THE VALUE OF HOMOGENEITY IN PREDICTING COLLEGE SUCCESS FROM HIGH SCHOOL GRADES

by

HARRY ENOCH ADELL

B. S., Kansas State Teachers College, Emporia, 1929

A THESIS

submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

KANSAS STATE COLLEGE
OF AGRICULTURE AND APPLIED SCIENCE

1931

TABLE OF CONTENTS

INTRODUCTION

Colleges and universities have for a number of years shown an ever increasing interest in and have been actively engaged in finding better means for selection and guidance of students who seek higher learning. They have come to use extensively various objective measurements, as teachers' estimates, high school grades, mental tests, achievement tests, aptitude tests, interest tests, etc.

Through statistical procedure educators have been able to find fairly satisfactory correlations between these various measurements and college grades. There is much disagreement as to the relative values of these different measures, but it is rather generally agreed that mental tests are perhaps the best single measure to date for predicting success in college. However, high school grades have been found very useful in combination with tests because they include certain factors which are not measured by intelligence tests, as interest, health, industry, outside employment, etc. These factors undoubtedly affect the quality and quantity of work done by a student.

A study of correlations of college grades with test scores and high school grades was made at the Kansas State

College in 1920. The tests used were the S. P. E. E. tests for college freshmen by Thurstone. One hundred sixteen freshmen from the engineering division were included in this study. The college grades referred to were the first semester means. The correlations found were (a) between high school grades and college grades, .319; (b) between test scores and college grades, .566; and (c) between college grades and combined high school grades and test scores, .703. The test scores in (c) were given six times as much weight as the high school grades. In this study[1] these correlations would all have been somewhat higher if all the engineering freshmen instead of only a selected group had been tested. Note that, although the high school grades when taken alone have a comparatively slight predictive value, the correlation coefficient becomes materially increased when high school grades are combined with test scores.

---

[1] The intelligence tests were purely voluntary on the part of the students, and approximately one-third of the engineering freshmen failed to respond to the call. It is safe to say that these students are, on the whole, considerably less intelligent than those who submitted to the tests. This inference is substantiated by a comparison of the attainments of the students who took only a part of the tests with those of the students who responded to all the tests. There were 51 students who took some but not all of the tests. Most of these individuals responded to the first three tests but failed to return the following day for the last three. In the tests to which they did respond, exactly two-thirds of these 51 students fell below the average of our standard group of 116 whose test records are complete.

It has often been supposed that the correlation between high school grades and college grades is considerably reduced in general by the fact that the high schools contributing to any college class are usually very diverse in their standards of grading. Our aim here is to test out this supposition.

Accordingly the problem of this study is to find the degree to which uniformity of high school standards in grading affects the correlation between high school grades and first semester college grades. This is done by comparing correlations of college grades in each division and also in all divisions combined with those of (1) a single high school and (2) a large number of high schools scattered over a number of states.

MATERIAL AND METHOD

Preparatory to this study a tabulation sheet was prepared upon which grades and credits in the various fields were to be recorded. The desired information was so arranged on this sheet as to provide a complete record for eight students' work on each sheet. A sample of the form used is given below:

Name _____ Age ___ Sex ___

English ___ ___ ___ ___ ___ ___ ___

Mathematics ___ ___ ___ ___ ___ ___ ___

Science ___ ___ ___ ___ ___ ___ ___

Social Science ___ ___ ___ ___ ___ ___ ___

Language ___ ___ ___ ___ ___ ___ ___

Miscellaneous ___ ___ ___ ___ ___ ___ ___

All Manhattan High School graduates who had entered Kansas State College as freshman during a period of six years, 1924 to 1929 inclusive, were first listed. Only such high school graduates as had completed at least three years of work in Manhattan High School and one semester or two summer sessions of college work were included in the study. There were a few students from this list who had to be dropped since they had not completed one semester of college work or its equivalent. The total number of cases available for study was 372.

All grades for each student listed were next collected from the cumulative records in the senior high school of Manhattan and were grouped under the following heads: English, mathematics, science, social science, language, miscellaneous, advanced credits, and the mean. The grades for all courses not included in any of the preceding groups were grouped under the head of "miscellaneous". In addition

the grades for the subjects recommended in the State Course of Study for the junior and senior years were grouped under the head of advanced credits.

From the foregoing it becomes evident that this group of students is a rather select group in as much as they have had the same type of administration and uniform grading system and standard applied to all. They will be referred to hereafter as the "homogeneous" group. The other group of students used in this study have come from schools where various grading systems and standards have been used. They will be referred to hereafter as the "heterogeneous" group.

