PREDICTION OF COLLEGE SUCCESS
FROM HIGH SCHOOL GRADES AND
INTELLIGENCE TEST SCORES

by

ALMA HOCHULI

B. S., Kansas State Agricultural College, 1927

A THESIS

submitted in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

KANSAS STATE AGRICULTURAL COLLEGE

1928

TABLE OF CONTENTS

PURPOSE

In public schools, juvenile courts, reform schools, police courts, prisons, and institutions for the defectives, the use of intelligence tests has become quite universal. Their value as educational measurements can hardly be overestimated. Tests give an essential part of the necessary information from which a pupil's possibilities of future mental growth can be foretold, and upon which his further education can be most profitably directed. This study deals with the predictive value not only of mental test results but also of high school grades or marks of achievement.

It must be realized that mental tests and high school grades cannot predict with perfect accuracy the degree of success of a college student. Neither takes a complete account of health, temperament, personality, and other factors of success, although high school grades tend to do so to some extent. Moreover, the college grades from which we infer success or failure are exceedinly variable and inaccurate measures of achievement. College grades would be a complete measure of achievement provided they represented the entire ability of a student to achieve plus his motivation. Since there are many interfering factors grades become variable measures. Grades reported from the high schools are very fluctuating from year to year due to the changing standards of the high schools. In 1919-1920 Dr. J. C. Peterson (1) working with 116 freshmen in the Division of Engineering at the Kansas State Agricultural College found a correlation of only .319 between high school grades and first semester freshman college grades. For students entering the Kansas State Agricultural College in the fall of 1921 Dr. V. L. Strickland (2) gives a correlation of .328 between high school scholarship and the first three years of college scholarship. In 1920 Dr. L. L. Thurstone (3) reported a correlation of .29 between average high school scholarship and first-term engineering scholarship at Carnegie Institute. Thus we must not expect too high a correlation between test scores and grade averages. Correlations reported from Columbia College (4) between the first two years of scholarship and Thorndike Intelligence scores is .672; between scholarship and Regents' Examinations .644.

This investigation is an attempt to find out any unexpressed relationship that may exist between the test results, scholarship in high school, and scholarship in college. The statistics are treated in such a manner as to secure the greatest possible correlation between grades received both in high school and college and data furnished by the Equation Completion test.

ACKNOWLEDGEMENTS

I desire to express my great indebtedness to Dr. J. C. Peterson who furnished the data upon which this investigation was made. It was only by his valuable suggestions throughout this study that I have been able to carry on the problem.

I am also under great obligations to Dr. W. H. Andrews for the instruction and aid that I received in his course in statistical methods.

SOURCE OF MATERIALS

All Freshmen entering the Kansas State Agricultural College are required to take an intelligence test at the beginning of the fall semester. Since 1925 the Equation-Completion test, Form A, has been a part of the intelligence examination given.

Instructions with the Equation-Completion test indicate that the score is recorded by five minute intervals for twelve periods, making the examination require sixty minutes. Assuming that the individual has become well enough acquainted with the test in the first fifteen minutes, we use the score recorded for the last forty-five minutes.

The test scores are those obtained from the freshmen in the Equation-Completion test in the fall of 1926. The high school grade averages used were taken from the transcripts sent to the registrar's office at the college. The college grades were those earned by freshmen during the first semester 1926-1927.

Kansas State Agricultural College uses the following system of grading:

E-Excellent

G-Good

M-Medium

P-Passing

F-Failure

In making the correlations the grades were weighted as follows:

E-5, G-4, M-3, P-2, F-1.

The grades were distributed approximately in the following proportions:

E-5%, G-22%, M-46%, P-22%, F-5%.

The above distribution of grades is in accordance with the normal distribution curve.

Of those who took the mental test in September, 1926, and for whom high school grades and first semester college grades could be obtained there were:

Division	Number of Cases
Agriculture	68
Engineering	287
General Science (Boys)	122
General Science (Girls)	88
Home Economics	109
Total	674

Since the number of cases in agriculture was very small this division was combined with the general science boys making a total of 190 cases with which to carry on correlation work.

