second years of college* [1]

Group	First year		Second year	
	Correlation coefficient	Students	Correlation coefficient	Students
I	.41	366	.32	334
II	.38	564	.34	524
III	.39	723	.32	696

[1] Adapted from Anderson and Spencer (2).

In the main the students were the same in these two groups. This shows a decided difference in favor of the first year as against the second. Corroborating evidence on this point is given by Miner (87) who found correlations given in table IX between Army Alpha scores and marks in successive semesters. The number of cases is not stated.

TABLE IX.—*Correlation coefficients found between Army Alpha scores and scholarship in later successive semesters* [1]

1st	2d	3d	4th	5th	6th	7th	8th
.50	.37	.07	.32	.29	.26	.30	.25

[1] Adapted from Miner (87).

A summary of the relationships found between single prediction items and college scholarships are given in tables X to XVI, inclusive.

TABLE X.—*Prediction of general college scholarship using general mental tests*

Predictive item	Reference	Correlation coefficients
American Council on Education, psychological examination (various editions).	Boucher (11)	.39
	Condit (27)	.45
	Crane (29)	.27, .58 women, .50 men
	Hartson (54)	.50 women, .53 men
	Hopkins (59)	.53
	Remmers (100)	.45
	Remmers (101)	.57
	Segel (110)	.48
	Stalnaker (114)	.57
	Stoddard (115)	.62
	Thurstone and Thurstone (126).	.36, .48, .46, .45, .62, .40, .44, .54, .52, .40, .41, .47, .59, .32, .51, .45, .40, .49, .49, .56
	Whitney and Goodman (132).	.33

TABLE X.—*Prediction of general college scholarship using general mental tests*—Continued

Predictive item	Reference	Correlation coefficients
		Median .48
Anderson	Crawford (31)	.42
Army Alpha	Anderson (1)	.38
	Bridges (13)	.35
	Colvin (26)	.44
	De Camp (34)	.41
	Ernst (42)	.41
	Jordan (69)	.48
	Miner (87)	.50
	Stoddard (115)	.49
	Stone (117 and 118)	.44, .50
	Terman (122)	.52, .43, .31
	Toll (127)	.33
	Van Wagenen (130)	.50
		Median .44
Brown University examination	MacPhail (83)	.37
Miller mental ability test	Binnewies (8)	.43
	MacPhail (83)	.49
		Median .46
Minnesota College ability test	Williamson (133)	.42
Ohio State University psychological examination.	Blair (9)	.50, .49
	Bryns (21)	.36
	Edgerton and Toops (40).	.45
	Guiler (53)	.47
	Neuberg (92)	.47
		Median .47
Otis intelligence test	Binnewies (8)	.39
	Guiler (53)	.40
	Hill (56)	.20, .34
	Odell (95)	.38
	Thompson and Russell (124).	.43, .40
	Toll (127)	.38, .34
		Median .38
Scholastic aptitude test of the college entrance examination board.	Brown (19)	.46
	Crawford (30)	.70
		Median .58
Terman group test	Binnewies (8)	.49
	Guiler (53)	.52
	Terman (122)	.54
	Toll (127)	.28
		Median .505
Thorndike intelligence examination for high-school graduates (various years and additions).	Bolenbaugh and Proctor (10).	.45, .37
	Bridges (14)	.40
	Cleeton (22)	.52, .47, .48
		.52, Part I .40
	Cleeton (23)	.50
	Columbia University (25).	.65
	Grauer and Root (52)	.39
	MacPhail (83)	.41

TABLE X.—*Prediction of general college scholarship using general mental tests*—Continued

Predictive item	Reference	Correlation coefficients
Thorndike intelligence examination for high-school graduates (various years and addition).	Nelson (90)	.36, .37, .53, .27, .38, .53
	Root (103)	.51
	Symonds (120)	.42
	Tallman (121)	.51
	Terman (122)	.41, .60
	Wood (134)	.55
		Median .46
University of Washington intelligence test.	Dvorak and Salyer (37).	.37
Yale classification tests (modified Alpha).	Anderson and Spencer (2).	.41, .38, .39
		Median .39

TABLE XI.—*Prediction of general college scholarship using general achievement tests*

Predictive item	Reference	Correlation coefficients
Achievement in high-school work measured by the Pennsylvania study sponsored by the Carnegie Foundation, 1927.	Hill (56)	.52, .66
		Median .59
College entrance examination board	Beatley (6)	.50
	Brigham (16)	.46, .39
	Crawford (30)	.64
	Crawford and Burnham (32).	.43
		Median .46
Iowa high-school content examination	Cleeton (22)	.49, .49
		Median .49
Iowa placement tests. All tests except ET-1.	Stoddard (115)	.75
New York regents examainations	Gilkey (48)	.50
	Jones (67)	.70 women, .40 men
		Median .50
Sones-Harry high-school achievement test	Boucher (11)	.57

TABLE XII.—*Prediction of general college scholarship using tests of specific traits, aptitudes, or achievements*

Predictive item	Reference	Correlation coefficients
Aptitude test for elementary- and high-school teachers.	Frits (46)	.63
Iowa High School content examination, English section.	Segel (110)	.35
Iowa High School content examination, mathematics section.	do	.22
Iowa High School content examination, science section.	do	.13
Iowa High School content examination, social-studies section.	do	.19

TABLE XII.—*Prediction of general college scholarship using tests of specific traits, aptitudes, or achievements*—Continued

Predictive item	Reference	Correlation coefficients
Iowa placement examinations, chemistry aptitude.	Remmers (101) Stoddard (115)	.50 .52 Median .51
Iowa placement examinations, chemistry training.	----do----	.35
Iowa placement examinations, English aptitude.	Remmers (101) Stoddard (115)	.05 .48 Median .265
Iowa placement examinations, English training.	Brown (19) Remmers (101) Stoddard (115)	.44 .14 .37 Median .37
Iowa placement examinations, mathematics aptitude.	Remmers (101) Stoddard (115)	.12 .42 Median .27
Iowa placement examinations, mathematics training.	Brown (19) Stoddard (115)	.38 .35 Median .365
Iowa placement examinations, combined physics aptitude and training.	Dvorak and Salyer (37).	.55
Iowa placement examinations, combined mathematics aptitude and training.	----do----	.58
Ohio State University psychological examination, arithmetic-problems section.	Blair (9)	.40, .26 Median .33
Ohio State University psychological examination, similarities-opposites section.	----do----	.44, .44 Median .44
Ohio State University psychological examination, analogies section.	----do----	.42, .43 Median .425
Ohio State University psychological examination, number-progression section.	----do----	.21, .28 Median .245
Ohio State University psychological examination, paragraph-meaning section.	----do----	.45, .39 Median .42
Study-performance test, Ohio State University, difficult comprehension.	Toops, Adkins, Meyer (126).	.48
Study-performance test, Ohio State University, foreign words and phrases.	----do----	.47
Study-performance test, Ohio State University, disarranged sentences.	----do----	.37
Study-performance test, Ohio State University, study-habit questions.	----do----	.27
Study-performance test, Ohio State University, note-taking test.	----do----	.40
Thurstone vocational guidance tests (with engineering students) geometry.	Beatty and Cleeton (7).	.30
Thurstone vocational guidance tests, technical information.	----do----	.25
Thurstone vocational guidance tests, algebra.	----do----	.42
Thurstone vocational guidance tests, arithmetic.	----do----	.38
Thurstone vocational guidance tests, physics	----do----	.34

TABLE XIII.—*Prediction of scholarship in specific college subjects using general mental tests*

Predictive item	Specific criterion item	Reference	Correlation coefficients
American Council on Education psychological examination.	Marks in English	Boucher (11)	.37
		Segel (110)	.40
		Thurstone and Thurstone (126).	.52 .41
			Median .405
	Marks in languages—foreign languages.	Segel (110)	.42
		Thurstone and Thurstone (126).	.38
			Median .40
	French	Boucher (11)	.35
	German	do	.32
	Spanish	do	.18
	Marks in mathematics	do	.45
		Remmers (101)	.46
		Thurstone and Thurstone (126).	.35 .33
			Median .40
	Marks in social studies:		
	General	Segel (110)	.33
	History	Boucher (11)	.47
		Thurstone and Thurstone (126).	.35
			Median .41
	Economics	Boucher (11)	.40
	Political science	do	.20
	Marks in science:		
	Biology	Thurstone and Thurstone (126).	.38
	Chemistry	Remmers (101)	.38, .41
		Thurstone and Thurstone (126).	.55
			Median .41
	Physics	Boucher (11)	.55
Army Alpha	Marks in English	Jordan (69)	.52
		Stone (118)	.50
			Median .51
	Marks in languages:		
	French	Stone (118)	.30
	German	do	.36
	Modern languages	Jordan (69)	.31
	Spanish	Stone (118)	.12
	Marks in mathematics	Jordan (69)	.31
		Stone (118)	.38
			Median .295
	Marks in science:		
	General	Jordan (69)	.45
	Biology	Stone (118)	.22
	Chemistry	do	.31
	Physics	do	.44
	Marks in social studies:		
	History	Jordan (69)	.54
		Stone (118)	.31