The grades for the first semester's college work were next tabulated both for the homogeneous and for the heterogeneous groups of students. These grades were in the main available in the psychology office where they had been assembled previously from year to year.

A random sampling method was used to determine the names of the students that were to be used for the heterogeneous group. The names of the students in each division were arranged alphabetically for each year. In each division the name of the student immediately following the name of a student from the homogeneous group was selected for the heterogeneous group. In case there were two or three names from the homogeneous group in succession, the

next two or three names following these would be used for the heterogeneous group. Thus, there were in all also 372 students selected for the heterogeneous.

The high school grades for the students in the heterogeneous group were in part available in the psychology office. The remainder were collected from the original transcripts at the Registrar's office. Since a varied system of grades and standards was used in this group it was necessary to reduce all to the same standard. In cases where numerals or letters were used as marks, the average for each mark was taken, and then expressed as a per cent.

The Manhattan High School and the College have a uniform system of grades. The grades were weighed by giving a certain number of points for each grade, as follows:

H. S. Grade	Col. Grade	Per Cent	Points
I	E	93 to 100	5
II	G	85 to 92	4
III	M	78 to 84	3
IV	P	70 to 77	2
F	F	Below 70	1

The first semester college mean grade for each student in both the homogeneous and the heterogeneous groups had been collected and changed to grade indexes and were available in the psychology office, except in a few cases.

For correlation purposes the high school grades for the heterogeneous group were divided into 13 equal intervals by including two grades in each interval. The high school grades for the homogeneous group and the college grades for both groups were also divided into 13 equal intervals. Each step from one whole number to the next whole number was divided into three steps: 1.00-1.33, 1.34-1.66, 1.67-1.99; etc.

The results of the correlations for the high school grades and the first semester college grades in each division for both the homogeneous and the heterogeneous groups and their probable errors are given in Tables I to IV. Those for the divisions combined are given in Table V. Ayers' formula as given in Volume I in the March number of the "Journal of Educational Research" was used in computing all the correlations except the means in each of the first five tables. These means were worked out on Holzinger's correlation sheet, and were also tested for linearity by the application of Blakeman's formula. The means and standard deviations for all comparisons are given in Tables VI to X inclusive.

Table I

Division of General Science

Correlation Between

High School Grades and First Semester College Grades

High School	Homogeneous			Heterogeneous		
Subject	N	R	P.E.	N	R	P.E.
English	211	.679	.025	203	.606	.03
Mathematics	208	.708	.023	207	.587	.03
Science	209	.630	.028			
Social Science	204	.672	.026			
Language	178	.595	.032	141	.592	.037
Miscellaneous	207	.586	.03			
Advance Credits	211	.726	.022			
Mean	211	.755	.02	211	.624	.023

Table II

Division of Home Economics

Correlation Between

High School Grades and First Semester College Grades

High School	Homogeneous			Heterogeneous		
Subject	N	R	P.E.	N	R	P.E.
English	63	.661	.048	38	.415	.07
Mathematics	63	.727	.04	63	.552	.06
Science	62	.707	.042			
Social Science	57	.579	.059			
Language	49	.757	.041	42	.453	.075
Miscellaneous	63	.523	.062			
Advance Credits	63	.729	.04			
Mean	63	.770	.035	63	.56	.058

Table III

Division of Engineering

Correlation Between

High School Grades and First Semester College Grades

High School	Homogeneous			Heterogeneous		
Subject	N	R	P. E.	N	R	P.E.
English	65	.744	.037	60	.498	.065
Mathematics	65	.719	.04	65	.399	.07
Science	64	.685	.045			
Social Science	65	.649	.048			
Language	52	.706	.047	42	.075	.104
Miscellaneous	65	.480	.065			
Advance Credits	65	.777	.034			
Mean	65	.780	.031	65	.506	.062

Table IV

Division of Agriculture

Correlation Between

High School Grades and First Semester College Grades

High School	Homogeneous			Heterogeneous		
Subject	N	R	P.E.	N	R	P.E.
English	33	.494	.091	30	.427	.101
Mathematics	33	.598	.078	33	.331	.109
Science	33	.303	.106			
Social Science	32	.468	.093			
Language	23	.296	.123	15	.414	.125
Miscellaneous	33	.231	.114			
Advance Credit	33	.564	.082			
Mean	33	.507	.089	33	.302	.109