PROCEDURE

The purpose of this problem is to obtain the predictive value of high school grade averages and the equation-completion test for success in college as evidenced by the scholarship of the first semester in the freshman year. This section is a brief explanation of the means of accomplishing this purpose.

For each division of the college we have the following data:

College grade averages

Equation-completion test scores

High school grade averages

Thus we have four separate groups with the agriculture and general science boys as one group. The divisions and sexes were treated separately in order to secure more homogeneous groups. Not only do standards of grading differ in the various divisions but there is a general tendency for girls to surpass boys in scholarship.

If we are to obtain the most accurate prediction of college success the high school grade averages and the equation-completion test scores must be properly weighted. This problem of finding the correlation resulting from the most favorable combination of the parts is solved by the statistical procedure known as multiple correlation. Upon the calculation of the partial coefficients of correlation we can determine the regression equation which will indicate the weight to be given to each independent variable. We can tell from the regression equation just what part each of the variables plays in determining the grades which will most probably be received. The probable error of estimate will then serve as an index of the reliability of our estimated grades.

RESULTS

Since so many of the zero, partial, and multiple coefficients indicate significant relationships, all the coefficients used in reaching the final results will be given. Each is followed by its probable error. The results for each division are stated separately.

Notation of the subscripts.

1-college grade averages

2-equation completion test scores

3-high school grade averages

General Science Girls

I. Coefficients obtained in computing $R_{1(23)}$

1. Coefficients of zero order

$r_{12} = .4447 \pm .058$

$r_{13} = .6158 \pm .045$

$r_{23} = .4106 \pm .060$

2. Partial coefficients

$r_{12.3} = .2670 \pm .067$

$r_{13.2} = .5304 \pm .052$

$r_{23.1} = .1937 \pm .069$

3. Multiple coefficient

$R_{1(23)} = .6507 \pm .042$

II. Regression equation and standard error of estimate.

1. Deviation form

$x_1 = .1104x_2 + .247x_3 \pm .7989$

2. Score form

$X_1 = .1104X_2 + .247X_3 - 46.78$

Home Economics

I. Coefficients obtained in computing $R_{1(23)}$

1. Coefficients of zero order

$$r_{12} = .4047 \pm .054$$

$$r_{13} = .5945 \pm .042$$

$$r_{23} = .3989 \pm .054$$

2. Partial coefficients

$$r_{12.3} = .2272 \pm .061$$

$$r_{13.2} = .5164 \pm .050$$

$$r_{23.1} = .2152 \pm .062$$

3. Multiple coefficient

$$R_{1(23)} = .6219 \pm .039$$

II. Regression equation and standard error of estimate.

1. Deviation form

$$x_1 = .09891x_2 + .1934x_3 \pm .7499$$

2. Score form

$$X_1 = .09891X_2 + .1934X_3 - 36.35$$

Engineering

I. Coefficients obtained in computing $R_{1(23)}$

1. Coefficients of zero order

$$r_{12} = .4424 \pm .032$$

$$r_{13} = .5651 \pm .027$$

$$r_{23} = .5469 \pm .028$$

2. Partial coefficients

$$r_{12.3} = .1930 \pm .038$$

$$r_{13.2} = .4304 \pm .032$$

$$r_{23.1} = .4012 \pm .033$$

3. Multiple coefficient

$$R_{1(23)} = .5871 \pm .026$$

II. Regression equation and standard error of estimate.

1. Deviation form

$$x_1 = .1095x_2 + .1955x_3 \pm .9773$$

2. Score form

$$X_1 = .1095X_2 + .1955X_3 \quad -40.655$$

Agriculture
and
General Science Boys

I. Coefficients obtained in computing $R_{1(23)}$

1. Coefficients of zero order

$$r_{12} = .3992 \pm .041$$

$$r_{13} = .6050 \pm .031$$

$$r_{23} = .3753 \pm .042$$

2. Partial coefficients

$$r_{12.3} = .2332 \pm .046$$

$$r_{13.2} = .5356 \pm .035$$

$$r_{23.1} = .1832 \pm .047$$

3. Multiple coefficient

$$R_{1(23)} = .6328 \pm .029$$

II. Regression equation and standard error of estimate.

1. Deviation form