TABLE XIII.—*Prediction of scholarship in specific college subjects using general mental tests*—Continued

Predictive item	Specific criterion item	Reference	Correlation coefficients
			Median .425
Brown University examination.	Marks in English	MacPhail (82)	.37
Otis group tests	Marks in languages:		
	French	Odell (95)	.32
		Tharp (123)	.32
			Median .32
	Latin	Odell (95)	.41
	Spanish	do	.27
	Marks in mathematics:		
	Algebra	do	.31
	Marks in science:		
	Biology	do	.20
	Botany	do	.42
	Zoology	do	.37
	Marks in social studies:		
	Economics	do	.28
	Political science	do	.33
Terman group test	Objective-test psychology	Nelson and Denny (91)	[1] .64
Thorndike intelligence examination for high-school graduates.	Marks in English	Lefever (80)	.26
		Root (103)	.36
		Stoddard (pt. I) (115)	.42, .36
			Median .36
	Marks in languages:		
	French	Lefever (80)	.19
		Root (103)	.40, .45
		Stoddard (pt. I) (115)	.25
			Median .325
	German	Root (103)	.50
	Spanish	Lefever (80)	.30
		Root (103)	.47
			Median .385
	Marks in mathematics	Lefever (80)	.24
		Root (103)	.39, .58, .51, .61
		Stoddard (pt. I) (115)	.23
			Median .45
	Marks in science:		
	Biology	Root (103)	.49, .52
			Median .505
	Chemistry	do	.43
		Lefever (80)	.24
			Median .335
	Physics	Root (103)	.50
	Zoology	Lefever (80)	.30

[1] Not used in the calculation of medians because the result was not obtained from marks.

TABLE XIV.—*Prediction of scholarship in specific college subjects, using general achievement tests*

Predictive item	Specific criterion item	Reference	Correlation coefficients
Iowa comprehension test.	Marks in English	Stoddard (115)	.47, .44 Median .455
	Marks in languages:		
	French	do	.31
	Marks in mathematics	do	.43
	Marks in science:		
	Chemistry	do	.33, .33 Median .33
Iowa high-school content examination.	Marks in English	Nelson (88)	.33
		Stoddard (115)	.39, .45 Median .39
	Marks in languages:		
	French	do	.42
	Marks in mathematics	do	.50
	Marks in science:		
	Chemistry	do	.24, .38 Median .31

TABLE XV.—*Prediction of scholarship in specific college subjects using tests of specific traits, aptitudes, or achievements*

AMERICAN COUNCIL ON EDUCATION PSYCHOLOGICAL EXAMINATION BY SECTIONS (all Segel and Gerberich (112))

Test	Marks in—		
	English	Foreign languages	Mathematics
Completion (I)	.41	.19	.33
Artificial language (II)	.54	.38	.43
Analogies (III)	.38	.25	.31
Arithmetic (IV)	.20	.14	.38
Opposites (V)	.51	.32	.36

Predictive item	Specific criterion item	Reference	Correlation coefficients
College entrance examination board, comprehensive chemistry.	Marks in chemistry	Crawford and Burnham (32).	.23 [1]
Comprehensive English.	Marks in English	Brigham (16)	.43
		Crawford and Burnham (32).	.30
		Whitman (131)	.39, average of 7 coefficients Median .39 [1]
Comprehensive French.	Marks in French	Brigham (16)	.44
		Brigham (15)	.28
		Whitman (131)	.47, average of 5 coefficients Median .47 [1]
Comprehensive mathematics.	Marks in mathematics	Brigham (16)	.32
		Crawford and Burnham (32).	.36

[1] See note p. 70.

TABLE XV.—*Prediction of scholarship in specific college subjects using tests of specific traits, aptitudes, or achievements*—Continued

AMERICAN COUNCIL ON EDUCATION PSYCHOLOGICAL EXAMINATION BY SECTIONS (all Segel and Gerberich (112))—Continued

Predictive item	Specific criterion item	Reference	Correlation coefficients
			Median .34[1]
Iowa High School content examination, English section.	Marks in English	Segel (110)	.36
		Stoddard (115)	.43
			Median .395[1]
	Marks in languages	Segel (110)	.29
	Marks in social studies	do	.35
Iowa High School content examination, mathematics.	Marks in English	do	.02
	Marks in languages	do	.28
	Marks in mathematics	Stoddard (115)	.28[1]
	Marks in social studies	Segel (110)	.10
Iowa High School content examination, science section.	Marks in English	do	.09
	Marks in languages	do	.17
	Marks in science: Chemistry	Stoddard (115)	.28[1]
	Marks in social studies	Segel (110)	.17
Iowa High School content examination, social-studies section.	Marks in English	do	.13
	Marks in languages	do	.05
	Marks in social studies	do	.23[1]
Iowa placement examinations, chemistry aptitude.	Marks in chemistry	Langlie (77)	.56, .56
		Miller (86)	.51, average of 29 coefficients
		Remmers (101)	.28, .55
		Stoddard (115)	.40
			Median .51[1]
Chemistry training.	Marks in chemistry	Langlie (77)	.43
		Miller (86)	.54, average of 18 coefficients
		Stoddard (115)	.34, .46
			Median .54[1]
English aptitude	Marks in English	Langlie (77)	.54, .46
		Miller (86)	.44, average of 24 coefficients
		Stoddard (115)	.42, .56
			Median .44[1]
English training	Marks in English	Langlie (77)	.63, .59
		Miller (86)	.40, average of 28 coefficients
		Nelson (88)	.48
		Stoddard (115)	.61, .50, .64
			Median .40[1]
Foreign language aptitude.	Marks in foreign languages.	Miller (86)	.48, average of 7 coefficients
	Marks in French	Tharp (123)	[2].47
	Test in French	do	[2].64
French training	Marks in French	Miller (86)	.52, average of 4 coefficients
		Stoddard (115)	.63

[1] See note p. 70.
[2] Not used in the calculation of medians because result was not obtained from marks.

TABLE XVI.—*Prediction of scholarship in specific college subjects using tests of specific traits, aptitudes, or achievements*—Continued

AMERICAN COUNCIL ON EDUCATION PSYCHOLOGICAL EXAMINATION BY SECTIONS (all Segel and Gerberich (112))—Continued

Predictive item	Specific criterion item	Reference	Correlation coefficients
Mathematics aptitude.	Marks in mathematics		Median [1].52
		Miller (86)	.44, average of 32 coefficients
		Stoddard (115)	.47
Mathematics training.	Marks in mathematics		Median [1].44
		Miller (86)	.48
		Remmers (101)	.66
		Stoddard (115)	.53, 60, 70
Physics aptitude	Marks in physics		Median [1].60
		Bear (5)	.50
		Miller (86)	.40, average of 7 coefficients
Physics training	Marks in physics		Median [1].40
		Miller (86)	[1].53, average of 4 coefficients
Nelson-Denny College reading test.	Marks in English	Nelson (88)	.37
New York Regents examinations, English.	do	Gilkey (48)	[1].49
Ancient languages	Marks in ancient languages.	do	[1].39
Mathematics	Marks in mathematics	do	[1].34
Social science	Marks in social studies	do	[1].34
Modern foreign languages.	Marks in modern foreign languages.	do	[1].32
Exact science	Marks in exact science	do	[1].15
Preliminary foods test.	Marks in foods	Brown (18)	[1].51
Sones-Harry High School achievement test:			
Language and literature section.	Marks in English	Boucher (11)	[1].45
Science section	Marks in chemistry	do	[1].59
	Marks in geology	do	[1].40
	Marks in physics	do	[1].60
Social studies section.	Marks in economics	do	[1].48
	Marks in history	do	[1].52
	Marks in political science.	do	[1].52

[1] See note p. 70.

SCHOLASTIC APTITUDE TEST OF THE COLLEGE ENTRANCE EXAMINATION BOARD BY SECTIONS (all Brigham, 17)

Section	Marks in—			
	Chemistry	English	Mathematics	Modern languages
Verbal	.39	.66	.33	.36
Mathematics	.38	.28	[r].59	.26

[r] See note p. 70.