Table V

Divisions Combined for Each Subject

Correlation Between

High School Grades and First Semester College Grades

High School	Homogeneous			Heterogeneous		
Subject	N	R	P.E.	N	R	P.E.
English	372	.684	.018	351	.543	.025
Mathematics	369	.799	.012	368	.551	.035
Science	360	.616	.022			
Social Science	358	.662	.02			
Language	302	.610	.024	247	.520	.031
Miscellaneous	366	.641	.021			
Advance Credits	372	.730	.017			
Mean	372	.743	.015	372	.571	.024

Table VI

Division of General Science

		Homogeneous					Heterogeneous			
		High School		College			High School		College	
Subject	N	M	S.D.	M	S.D.	N	M	S.D.	M	S.D.
English	211	3.639	1.095	3.050	1.011	203	87.48	4.716	2.974	.986
Math.	208	3.486	1.159	3.032	1.015	207	86.69	5.721	2.874	.991
Sci.	209	3.301	1.033	3.027	1.011					
Soc. Sci.	204	3.488	1.092	3.051	1.020					
Lang.	178	3.376	1.258	3.185	.947	141	86.21	6.346	2.908	.967
Misc.	207	3.641	.799	3.016	1.006					
Adv. Cr.	211	3.452	.989	3.050	1.011					
Mean	211	3.450	.955	3.031	1.009	211	85.83	4.777	2.846	.993

Table VII

Division of Home Economics

		Homogeneous					Heterogeneous			
		High School		College			High School		College	
Subject	N	M	S.D.	M	S.D.	N	M	S.D.	M	S.D.
English	63	3.675	1.001	2.823	.785	58	87.38	4.860	2.955	.511
Math.	63	3.616	1.217	2.823	.785	63	86.30	6.008	2.828	.557
Science	62	3.111	1.072	2.836	.778					
Soc. Sci.	57	3.398	1.109	2.816	.928					
Lang.	49	3.286	1.345	2.942	.826	49	86.55	5.814	2.855	.578
Misc.	63	3.647	.813	2.823	.785					
Adv. Cr.	63	3.446	.850	2.823	.785					
Mean	63	3.472	.847	2.823	.785	63	87.27	4.652	2.836	.560

Table VIII

Division of Engineering

		Homogeneous					Heterogeneous			
		High School		College			High School		College	
	N	M	S.D.	M	S.D.	N	M	S.D.	M	S.D.
English	65	3.050	.929	2.706	.781	60	84.40	5.050	2.772	.855
Math.	65	3.458	.966	2.706	.813	63	85.89	4.680	2.752	.872
Science	64	3.449	.956	2.712	.807					
Soc. Sci.	65	3.355	.929	2.706	.813					
Lang.	52	2.971	1.226	2.856	.781	42	85.00	6.760	2.900	.782
Misc.	63	3.642	.775	2.700	.804					
Adv. Cr.	65	3.190	.869	2.706	.781					
Mean	65	3.268	.802	2.744	.802	65	86.27	5.786	2.752	.847

Table IX

Division of Agriculture

		Homogeneous					Heterogeneous			
		High School		College			High School		College	
Subject	N	M	S.D.	M	S.D.	N	M	S.D.	M	S.D.
English	33	2.512	.931	2.420	.716	30	83.53	5.214	2.606	.795
Math.	33	2.801	1.181	2.420	.716	33	83.54	5.082	2.640	.784
Science	33	3.141	.818	2.420	.716					
Soc. Sci.	32	2.596	.991	2.390	.718					
Lang.	23	2.495	1.179	2.557	.713	15	84.20	7.303	2.820	.804
Misc.	33	3.121	.676	2.420	.716					
Adv. Cr.	33	2.842	.792	2.420	.716					
Mean	33	2.828	.753	2.421	.715	33	84.77	4.920	2.637	.787

Table X

All Divisions Combined

		Homogeneous					Heterogeneous			
		High School		College			High School		College	
Subject	N	M	S.D.	M	S.D.	N	M	S.D.	M	S.D.
English	372	3.437	1.101	2.886	.938	351	86.60	5.064	2.831	.888
Math.	369	3.411	1.012	2.886	.940	368	86.20	5.616	2.824	.890
Science	368	3.379	1.016	2.886	.938					
Soc. Sci.	358	3.376	1.081	2.885	.950					
Lang.	302	3.222	1.286	3.033	.944	247	25.93	6.324	2.891	.861
Misc.	366	3.596	.800	2.876	.791					
Adv. Cr.	372	3.341	.944	2.886	.935					
Mean	372	3.371	.909	2.884	.937	372	87.01	4.712	2.817	.894