$$x_1 = .11078x_2 + .25009x_3 \pm 1.002$$

2. Score form

$$X_1 = .11078X_2 + .25009X_3 - 43.663$$

While the difference in the mean score obtained by the various divisions is not essential to a statement of our problem such differences are of general interest and may be summarized so briefly that they are given below.

Mean Equation-Completion Test Score of Various Groups of Freshmen

General Science Girls	252.20 ± 5.856
General Science Boys	245.49 ± 5.207
Engineering Boys	240.17 ± 3.092
Home Economics Girls	224.00 ± 4.617
Agriculture Boys	193.01 ± 6.465

The differences between these mean scores are as follows:

	Mean Dif.	P.E.Dif.	Dif.:P.E.Dif.
G. S. Girls minus Ag. Boys	59.19	± 8.715	6.78:1
Eng. Boys minus Ag. Boys	47.16	± 7.160	6.58:1
G. S. Boys minus Ag. Boys	52.48	± 8.292	6.32:1
H. E. Girls minus Ag. Boys	30.99	± 7.936	3.90:1
G. S. Girls minus H. E. Girls	28.20	± 7.418	3.80:1
G. S. Boys minus H. E. Girls	21.49	± 6.949	3.09:1
Eng. Boys minus H. E. Girls	16.17	± 5.549	2.91:1
G. S. Girls minus Eng. Boys	12.03	± 6.615	1.81:1
G. S. Boys minus Eng. Boys	5.32	± 6.048	.87:1
G. S. Girls minus G. S. Boys	6.71	± 7.827	.85:1

According to Rugg, "Statistical Methods Applied to Education," a difference which is five times its probable error would disappear in only one out of 1315 cases on the average. A difference of four times its probable error would disappear in 143 cases. Only three of the differences are more than four times their probable error, but three others are very close to four times so that these also may be considered fairly significant.

As a sideline to this study a problem was worked out with 109 cases in the Home Economics division to determine the effect of the rights, errors, and omissions of the equation-completion test in predicting college grades.

Notation of the subscripts.

1-college grade averages

2-rights of equation-completion test

3-errors of equation-completion test

4-omissions of equation-completion test

I. Coefficients obtained in computing $R_{1(234)}$

1. Coefficients of zero order

$r_{12} = .4047 \pm .054$

$r_{13} = -.289 \pm .059$

$r_{23} = -.2787 \pm .060$

$r_{14} = -.278 \pm .060$

$r_{24} = -.4189 \pm .054$

$r_{34} = .2616 \pm .060$

2. Partial coefficients

a. First order

$$r_{12.3} = .3525 \pm .057$$

$$r_{13.2} = -.2006 \pm .062$$

$$r_{14.3} = -.2190 \pm .061$$

$$r_{14.2} = -.1306 \pm .063$$

$$r_{24.3} = -.3732 \pm .056$$

$$r_{34.2} = .1661 \pm .063$$

b. Second order

$$r_{12.34} = .2991 \pm .059$$

$$r_{13.24} = -.1829 \pm .062$$

$$r_{14.23} = -.1006 \pm .064$$

3. Multiple coefficient

$$R_{1(234)} = .4533 \pm .052$$

II. Regression equation and standard error of estimate.

1. Deviation form

$$x_1 = .1556x_2 + .0767x_3 + .0422x_4 \pm .0513$$

2. Score form

$$X_1 = .1556X_2 + .0767X_3 + .0422X_4 - 32.805$$

CONCLUSIONS

In each of the divisions except the division of Agriculture it will be noticed that the correlation between college grade averages and high school grade averages is somewhat greater than that found between college grade averages and test scores. This would indicate that high school grades are more valuable than equation-completion test scores for predicting college success. It should be noted here though that as the entire freshmen test is now made up, the equation-completion test comprises less than one-fourth of the total. It is however, the only test of the ones given in the past which will be used in the future. Thus we cannot conclude that high school grades have more predictive value than intelligence tests since we have only the equation-completion test upon which to base such an assumption. Perhaps the condition found in this study is due to the fact that high schools are coming more and more to the idea of using objective tests, such as colleges are now employing, as a model. College tests have become more efficient so that tests in high schools are correspondingly causing the teacher to grade more closely and recognize the success of his pupils.