STRONG VOCATIONAL INTEREST BLANK (all, Segel and Brintle, 111)

Interest in—	Marks in—			
	English	History	Languages	Mathematics and science
Engineering	−.10	−.47	−.09	.14
Medicine	−.08	−.13	−.01	.06
Law	.04	−.10	.60	−.22
Life-insurance salesmanship	−.15	−.05	.01	−.22
Personnel management	−.14	.27	−.17	−.01
Purchasing agent	−.22	−.17	−.14	−.02

THORNDIKE INTELLIGENCE EXAMINATION FOR HIGH-SCHOOL GRADUATES BY SECTIONS DIVIDED INTO LINGUISTIC, MATHEMATICAL, AND READING MATERIAL (all Lefever, 80)

Marks in—	Linguistic ability	Mathematical ability	Reading comprehension
Botany	.18	.23	.27
Chemistry	.12	.51	−.06
Economics	.21	.12	.27
English	.31	.19	.26
French	.21	.29	.16
History	.39	.26	.32
Mathematics	.09	[r].42	.11
Sociology	.34	.01	.25
Spanish	.38	.20	.29
Zoology	.17	.09	.22

TABLE XVI.—*Prediction of general college scholarship using average high-school marks* [1]

References	Correlation coefficients
Blair (9)	.66, .44, .62, .58
Brammell (12)	.52, .53
Cocking and Holy (24)	.55
Columbia University (25)	.29, 45, 47, .38, .61
Crawford (30)	.61
Edds and McCall (39)	.65
Edgerton and Toops (40)	.44
Goldthorpe (50)	.62
Hawks (55)	.66, .77, .72, .64, .69
Jones (67)	.48, women, .45, men
Lauer and Evans (78)	.63, women, .45, men
Lincoln (81)	.49
Odell (94)	.68, .58
Odell (95)	.55
Pierson and Netteis (97)	.54
Potthoff (98)	.52
Proctor (99)	.60, .62
Scates (105)	.41, .52
Seashore (106)	.61
Symonds (119)	.35
Symonds (120)	.55, .39
Terman (122)	.39
Williamson (133)	.53, .63, .54, .54, .54, 69, .69
	.65

[1] Median .55.

The median coefficients for this data grouped into six categories on the basis of kinds of predictive items and college success criteria using the median coefficient or single coefficients for each test are as follows (marks used in all cases as a criterion of college success):

TABLE XVII.—*Summary table of predictive relationships*

GENERAL COLLEGE SCHOLARSHIP

General mental tests	General achievement tests	Tests of specific traits, aptitudes, or achievements
.44	.545	.37

TABLE XVII.—*Summary table of predictive relationships*—Continued

SCHOLARSHIP IN SPECIFIC COLLEGE SUBJECTS OR SUBJECT GROUPS

	General mental tests	General achievement tests	Tests of specific traits, aptitudes, or achievements [1]
Foreign language	.355	....	.355
French	.322	.365	.47
German	.36	....	....
Latin	.41	....	....
Spanish	.225	....	....
English	.387	.422	.42
Mathematics	.355	.465	.42
Science	.45	....	....
Exact science	....	....	.15
Biology	.30	....	....
Botany	.42	....	....
Chemistry	.335	.32	.51
Physics	.50	....	.53
Zoology	.335	....	....
Geology	....	....	.40
Social studies	.33	....	.285
Economics	.34	....	.48
History	.417	....	.52
Political science	.265	....	.52

[1] Only tests which may be assumed to be prognostic of success in each of the subjects involved are summarized in this column. The test results used have been marked in the detailed data given previously in this section.

GENERAL COLLEGE SCHOLARSHIP AND AVERAGE HIGH-SCHOOL MARKS

Median .55

Referring to the summarization of the data on correlation coefficients it is seen that general achievement tests at the end of the high-school course are more prognostic of general college scholarship than general mental tests whereas individual tests of specific traits, aptitudes, and achievements are lowest of all for this purpose. The coefficients are .535, .44, and .367, respectively. The median correlation found between average high-school marks and general college scholarship is .55, which is slightly higher than the highest of the three median coefficients found by the use of single tests. Other writers (Douglass, 35 and Odell, 95) have noted the fact that the coefficient of correlation when college success is predicted from average high-school marks is higher than the corresponding coefficient obtained with general mental tests. When general achievment tests are used the difference in favor of average high-school marks is not great. There is also a factor in this situation which should be taken into account before making any general conclusions. It is this that predictive correlation coefficients using average high-school marks are more variable than coefficients obtained through the use of general scholastic aptitude or general-achievement tests. The results from these different schools are more stable when test results are considered. Just what marks from any one high school means is not known accurately since the correlation coefficients for marks for any high school or for the entrants to any college may very likely be anywhere from .45 to .65 and can be much less or somewhat more than these figures. As high-school instructors become more proficient in assigning marks these differences in coefficients will to some extent disappear.

The general trends of these coefficients of correlation are important. These show that for predicting general college scholarship the best tests are those testing general achievement and that for predicting scholarship in specific college subjects, tests of specific aptitudes or achievements are the best.

CHAPTER III: DIRECT RELATIONSHIPS FOUND BETWEEN COMBINATIONS OF PREDICTIVE ITEMS AND VARIOUS CRITERIA OF COLLEGE SUCCESS

THIS section will present the results of studies where two or more items have been combined through the medium of the regression equation to predict success in a college subject or group of college subjects. The multiple correlation coefficient expresses the efficiency of the resulting relationship. We will give first, a table of results in which the criterion of success was general college scholarship (table XVIII) and second, the results when predicting attainment in particular college subjects.

TABLE XVIII.—*Prediction of general college scholarship using a combination of predictive items*

Predictive item combination	Reference	Correlation coefficient
(a) Ohio psychological examination (b) High-school marks average	Blair (9)	.69, .56
(a) Ohio psychological examination (b) Specific high-school subject marks	Blair (9)	.70, .64
(a) High-school English marks (b) High-school natural science marks (c) High-school mathematics marks (d) University of Washington intelligence tests (e) Iowa mathematics aptitude and training tests (f) Iowa physics aptitude and training tests	Dvorak and Salyer (37)	.68
(a) College entrance board examination (b) Average high-school marks (c) Scholastic aptitude test (d) Age at entrance	Crawford (30)	.74
(a) Average high-school marks (b) Otis group intelligence test (c) Cross English test	Edds and McCall (39)	.81
(a) Average high-school marks (b) American council psychological examination	Douglass (35)	.63
(a) Average high-school marks (b) Ohio State University intelligence examination (c) Ohio study performance test	Hartson (54)	.75
(a) Rank in high-school class (b) Minnesota intelligence examination	Johnston (62) (63) (64) (65)	.67

TABLE XVIII.—*Prediction of general college scholarship using a combination of predictive items*—Continued

Predictive item combination	Reference	Correlation coefficient
(a) Average high-school marks (b) Otis group intelligence test	Jordan (68)	.58
(a) Average high-school marks (b) American council psychological examination	May (85)	.63
(a) Ohio University intelligence examination (b) Average high-school marks	Newberg (92)	.67
(a) Average high-school marks (b) Otis intelligence test	Odell (95)	.58
(a) Terman group (b) Average high-school marks (c) Character ratings	Pierson and Nettels (97)	.65
(a) Average high-school marks (b) Thorndike intelligence test	Proctor (99)	.68
(a) Average high-school marks (b) American council psychological examination	Symonds (119)	.59
(a) Average high-school marks (b) New York regents examination (c) Thorndike intelligence examination	Wood (134)	.66

Odell (95) determined that there was little difference as to predictive coefficients obtained by using point scores (raw scores) on the Otis by using the I.Q. calculated from the point score. Odell calculated zero order coefficients and multiple correlations between various high-school subjects, average high-school marks, and point score on the Otis self-administering test of mental ability, higher examination, on the one hand, and marks in particular college subjects on the other. He found very little differential predictive value in the marks in various high-school subjects and not much of an increase in the multiple correlation coefficient using 2 or 3 factors over the simple zero order correlation coefficient. Table XIX gives some of Odell's results.

TABLE XIX.—*Relationship between the combination of certain high-school records and scholarship in certain college freshmen subjects*[1]

High-school predictive items (all average of marks mentioned)	Subject in which scholarship was predicted	Multiple correlation coefficient
(a) Mathematics (b) Average	Algebra	.53
(a) Chemistry (b) Science (c) Intelligence point score	Chemistry	.53
(a) Average (b) History (c) Intelligence point score	Economics	.43
(a) Average (b) French (c) English (d) Intelligence point score	French	.60
(a) Geometry (b) Mathematics (c) Average or intelligence point score	Geometry	35
(a) English (b) Foreign language (c) Intelligence point score	German	.50
(a) Average (b) History (c) Intelligence point score	History	.46
(a) English (b) Intelligence point score (c) Average or Latin	Rhetoric	.52

[1] Adapted from Odell, (95).

Stoddard (115) using the Iowa high-school content examination, the Iowa comprehension, part I of the Thorndike Intelligence Examination for High-School Graduates, and the Iowa placement tests found multiple correlations higher than those found by Odell for most of the college subjects upon which he obtained data. His results are given in brief form in table XX.

TABLE XX.—*Multiple correlation coefficients between certain test results and marks in particular college subjects*[1]

Tests	Criterion	Multiple correlation coefficient
(a) Iowa high-school content (b) Iowa comprehension (c) Thorndike, pt. I (d) Iowa placement, French training	French	.76

[1] From Stoddard, (115).