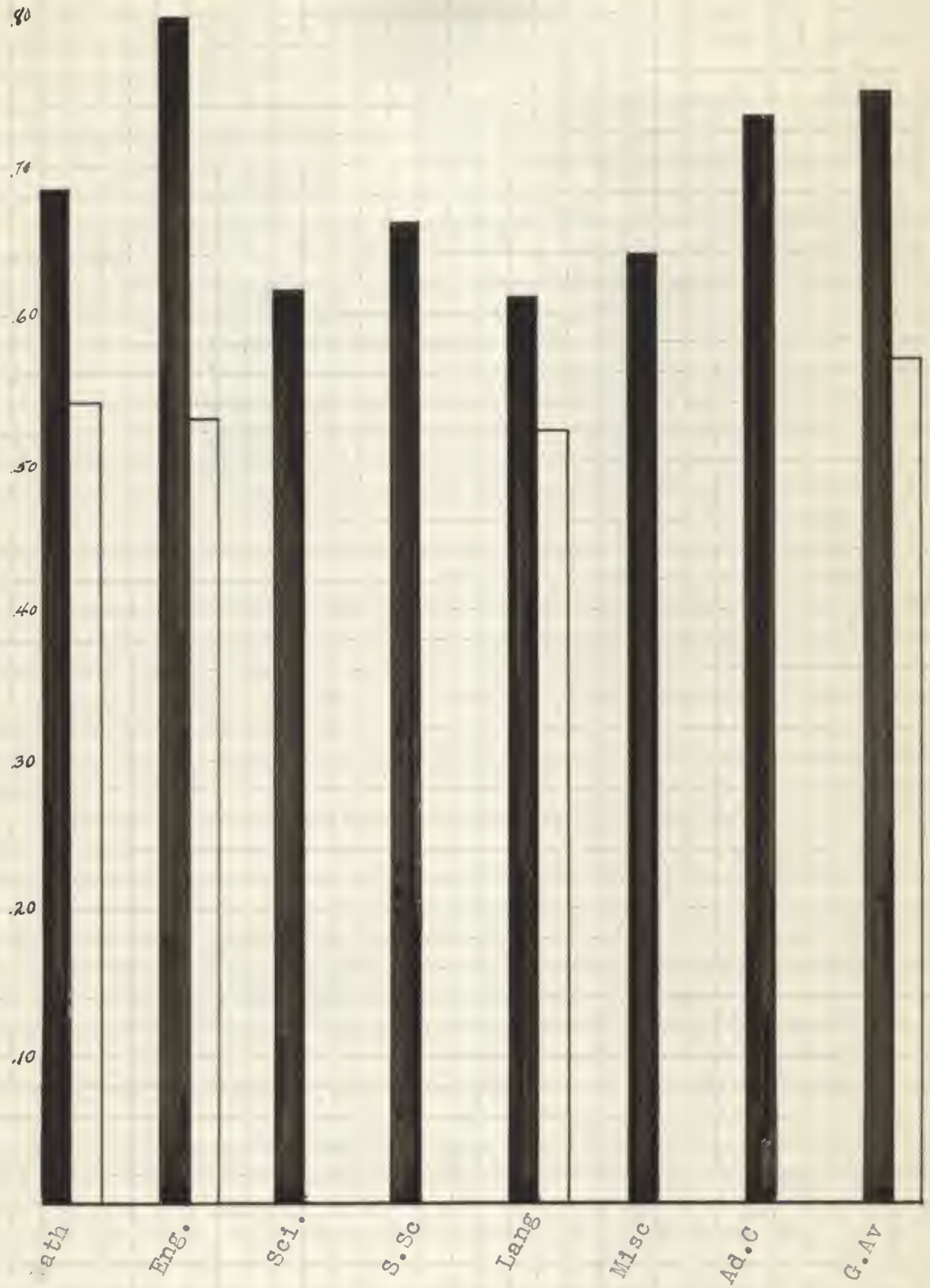

Fig I. Graphical representation for correlations in table V.