The multiple correlation coefficients of college grades with the other two variables are quite a lot higher than the zero coefficients of college grades with high school grade averages so that it is well to take into account the equation completion test scores.

Of course we cannot assume that the predicted grades are wholly reliable. In interpreting standard errors we always make the assumption that measures obtained from successive samples are distributed according to the normal probability curve. This assumption is only true when the number of cases is large. The number of cases is large enough to indicate a fair degree of reliability in this study. There being only five grade steps though has probably cut down the accuracy of the results.

In the study made of the Home Economics division to determine the effect of considering errors and omissions as well as rights of the equation completion test in predicting college grades it was found that the correlation was raised nearly five points above the correlation of rights only with college grade averages.

APPENDIX

Statistics Involved

The writer in preparing this study has assumed that the reader has a sufficient knowledge of statistical procedure to follow through the methods used in partial and multiple correlations.

The formula used for the zero order correlations is Pearson's Product Moment Method formulated by Dr. Leonard P. Ayres (5). All results have been checked by a Monroe Calculator, or in some cases by a different formula to obtain the same constant.

To compute $R_{1(23)}$ all the possible zero order correlations and partial correlations are needed.

r_{12}	$r_{12.3}$
r_{13}	$r_{13.2}$
r_{23}	$r_{23.1}$

The multiple $R_{1(23)}$ can be obtained by rewriting the order of the subscripts. This gives a check on the constants.

$$R_{1(23)} = \sqrt{1-(1-r^2_{12})(1-r^2_{13.2})}$$

$$R_{1(32)} = \sqrt{1-(1-r^2_{13})(1-r^2_{12.3})}$$

The necessary formulae for the standard deviations, regression coefficients, regression equation, and the probable error of estimate follow:

Standard Deviations

$$S.D._{1.23} = S.D._1 \sqrt{1-r^2_{12}} \sqrt{1-r^2_{13.2}}$$

$$S.D._{2.13} = S.D._2 \sqrt{1-r^2_{12}} \sqrt{1-r^2_{23.1}}$$

$$S.D._{3.12} = S.D._3 \sqrt{1-r^2_{13}} \sqrt{1-r^2_{23.1}}$$

Regression Coefficients

$$b_{12.3} = r_{12.3} \frac{S.D._{1.23}}{S.D._{2.13}}$$

$$b_{13.2} = r_{13.2} \frac{S.D._{1.23}}{S.D._{3.12}}$$

Regression Equation

$$X_1 = b_{12.3} X_2 + b_{13.2} X_3 + K$$

Standard Error of Estimate

$$P.E._{est.X_1} = .6745 \ S.D. \ X_1$$

BIBLIOGRAPHY

1. PETERSON, J. C. An Unpublished Report to the President of the Kansas State Agricultural College. 1920.

2. STRICKLAND, V. L. Relative Levels of Intelligence in College and High School. Kansas State Agricultural College Bulletin, 10, Number 4, November 15, 1926, p. 69.

3. Report of Committee Number 22 on Intelligence Tests. Engineering Education, Proceedings of the Society for the Promotion of Engineering Education--Ann Arbor Meeting, 1920, pp. 349-361.

4. WOOD, BEN D. Measurement in Higher Education. New York, 1923, p. 85.

5. AYRES, LEONARD P. The Correlation Ratio. Journal of Educational Research, II, p. 452.

Date Due

May21 '58X		
Nov 21 '59T		