TABLE XX.—*Multiple correlation coefficients between certain test results and marks in particular college subjects*—Continued

Tests	Criterion	Multiple correlation coefficient
(a) Iowa high-school content (b) Iowa comprehension (c) Thorndike, pt. I (d) Iowa placement, chemistry training	Chemistry	.44
(a) Iowa high-school content (b) Iowa comprehension (c) Thorndike, pt. I (d) Iowa placement, chemistry placement	Chemistry	.46
(a) Iowa high-school content (b) Iowa comprehension (c) Thorndike, pt. I (d) Iowa placement, English training	English	.64
(a) Iowa high-school content (b) Iowa comprehension (c) Thorndike, pt. I (d) Iowa placement, English aptitude	English	.52
(a) Iowa high-school content, English section. (b) Iowa placement, English training	English	.63
(a) Iowa high-school content (b) Iowa comprehension (c) Thorndike, pt. I (d) Iowa placement, mathematics training	Mathematics	.60
(a) Iowa high-school content, mathematics section. (b) Iowa placement, mathematics training.	Mathematics	.63

In addition to the results cited, the work of Mann (84) should be mentioned. Using the Iowa placement tests, CA, EA, ET, MA, and MT, and the School of Mines Engineering drawing placement examination he finds—

The correlations we have secured between raw test scores and course grades have been good, ranging from .626 to .810, according to the course and department. Correlations between averages for all tests and first semester scholarship range upward from .750 for various classes. Correlation between first and second semester grades is .940, which indicates rather a high reliability for the grades being awarded.

It will be noted that coefficients using a combination of items are higher than those given for single predictive items as given in part II, chapter II. This is an important point.

CHAPTER IV : STUDIES OF DIFFERENTIAL PREDICTION

IN THE preceding chapter the results of studies of direct prediction were given in each case in which one or more factors were used in the prediction of success in some one criterion. This criterion has been either the average achievement in all college subjects or the achievement in some particular subject or closely allied group of subjects. In the problem of placement and guidance, after a student has been admitted to college we are interested mainly in how well the student will do in individual subjects in comparison one with the other.

At first glance it would seem that if the direct prediction from two subjects is at hand, it would be easy to determine in the case of any one individual the subject in which he would do best. This, however, is not true with the present methods of evaluating achievement in college because the marks given in the different departments vary considerably in their central tendency and also because the capacities of the students in different departments differ materially. This has been shown by Eells (41) for one university. These variations by departments are not accidental variations. Kuenzel and Toops (75) determined that the level of the intelligence of the students in different divisions of college study was comparatively stable over a period of years.

The most accurate comparison between predicted success in two or more subjects must take into account these variations both as to variations in marks and in the ability of the students. This more accurate method has been called differential prediction to distinguish it from direct prediction. It must be noted, however, that the comparison of the direct predictions on two or more different subjects is after all also differential prediction even though the accuracy in this case is undetermined.

In order to show more clearly the procedures involved in using the more accurate methods of differential prediction, there follows a rather full description of a comprehensive investigation made by Segel (109) in 1932.

The tests selected were those which seemed to have possibilities for the differentiation of success in different college subject groups. These tests were: Van Wagenen Reading History Scale A; Cross English Test A; Illinois Standardized Algebra Test; Standard Interest Report Blank (also called the Cowding Interest Questionnaire);[1] and Wyman Association Test for Interests. The Wyman Association Test for Interests was omitted from consideration early in the investigation because the reliability coefficient was found to be about .00. This does not impugn the value of the test for the level for which it was constructed, i.e., the upper elementary school level.

The tests were given to 200 graduating seniors (class of 1925, Polytechnic High School, Long Beach, Calif.). The records made by these students in their college work were obtained insofar as was possible. The records for the academic years 1925–26 and 1926–27 were asked for and used. College records for 97 students were found to be usable in this study out of 115 reports on students. It is believed that the reports on more than 95 percent of the students enrolling in colleges during the 2 years aforementioned were obtained from the original group of 200 that were tested. The college records were analyzed for subjects which seem to fall into natural ability groups.

After considerable preliminary investigation in the differentiation of various college groups the following were finally determined upon as giving the best possibilities: English, languages, economics, history, biological science, and physical science.

The language group included all foreign languages. It was composed mostly, however, of French, Spanish, and German. The other subject groups are self-explanatory. From the marks in the six subject groups the reliabilities of and intercorrelations between the groups were calculated. The raw data for this work were given in the following form:

Student	Languages	History	Economics	Biological science	Physical science	English
E. B.	3 CCC	3.5 CD	3 C	3 C	3.3 DCC	2.7 CCBC

[1] This questionnaire was later revised and extended in form by Dr. Edward K. Strong, Jr.

Each letter is the mark in some particular subject in that subject group. To find mean marks the system was adopted where A equals 1, B equals 2, C equals 3, D equals 4, and all failure marks or F's equal 5. The reliability of the single marks in each subject is given in table I. All the coefficients of correlation reported in table I and in the tables following are based on various sized populations taken in all cases from the records of the 97 students for whom we have college records. In some cases, overlapping populations have been used where identical populations would have been more desirable because they would have made certain correlations higher. For most of the tables the range of P.E.'s is given.

Table XXII gives the intercorrelations between the college subject groups using individual marks. All the correlations reported in tables XXI and XXII have been calculated by the Pearson product moment method, corrected for broad categories. (Kelley (71), p. 167.)

The data were now available for making a good estimate of the existence of differential groups of college subjects among the particular ones investigated. This was done by the application of the formula for the correction for attenuation of the correlations reported in table XXII. These correlations are reported in table XXIII.

These correlations represent estimates of the true correlation as to success between these college subject groups when measured by the individual marks since they are correlations resulting when the inadequacies of the measuring instrument have been taken into account and allowed for.

TABLE XXI.—*Reliabilities of the individual marks in college subject groups* [1]

Languages	.664
Economics	.571
History	.370
Biological sciences	.606
Physical sciences	.519
English	.406

The P.E.'s for this table have a range of .02 to .06.

[1] From Segel (109).

TABLE XXII.—*Intercorrelation coefficients between college subject groups using individual marks* [1]

Subject	Economics	History	Biological science	Physical science	English
Languages	.217	.523	.392	.279	.517
Economics		.296	.436	.546	.187
History			.347	.165	.327
Biological science				.415	.196
Physical science					.320

[1] From Segel (109).

The P.E.'s for this table have a range of .02 to .05.

TABLE XXIII.—*Intercorrelation coefficients corrected for attenuation* [1]

Subject	Economics	History	Biological science	Physical science	English
Languages	.352	1.054	.618	.475	.996
Economics		.643	.741	1.004	.388
History			.734	.377	.845
Biological science				.741	.395
Physical science					.697

[1] From Segel (109).

These correlations corrected for attenuation differ considerably. Any pair of college subject groups that produce a correlation near 1.00 indicates that the abilities involved in the two groups were nearly the same and that therefore no differentiation can be expected. From a consideration of the table the following pairs of subject groups were set down as differentiable: Languages and economics, languages and biological science, languages and physical science, economics and history, economics and biological science, economics and English, history and biological science, history and physical science, biological science and physical science, biological science and English, and physical science and English.

The two equations which are applicable are (39) and (40) given in part II, chapter I. These are:

$$\overline{X}_d = b_{d1\cdot 23\cdots\cdots\cdots n}X_1 + b_{d2\cdot 13\cdots\cdots\cdots n}X_2 \cdots\cdots + b_{dn\cdot 12\cdots\cdots\cdots(n-1)}X_n + C$$

and

$$C = M_d - b_{d1\cdot 23\cdots\cdots\cdots n}M_1 - b_{d2\cdot 13\cdots\cdots\cdots n} - b_{dn\cdot 12\cdots\cdots\cdots(n-1)}M_n$$

and

$$\sigma_{d\cdot i} = b_{d1\cdot 23\cdots\cdots\cdots n}\,\sigma_1\sqrt{1-r_{II}} + b_{d2\cdot 13\cdots\cdots\cdots n}\,\sigma_2\sqrt{1-r_{2II}} + \cdots\cdots b_{dn\cdot 12\cdots\cdots\cdots(n-1)}\,\sigma_n\sqrt{(1-r_{nN})}$$

The zero order correlations with the criteria, i.e., the correlations of the test and questionnaire scores with the differences in marks in the pairs of subject groups selected are given in table XXIV. In order to proceed with the calculations it was also necessary to obtain the intercorrelations between the test items and the reliabilities of the test items. The reliabilities calculated from these data cannot be compared directly to reliabilities published for these tests and questionnaires since this group is a somewhat selected group. In most cases our reliabilities will be found to be lower. These reliabilities were found by splitting the tests in half and using the Spearman-Brown formula. The population in all cases for this purpose was 50. The reliabilities are given in table XXV. The intercorrelations are given in table XXVI.

Next were found the regression coefficients or constants and the multiple correlation coefficients for each of the college subject group pairs for which a difference was to be estimated. All items for the calculation of constants were omitted where the correlation with the criterion was less than .10 since, in general, such low correlations will result in little or no value in a regression equation. The constants and the multiple correlation coefficients were calculated by the method of successive approximations. (Kelley (71), p. 302.)

TABLE XXIV.—*Correlations of different items with certain criteria of college success* [1]

Criteria	Reading	English	Algebra	Medicine	Law	Engineering
Difference in marks received in—						
Language, physics	.037	—.333	.195	—.219	—.236	.388
Language, economics	.170	—.451	.147	—.140	—.122	.299
Language, biological science	.155	.010	.150	—.031	—.287	.098
Economics, history	—.478	.152	—.316	.221	—.023	—.066
Economics, English	—.344	.067	—.029	—.006	.209	.000
History, physical science	.325	.182	.186	.223	—.277	.232
Biological science, English	—.046	.022	.122	—.257	.316	—.163
Physical science, English	—.179	.122	.000	—.032	.357	—.061
Economics, biological science	—.208	.151	—.042	.317	—.183	.000
History, biollogical science	.325	.142	.075	.162	—.267	—.070
Biological science, physical science	.000	—.065	—.074	—.346	.214	.253

[1] From Segel (109).

The P.E.'s for the correlations of greater than ±.100 for this table range from .08 to .10.

TABLE XXV.—*Reliability of test items used in the prediction study*[1]

Test:	
Van Wagenen reading	.769
Cross English	.642
Illinois algebra	.953
Cowdery questionnaire medicine	.431
Cowdery questionnaire law	.318
Cowdery questionnaire engineering	.465

[1] From Segel (109).

TABLE XXVI.—*Intercorrelation coefficients between test items*[1]

Test	Cross English	Illinois algebra	Cowdery medicine	Cowdery law	Cowdery engineering
Van Wagenen reading	.330	.204	−.012	−.015	−.007
Cross English		.062	.170	.060	.106
Illinois algebra			−.164	−.165	.287
Cowdery medicine				−.244	−.143
Cowdery law					−.550

[1] From Segel (109).

The P.E.'s for this table range from .05 to .07.

This method gives constants and multiple correlation coefficients to any degree of accuracy required. We have calculated our multiple regression coefficients to within .001 of the optimum value. The constants are therefore as accurate as necessary for the work involved. The pair economics and English have only two correlations with test and questionnaire items which are more than .10 and therefore we have omitted this pair in our further study. The constants and the multiple regression coefficients are given in table XXVII. The multiple correlation coefficients are in general very superior to the zero order coefficients of correlation between the various test items and the criteria. This is due to the low intercorrelations between the test items. If these intercorrelations were not low our multiple correlation coefficients would not be nearly as significant as they are now.

TABLE XXVII.—*Regression coefficients and multiple correlation coefficients for prediction pairs*[1]

Item	Languages and physical science	Languages and economics	Languages and biological science	Economics and history
Reading test	----------	.394	.134	−.536
English test	−.369	−.663	.087	.323
Algebra test	.069	.014	----------	−.205
Cowdery medicine	−.086	.094	−.271	.126
Cowdery law	.011	.221	----------	----------
Cowdery engineering	.428	.481	----------	----------
Multiple correlation coefficient resulting	.555	.694	.333	.632

[1] From Segel (109).

TABLE XXVII.—*Regression coefficients and multiple correlation coefficients for prediction pairs*—Continued

Items	History and physical science	Biological science and English	Physical science and English
Reading test	.254		−.203
English test	.172		.142
Algebra test		.169	
Cowdery medicine	−.342	−.202	
Cowdery law	−.378	.214	.345
Cowdery engineering	−.037	−.112	
Multiple correlation coefficient resulting	.548	.400	.427
	Economics and biological science	History and biological science	Biological science and physical science
Reading test	.268	.304	
English test	.206	.038	
Algebra test			
Cowdery medicine	.245	.098	−.203
Cowdery law	−.133	−.245	.344
Cowdery engineering			.414
Multiple correlation coefficient resulting	.438	.432	.480

We have chosen a multiple correlation coefficient of .500 as a critical coefficient for determining whether or not we should set up prediction equations. This is of course arbitrary although it is a common practice to consider coefficients of less than .50 to indicate prediction efficiency too low to be of practical value. Even higher correlations may not bring about efficiency of a practical sort. It depends upon other factors in the situation. Since we are primarily concerned in establishing a method, no further discussion of this point will be carried on. There are four pairs of groups that have multiple correlation coefficients above .50. The multiple regression equations for these four groups and the probable errors of the resulting predictive values have been calculated by the use of formulas 39 and 40.

The equation for predicting a difference in achievement in marks between economics and history in favor of history is as follows:

$$X_{d(Ec-H)} = -.474X_1 + .192X_2 - .138X_3 + .015X_4 + 24.5$$

$$\sigma_{d.i} = 2.53$$

$$P.E._{d.i} = 1.71$$

Similarly the predictive equations and the probable errors for the other three pairs of college subject groups were found to be as follows:

Languages and economics in favor of economics

$$\bar{X}_{d(L-Pe)}=.391X_1-.371X_2+.010X_3+.012X_4+.018X_5+.044X_6+26.4$$
$$\sigma_{d.i}=4.29$$
$$P.E._{d.i}=2.89$$

Languages and physical science in favor of physical science

$$\bar{X}_{d(L-P)}=-.161X_2+.037X_3-.008X_4+.001X_5+.029X_6+36.3$$
$$\sigma_{d.i}=2.19$$
$$P.E._{d.i}=1.48$$

History and physical science in favor of physical science

$$\bar{X}_{d(H-P)}=.187X_1+.087X_2-.031X_4-.029X_5-.003X_6-.2$$
$$\sigma_{d.i}=2.65$$
$$P.E._{d.i}=1.79$$

The practical application of the differential prediction equations.—The application of the four prediction equations obtained in the preceding chapter will be illustrated. In order to lessen confusion in the arithmetical computation and for ease of interpretation, the differences in marks between any two subject groups have been transformed so that the final predictive value is positive. This transformation scale is given in table XXVIII. In the calculations of this study these transformed values have been used. For this illustration the equations have been applied to the original records used in this study. For this purpose a sample of the students were taken who had complete test and questionnaire data and at least one actual difference between subject groups out of the four differences predicted. Fifty-six of our students came under this classification. The use of the original records for this purpose causes a somewhat better prediction, no doubt, than would be the case if new records were to be used.

The equation for economics and history in favor of history is

$$\bar{X}_d \doteq -.474X_1+.192X_2-.138X_3+.015X_4+24.5$$

TABLE XXVIII.—*Table for changing differences in marks between college subject groups to a positive value* [1]

Difference	Transformed score	Difference	Transformed score
−1.3	1	.5	19
−1.2	2	.6	20
−1.1	3	.7	21
−1.0	4	.8	22
−.9	5	.9	23
−.8	6	1.0	24
−.7	7	1.1	25
−.6	8	1.2	26
−.5	9	1.3	27
−.4	10	1.4	28
−.3	11	1.5	29
−.2	12	1.6	30
−.1	13	1.7	31
.0	14	1.8	32
.1	15	1.9	33
.2	16	2.0	34
.3	17	etc.	
.4	18		

[1] From Segel (109).

Substituting the values from the tests for the first student, V.A., in the equation gives us the following:

$$-.474\ (71)+.192\ (107)-.138\ (51)\ +.015\ (160)+24.5=\bar{X}_d$$
$$\text{or}\ -33.7+20.5-.7+2.4+24.5=\bar{X}_d=13.0$$

Similarly, the values for the other students have been calculated for this equation and for the other three prediction equations. Table XXIX gives the values in the 4 prediction equations after substitution of the test scores for the first 4 cases of the 56 cases selected as aforementioned. The predicted score, i.e., a predicted difference is given in the last column of the table.

Since these predicted differences are regressed values they cannot be compared directly with the differences actually found. However, rank placement comparisons may be made. Comparisons for all 56 cases mentioned were made. The number of times the predicted and actual differences were above or below the mean and the number of times the predicted and actual scores were on opposite sides of the mean were counted. By chance alone the number of cases in each category would be 50 percent. It was found that about 70 percent of the cases for these four equations were correctly classified. This is in accordance with the relationship indicated by the multiple correlation coefficients.

TABLE XXIX.—*Table showing calculations for the first 4 cases on the first 2 prediction equations* [1]

ECONOMICS AND HISTORY

Student	$-.474X_1$	$+.192X_2$	$-.138X_3$	$+.015X_4$	$+24.5$	$=X_d$
V.A	−33.7	+20.5	−.7	+2.4	+24.5	=13.0
E.Bri	−43.6	+26.9	−1.1	+3.2	+24.5	=9.9
T.B	−38.9	+31.9	−1.0	+4.3	+24.5	=20.8
E.C	−43.1	+32.3	−2.6	+2.6	+24.5	=13.7

LANGUAGES AND ECONOMICS

Student	$.391X_1$	$-.371X_2$	$+.010X_3$	$+.012X_4$	$+.018X_5$	$+.044X_6$	$+26.4$	$=X_d$
V.A	27.8	−39.7	+0	+1.9	+5.8	+5.3	+26.4	=27.5
E.Bri	36.0	−51.9	+.1	+2.5	+6.1	+1.4	+26.4	=20.6
T.B	32.1	−61.6	+.1	+1.1	+.5	+4.1	+26.4	=2.7
E.C	35.6	−62.3	+.2	+2.1	+6.1	+4.2	+26.4	=12.3

[1] From Segel (109).

The method of success prediction for the first student of our 56 cases has been illustrated graphically.

The accuracy of the prediction made by the differential prediction method can be determined in an individual case by comparing the predicted difference with its probable error. As better measures are used this method of differential prediction will become more and more useful. With the prediction items used in this study many of the predictions are of value. Where there is room for doubt other corroborative evidence in the form of scholastic success, expressed interests, etc., can be used to give some accuracy to the guidance of a student.

The probable error given herein is an average value and is applicable to each one of the differences of the total population. It is to be noted that if we deliberately choose differences from a population because of some particular quality such as the picking out of especially large differences in one direction the probable error is too small or too large (Kelley (71), pp. 184–85).

The differential prediction for V.A., the first student on the list, is shown in the figure. This shows no reliable differentiation as between economics and history since the predicted score of 13 is only 1 point below the mean, and the probable error is 1.7. The differentiation in favor of economics as opposed to languages is fairly significant because the predicted score is 9.5 above the mean of 27.5 which is 3.3 times the P.E. There is a differentiation in favor of physical science over languages of 3.8 which is about 2½ times the P.E. of 1.48. This differentiation can be considered as indicating a possibility of being better in physical science than in languages. The differentiation of history and physical science in favor of history is quite marked since the predicted score is 7.4 below the mean which is more than 4 times the P.E., of 1.79. From a study of the pattern of prediction presented one can go further and indicate the

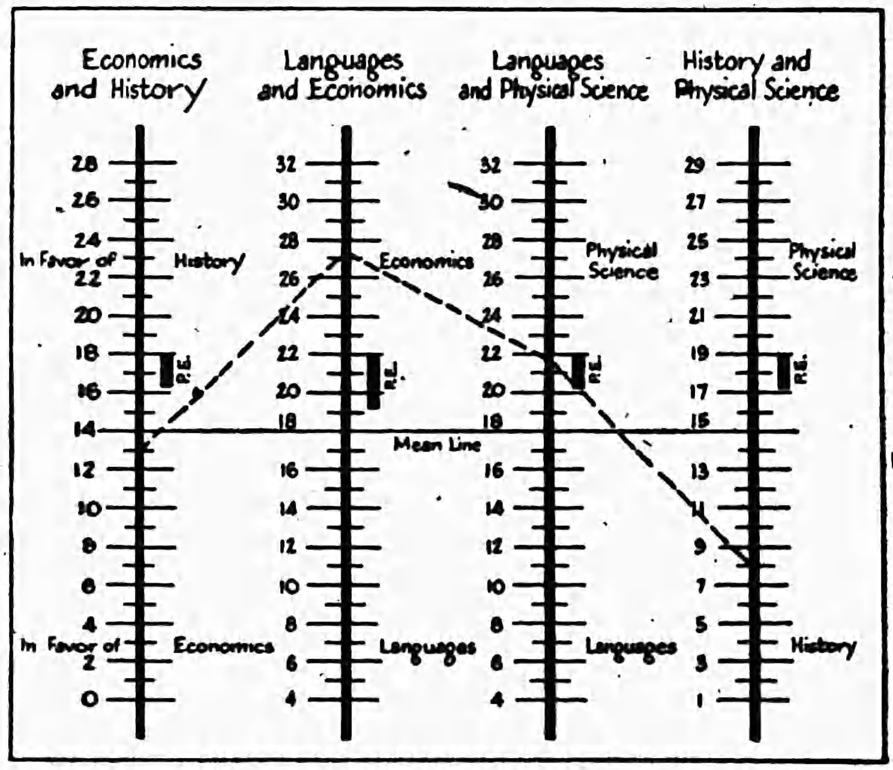


FIGURE 4.—Differential prediction for V. A.

probable arrangement of the four subjects as to success in them in a descending order as follows: Economics, history, physical science, and languages.

Another later study (Segel, 108) on the same level but using the Strong interest blank, the American Council on Education psychological examination, and the parts of the Iowa High School content examination was made. In this study the relationship between each of these tests and the differences between marks in certain subject groups were obtained. These are given in table XXX.

TABLE XXX.—*Zero order correlation coefficients between the scores on various prediction factors and differences between marks in certain college subject groups (differential prediction correlation coefficients)*[1]

Differences in marks between—	American council 1	Iowa English literature 2	Iowa mathematics 3	Iowa science 4	Iowa history and social science 5
1	2	3	4	5	6
English—Languages (62 cases)	+.23	+.36	+.12	+.14	+.21
Languages—Mathematics and science (68 cases)	-.09	-.27	-.37	-.26	-.09
Language—History (41 cases)	-.30	-.36	+.02	+.07	-.26
English—Mathematics and science (66 cases)	+.14	+.09	-.25	-.12	+.12
English—History (42 cases)	-.07	.00	+.14	+.21	-.05
Mathematics and science—History (44 cases)	-.21	-.09	+.39	+.33	-.17

Differences in marks between—	Strong engineering 6	Strong medicine 7	Strong law 8	Strong life insurance 9	Strong personnel management 10	Strong purchasing agent 11
1	7	8	9	10	11	12
English—Languages (62 cases)	-.01	-.07	+.04	-.16	-.07	-.08
Languages—Mathematics and science (68 cases)	-.23	-.07	+.22	+.23	-.06	-.12
Language—History (41 cases)	+.38	+.12	+.10	+.06	-.34	+.03
English—Mathematics and science (66 cases)	-.24	-.14	+.26	+.07	-.13	-.20
English—History (42 cases)	+.37	+.05	+.14	-.10	-.05	-.41
Mathematics and science—History (44 cases)	+.61	+.18	-.12	-.17	-.28	+.15

[1] From Segel (108).

The differential prediction coefficients in this study were obtained by the use of formula (38)

$$r_{(a-b)x} = \frac{r_{ax}\sigma_a - r_{bx}\sigma_b}{\sqrt{\sigma_a^2 + \sigma_b^2 - 2r_{ab}\sigma_a\sigma_b}}$$

since the direct correlations had been calculated for a study of direct prediction.

This table bears close examination. The results with the American Council test gives a differential meaning to its scores in favor of English (English—languages, +.23; English—mathematics and science, +.14) and history (languages—history, -.30; mathematics and science—history, -.21) as opposed to languages on the one hand and mathematics and science on the other, but does not show a differential as between English and history (English—history −.07).

The English literature test results give a differential in favor of English over languages (+.36); mathematics and science over languages (−.27); mathematics gives a differential prediction particularly well as between mathematics and science marks as opposed to languages, English, and history (−.37, −.25, +.39). Similarly each of these tests may be analyzed to discover its particular value in predicting differential success. The probable errors for these coefficients are substantial, so that the smaller coefficients cannot be considered too seriously.

Although many of these coefficients may be of sufficient size to be taken into account in theoretical discussions, each of them (with one possible exception) may be considered to have little value as an individual correlation coefficient to be used for purposes of guidance. An examination of table XXX indicates that the first criterion, English-languages had very few correlation coefficients of any magnitude.

This difference criterion was omitted from our further consideration. The other criteria seemed to have possibilities in this connection. The multiple correlation coefficients indicating the efficiency of predicting these five criteria were obtained using those factors in each case which had a significant relationship with the criterion. These results (given in table XXXI) show that these factors when combined do have important relationships to the criteria. It is to be remembered that these criteria consist of teachers' marks which in turn have been based on various marks and ratings, mostly subjective, made by the teacher on a scale of five points.

TABLE XXXI.—*Multiple correlation coefficients for the prediction of differences in college marks using various factors* [1]

Criterion (differences in marks)	Factors	Multiple correlation coefficient
Languages, mathematics, and science	1. Iowa English literature 2. Iowa mathematics 3. Iowa science 6. Strong engineering 8. Strong law 9. Strong life insurance	.49

[1] From Segal (108).

TABLE XXXI.—*Multiple correlation coefficients for the prediction of differences in college marks using various factors*—Continued

Criterion (difference in marks)	Factors	Multiple correlation coefficient
Languages, history	1. American Council 2. Iowa English literature 5. Iowa history and social science 6. Strong engineering 10. Strong personnel management	.58
English, mathematics, and science	3. Iowa mathematics 6. Strong engineering 8. Strong law 10. Strong purchasing agent	.40
English, history	4. Iowa science 6. Strong engineering 11. Strong purchasing agent	.48
Mathematics and science, history	1. American Council 3. Iowa mathematics 4. Iowa science 6. Strong engineering 10. Strong purchasing agent	.72

The American Council psychological examination was analyzed (Segel and Gerberich, 112) for its possible use in differentiating the ability of students to do work in the specific subjects of English, foreign languages, and mathematics. This was done by calculating differential correlation coefficients between scores on the various parts of the American Council test and differences between college marks in English, foreign languages, and mathematics. These were all found to be low. The correlations varied in value sufficiently to show that different parts of the American Council test measured somewhat different traits.

87924—35——7

BIBLIOGRAPHY

1. Anderson, J. E. Intelligence tests of Yale freshmen. School and society, 11: 417–20, Apr. 3, 1920.
2. ——— *and* Spencer, L. T. The predictive value of the Yale classification tests. School and society, 24: 305–12, Sept. 4, 1926.
3. Arlitt, A. H., *and* Hall, Margaret. Intelligence tests versus entrance examinations as a means of predicting success in college. Journal of applied psychology, 7: 330–38, December 1923.
4. Barr, A. S. The coefficient of correlation. (Editorial.) Journal of educational research, 23: 55–60, January 1931.
5. Bear, R. M. The predictive value of the Iowa physics aptitude placement test. Journal of applied psychology, 11: 381–84, October 1927.
6. Beatley, B. The relative standing of students in secondary school, in comprehensive entrance examinations and in college. School review, 30: 141–47, February 1922.
7. Beatty, J. D., *and* Cleeton, Glen U. Predicting achievement in college and after graduation. Personnel journal, 6: 344–51, February 1928.
8. Binnewies, W. G. Freshmen grades and mental tests. Educational administration and supervision, 9: 161–62, March 1923.
9. Blair, J. L. Significant factors in the prediction of the success of college freshmen. University of Chicago Libraries, Chicago, 1933.
10. Bolenbaugh, L., *and* Proctor, W. M. Relation of subjects taken in high school to success in college. Journal of educational research, 15: 87–92, February 1927.
11. Boucher, C. S. Some studies of freshmen admission at the University of Chicago. Mimeographed. 1932.
12. Brammell, P. Roy. Articulation of high school and college. Washington, D. C., National Survey of Secondary Education, Monograph, No. 10. Office of Education Bulletin 1932, No. 17.
13. Bridges, J. W. The correlation between college grades and the Alpha intelligence tests. Journal of educational psychology, 11: 361–67, October 1920.
14. ——— The value of intelligence tests in universities. School and society, 15: 295–303, Mar. 18, 1922.
15. Brigham, C. C. The prognosis of college work in French. *In* Prognosis tests in the modern foreign languages. New York, N. Y., Macmillan Co., 1929.
16. ——— Correlation of the examination of the College Entrance Examination Board with college standing. Journal of engineering education, 15: 653–89, May 1925.
17. ——— A study of error. New York, N.Y. College Entrance Examination Board, 1932.
18. Brown, Clara M. An experiment in sectioning. Journal of higher education, 1: 269–73, May 1930.

19. Brown, Ruth A. A comparison of data on freshmen entering the University of Michigan in the fall of 1927. Ann Arbor, Mich., University of Michigan, 1930. (University of Michigan Administrative Studies I, No. 1.)

20. Brown, William, *and* Thomson, C. H. The essentials of mental measurement. London, Cambridge university press, 1921.

21. Byrns, Ruth K. Scholastic aptitude and freshmen achievement. School and society, 35: 713–18, May 21, 1932.

22. Cleeton, Glen U. The predictive value of certain measures of ability in college freshmen. Journal of educational research, 15: 357–70, May 1927.

23. ——— Report on the Thorndike intelligence test. Pittsburgh, Pa., Department of Education and Psychology, Carnegie Institute of Technology, 1924.

24. Cocking, W. D., *and* Holy, T. C. Relation of intelligence scores to high-school and university marks. Educational research bulletin, 6: 383–84, Dec. 7, 1927.

25. Columbia University. Report on the use of intelligence examinations in Columbia University. 1927.

26. Colvin, S. S. The use of intelligence tests. Educational review, 62: 134–48, September 1921.

27. Condit, P. M. The prediction of scholastic success by means of classification examinations. Journal of educational research, 19: 331–35, May 1929.

28. Cowley, W. H. The personnel bibliographical index. Columbus, Ohio, Ohio State University, 1932.

29. Crane, E. The results of some psychological tests at Bryn Mawr College. School and society, 25:640–44, May 28, 1927.

30. Crawford, A. B. Forecasting freshmen achievement. School and society, 31:125–30, Jan. 25, 1930.

31. ——— Incentives to study. New Haven, Conn., Yale university press, 1929.

32. ——— *and* Burnham, Paul S. Entrance examinations and college achievement. School and society, 36:344–52 and 378–84, Sept. 10 and 17, 1932.

33. Cuff, Noel B. Prognosis and diagnosis of success in college. Journal of applied psychology, 14:612–19, December 1930.

34. De Camp, J. E. Studies in mental tests. School and society, 14:254–58, Oct. 1, 1921.

35. Douglass, Harl R. The relation of high-school preparation and certain other factors to academic success at the University of Oregon. University of Oregon Publications III, No. 1, September 1931.

36. ——— Relation of pattern of high-school credits to scholastic success in college. North Central Association quarterly, 6:283, December 1931.

37. Dvorak, A., *and* Salyer, R. C. Significance of entrance requirements for the engineering college at the University of Washington. Journal of engineering education (new series), 23:618–23, April 1933.

38. Dunlap, Jack W., *and* Kurtz, A. K. Handbook of statistical monographs, tables, and formulas. Yonkers, N.Y., World Book Co., 1932.

39. Edds, J. H., *and* McCall, W. M. Predicting the scholastic success of college freshmen. Journal of educational research, 27:127–30, October 1933.

40. Edgerton, H. A., *and* Toops, H. A. Academic Progress, Columbus, Ohio, Ohio State university press, 1929.

41. Eells, Walter C. University grading with special reference to intelligence tests. Doctor's thesis (unpublished). Stanford University library, 1927.

42. Ernst, John L. Psychological tests versus the first semester marks as a means of academic prediction. School and society, 18:419–40, Oct. 6, 1923.

43. Ezekiel, Mordecai. Methods of correlation analysis. New York, N.Y., John Wiley & Sons, 1930.

44. Frasier, Geo. W. *and* Heilman, J. D. Experiments in teacher-college administration, III: intelligence tests. Educational administration and supervision, 14:268–78, April 1928.

45. Freeman, Frank S. Predicting academic survival. Journal of educational research, 23:113–23, February 1931.

46. Fritz, R. A. Predicting college marks and teaching success for students in a teachers college. Journal of applied psychology, 17:439–46, August 1933.

47. Garrett, H. E. Statistics in psychology and education. New York, N.Y., Longmans, Green & Co., 1926.

48. Gilkey, Royal. The relation of success in certain subjects in high school to success in the same subjects in college. School review, 37:576–88, October 1929.

49. Gilliland, A. R. *and* Jordan, R. H. Educational measurements and the classroom teacher. New York, N.Y., Century Co., 1924.

50. Goldthorpe, J. H. The relative rank in high school and the first two years of the university. School and society, 30:130–34, July 27, 1929.

51. Gowen, J. W., *and* Gooch, M. The mental attainments of college students in relation to previous training. Journal of educational psychology, 16:547–68, November 1925.

52. Grauer, David, *and* Root, W. T. The Thorndike intelligence test and academic grades. Journal of applied psychology, 11:297–318, August 1927.

53. Guiler, W. S. The predictive value of group intelligence tests. Journal of educational research, 16:365–74, December 1927.

54. Hartson, L. D. The most valid combination of 23 tests for predicting freshmen scholarship at Oberlin College. Columbus, Ohio, Ohio College Association bulletin, No. 58, 1928.

55. Hawks, Lena J. Certain relationships between scholarship in high school and in college. Baltimore, Md., Johns Hopkins press, 1929.

56. Hill, H. D. The relative weight of certain factors in predicting academic success in college. Pennsylvania State studies in education in Abstracts of studies in education at Pennsylvania State College.

57. Hines, H. C. A guide to educational measurements. Boston, Mass., Houghton Mifflin Co., 1923.

58. Holzinger, Karl J. Statistical methods for students in education. Boston, Mass., Ginn & Co., 1928.

59. Hopkins, L. T. The reliability of the Wilkins prognosis test for predicting success in modern foreign languages. *In* Prognosis tests in modern foreign languages. New York, N.Y., Macmillan Co., 1929.

60. Hull Clark L. Aptitude testing. Yonkers, N.Y., World Book Co., 1928.

61. Jarvis, P. G. H. Mental tests as entrance requirements to State-supported institutions from the legal viewpoints. School and society, 21:280–81, Mar. 7, 1925.

62. Johnston, J. B. Predicting college success for the high-school senior. Vocational guidance magazine, 6:289–94, April 1928.

63. ——— Predicting success in college at the time of entrance. School and society, 23:82–88, Jan. 16, 1926.

64. ——— Predicting success or failure in college at the time of entrance. School and society, 19:82–88, Jan. 16, 1924.

65. ——— The selection of college students. *In* Problems of college education. Minneapolis, University of Minnesota press, 1928.

66. Jones, E. S. Comprehensive examinations in American colleges. New York, N.Y., Macmillan Co., 1933.

67. ——— Predictions from high-school performance. School and society, 27:339–40, Mar. 17, 1928.

68. Jordan, A. M. Correlations of four intelligence tests with the grades of students and with each other. Psychological Bulletin, 19:93–94, February 1922.

69. ——— Some results and correlations of Army Alpha tests. School and society, 11:354–58, May 20, 1920.

70. Kelley, Truman L. Interpretation of educational measurements, Yonkers, N.Y., World Book Co., 1927.

71. ——— Statistical method. New York, N.Y., Macmillan Co., 1924.

72. Kellogg, Chester, E. Relative value of intelligence tests and matriculation examinations as means of estimating probable success in college. School and society, 30:893–96, Dec. 28, 1929.

73. Kelly, F. J. Fusing of high school and college. Junior college journal, 2:413–20, May 1932.

74. Keys, Noel. The improvement of measurement through cumulative testing. New York, N.Y. Bureau of publications, Teachers college, Columbia university, 1928. (Contributions to education no. 321.)

75. Kuenzel, Myra W., *and* Toops, H. A. The relative stability of intelligence levels in colleges. School and society, 22:176–84, Aug. 8, 1925.
76. Kurani, H. A. Selecting the college student in America. New York, N.Y., Bureau of publications, Teachers college, Columbia university, 1931. (Contributions to education no. 503.)
77. Langlie, T. A. The Iowa placement examinations at the University of Minnesota. Journal of engineering education, 17:842–60, May 1927.
78. Lauer, Alvah R., *and* Evans, J. E. The relative predictive value of different high-school subjects on college grades. School and society, 31:159–60, Feb. 1, 1930.
79. Lee, J. Murray, *and* Segel, David. The utilization of data from simple or direct prediction in the development of regression equations for differential prediction. Journal of educational psychology, 24: 550–54, October 1933.
80. Lefever, David W. The prognostic values of certain groupings of the test elements of the Thorndike examination for high-school graduates. Los Angeles, University of Southern California press, 1930. (University of Southern California Studies, Education Series no. 9.)
81. Lincoln, Edward, A. The relative standing of pupils in high school, in early college, and on college-entrance examinations. School and society, 5:417–20, Apr. 7, 1917.
82. MacPhail, A. H. Classification of freshmen at Brown University. Journal of educational research, 14:365–69, December 1926.
83. ——— The intelligence of college students. Baltimore, Md., Warwick and York, 1924.
84. Mann, Clair V. Measurement of engineering students. Rollo, Mo., 1933. (mimeographed.)
85. May, Mark A. Predicting academic success. Journal of educational psychology, 14: 429–40, October 1923.
86. Miller, L. W. An experimental study of the Iowa placement examination. Iowa City, Iowa, University of Iowa, 1930. (University of Iowa Studies in Education, V, no. 6, first series, no. 181.)
87. Miner, J. B. Scholarship and intelligence. Personnel journal, 6: 113–18, August 1927.
88. Nelson, M. J. Some data from the freshmen tests. School and society, 31: 772–74, June 7, 1930.
89. ——— A study in the value of entrance requirements at Iowa State Teachers College. School and society, 37: 262–64, Feb. 25, 1933.
90. ——— The Thorndike intelligence examinations for high-school graduates as a criterion for predicting college success. Madison, Wis., University of Wisconsin. Unpublished M.A. thesis.
91. ——— *and* Denny, E. C. The Terman and Thurstone group tests as criteria for predicting college success. School and society, 26: 501–502, Oct. 15, 1927.

92. Neuberg, Maurice, J. The relative values of several means of prognosticating success in college. Columbus, Ohio, Ohio State University, Ohio College Association Bulletin no. 60.

93. The Nineteen hundred and thirty-three college sophomore testing program. Educational record, 14: 522–71, October 1933.

94. Odell, Charles W. Attempt at predicting success in the freshman year at college. School and society, 25: 702–706, June 11, 1927.

95. ——— Predicting the scholastic success of college freshmen. Urbana, Ill., University of Illinois Bulletin, XXV, no. 2. (College of education, Bureau of educational research, Bulletin no. 37.)

96. Otis, A. S. Statistical method in educational measurement. Yonkers, N.Y., World Book Co., 1926.

97. Pierson C. D., *and* Nettels, C. H. A study of high-school seniors to determine who shall be recommended to college. School and society, 28: 215–16, Aug. 18, 1928.

98. Potthoff, E. The effectiveness of certain requirements in the selective admission of college students. School review, 37: 519–28, September 1929.

99. Proctor, William M. The high-school's interest in methods of selecting students for college admission. School and society, 22: 441–48, Oct. 10, 1925.

100. Remmers, H. H. A report on progress of a diagnostic and remedial study of failing students at Purdue University. Lafayette, Ind. Purdue University, 1927. (Purdue University Bulletin, XXVII, no. 14, and educational Reference Circular no. 1.)

101. ——— A study of freshmen placement tests at Purdue University, 1926–29. Lafayette, Ind., Purdue University, 1929. (Purdue University Bulletin, XXIX, no. 13.)

102. Rietz, H. L. Handbook of mathematical statistics. Boston, Mass., Houghton, Mifflin Co., 1924.

103. Root, W. T. The freshmen: Thorndike College entrance tests, first semester grades, Binet tests. Journal of applied psychology, 7: 77–92, March 1923.

104. Rugg, H. O. Statistical methods applied to education. Boston, Mass., Houghton Mifflin Co., 1917.

105. Scates, Douglas E. A study of high-school and first-year university grades. School review, 33: 182–92, March 1924.

106. Seashore, C. E. Notes on the use of mental tests in college. National Research Council, Washington, D.C. Dec. 30, 1921.

107. Segel, David. Differential diagnosis of ability in school children. Baltimore, Md., Warwick and York, 1934.

108. ——— Differential prediction of scholastic success. School and society, 39: 91–96, Jan. 20, 1934.

109. ——— Differential prediction of ability as represented by college subject groups. Journal of educational research, 25: 14–26, 93–98, January and February 1932.

110. ——— Prediction of success in junior college. Junior college journal, 1: 499–502, May 1931.

111. Segel, David, *and* Brintle, S. L. The relation of occupational interest scores as measured by the Strong Interest Blank to achievement test results and college marks in certain college subject groups. Journal of educational research. **27**: 442–45, February 1934.

112. ——— *and* Gerberich, J. R. The value of the American Council on Education Psychological Examination in the differential prediction of achievement in certain college subjects. Journal of applied psychology, 17: 637–45, December 1933.

113. Sorenson, Herbert. High-school subjects as conditioners of college success: Implications and theories concerning mental factors and faculties. Journal of educational research, **19**: 237–54, April 1929.

114. Stalnaker, John M. American Council Psychological Examination for 1926 at Purdue University. School and society, **27**: 86–88, January 21, 1928.

115. Stoddard, G. D. Iowa placement examinations. Iowa City, Iowa, University of Iowa, 1925. (University of Iowa Studies in Education, III, No. 2.)

116. ——— The use of qualitative measurements in inducting the student into institutions of higher learning and in the prediction of academic success. Studies in Education, 18th Yearbook of the National Society of College Teachers of Education, 1930, pp. 88–120.

117. Stone, C. L. The disparity between intelligence and scholarship. Journal of educational psychology, 13: 241–44, April 1922.

118. ——— The significance of Alpha in college. Journal of educational psychology, 13: 298–302, May 1922.

119. Symonds, P. M. Measurement in secondary education. New York, N.Y., Macmillan Co., 1927.

120. ——— The significance of intelligence tests in the University of Hawaii. School and society, 20: 601–606, Nov. 8, 1924.

121. Tallman, R. W. A critical analysis of student persistence at the State University of Iowa. Iowa City, Iowa, University of Iowa. (University of Iowa Studies IV, No. 1, 1927.)

122. Terman, L. M. Intelligence tests in colleges and universities. School and society, 13: 481–94, Apr. 23, 1921.

123. Tharp, James B. Sectioning classes in romance languages. Modern language journal, 12: 95–114, November 1927.

124. Thompson, C. O., *and* Russell, Dale. Some factors affecting the scholastic success of teachers of college freshmen. Bulletin of Indiana State Normal School, Eastern Division II, No. 4, 1927.

125. Thurstone, L. L. The fundamentals of statistics. New York, N.Y., Macmillan Co., 1925.

126. ——— *and* Thurstone, Thelma G. The 1930 Psychological examination. Educational record, 12: 160–78, April 1931.

127. Toll, C. H. Scholastic aptitude tests in Amherst College. School and society, 28: 524–28, Oct. 27, 1928.

128. Toops, H. A., Atkins, D., *and* Meyer, Lorenz. A follow-up investigation of a study performance test at Ohio State University. Columbus, Ohio, Ohio State University, Ohio College Association Bulletin No. 66.

129. Tyler, Ralph W. Techniques for evaluating behavior. Educational research bulletin, 13: 1–11, Jan. 17, 1934.

130. Van Wagenen, M. J. Some results and influences derived from the use of the Army tests at the University of Minnesota. Journal of applied psychology, 4: 59–72, March 1920.

131. Whitman, A. D. The value of the examination of the College Entrance Examination Board as predictions of success in college. Bureau of publications, Teachers college, Columbia University. 1926.

132. Whitney, F. L., *and* Goodman, A K. The relationship of intelligence and classification scores to mortality and scholastic ratings. Department of Research, Colorado State Teachers College, Greeley, Colo., August 1930.

133. Williamson, E. G. An analysis of scholastic aptitude of freshmen in some private colleges in Minnesota. School and society, 34: 674–80, Nov. 14, 1931.

134. Wood, Ben D. Measurement in higher education. Yonkers, N.Y., World Book Co., 1923.

135. Zook, George F. Accrediting schools and colleges. Educational record, 15: 10–26, January 1934.

O