DISCUSSION AND CONCLUSION

The results of the correlations in each of the four divisions have a general trend. At a glance it will be seen that mathematics and English correlate more highly with college grades than do any of the other subjects. The reason for this is perhaps that each of these subjects is fairly continuous and similar in subject matter from the lowest grade of the elementary school to the corresponding courses in college. In the light of this fact it is to be expected that the subjects which have the least continuity and similarity would have the lowest correlation with college grades, and this is what happened. The miscellaneous group of subjects has the lowest correlation in each division. It will also be seen from the tables that the advanced credits group, which comprises the subjects of the junior and senior years only, has the highest correlation in each of the divisions. This perhaps indicates a number of things: The high school curriculum offers a greater opportunity for selection of subjects in the junior and senior years, and consequently there will be greater similarity between the advanced subjects and the college subjects. Students are getting better adjusted to the

subjects, teachers, and school requirements, and are acquiring better study habits.

In each of the subjects and the mean of the homogeneous group we have a higher correlation than in the corresponding heterogeneous group, with the exception of language in the division of agriculture. There were only 15 students from the heterogeneous group and 23 from the homogeneous group who had taken language in this division. These are not sufficient numbers to make any comparison significant. Also, according to Rugg (10), the correlation coefficient should be at least four times its probable error to be significant, and in neither case does the correlation coefficient meet this test.

In Figure 1 we have a graphical representation of the results of the correlations as tabulated in Table V, where the divisions are combined for each subject. It will be seen at a glance that the subjects in the homogeneous group have a higher correlation than any of those from the other group.

Rugg, after having made an exhaustive examination of many correlation tables, has classified correlations below .20 as negligible, from .20 to .40 as significant, from .40 to .60 as marked, and from .60 to .70 or above as high. In the light of this classification, referring to Table V

again, we note that the correlations for the subjects in the homogeneous group are high, ranging from .61 to .799. Those for the heterogeneous group are marked, ranging from .52 to .571. Language has the lowest correlation in each group. Referring also to Table X, we find for both groups that in language we have the lowest mean and the widest standard deviation of high school grades while for the college grades the highest mean. This indicates that of the many who elect language in high school a goodly number fail or do poor work in it, drop out, and consequently the mean for college grades is raised. In Table V again, we note that mathematics and English have the highest correlations in each group. Mathematics in the homogeneous group shows .268 higher correlation than in the heterogeneous. English in the homogeneous group shows .14 higher correlation than in the heterogeneous group. Language in the homogeneous group shows .09 higher correlation than in the heterogeneous group. The correlation between the high school mean in the homogeneous group with the college mean reached the high correlation of .743 while that of the heterogeneous group reached only .571.

These correlations of .743 for the homogeneous group and .571 for the heterogeneous group are found to be somewhat low as revealed in the results of the application of

Blakeman's test for linearity. Consequently, the values for eta, which are .774 for the homogeneous group and .617 for the heterogeneous group, are more nearly right.

The correlations then obtained in this study between high school and college grades are much higher than those found earlier at this college. This is probably due in part to a nearer approach to the standardization of our schools in the last few years. Eastern institutions have made progress along this line earlier, and consequently other investigators have found somewhat higher correlations in this field. The results of a few of these studies might well be mentioned here.

Dearborn (5) of the University of Wisconsin in his study found a correlation of .81. His study was made in 1909 and included 472 cases. This result seems to be unusually high and is undoubtedly an exception. Lincoln (6) of Harvard University in 1917 in his study of 253 cases chosen at random and representing 30 schools, found a correlation of .69 between high school grades and college grades, a correlation of .47 between entrance examinations and college grades, and a correlation of .46 between entrance examinations and high school grades. He concluded that the high school record was a better forecast of success in college than was the record made on entrance examinations. Beatley (2), also of Harvard University, in

1922 made a study which supplemented Lincoln's earlier study in which he included 423 cases representing 208 schools and 36 states. He found a correlation of .56 between high school grades and first year college grades, a correlation of .50 between the comprehensive examination and first year college grades, and a correlation of .37 between the comprehensive examination and high school grades. He concluded that "the school record gives a somewhat better indication of the quality of work that man will do in college than does the record on the comprehensive examination".

Strickland (13) in a study of 935 cases representing 21 schools found a correlation of .45 between high school grades and test scores. Terman (14) of Stanford University reporting on results found in 25 colleges, found correlations ranging from .33 to .74 between high school and college marks and of .29 to .63 between test scores and college marks. He concluded that "the Thorndike intelligence examination is probably the best of those available for the purpose of predicting scholastic success in college". Roberts (10) of the University of Washington reports correlations of .53 to .69 between high school and college marks and of .31 to .60 between test scores and college marks.

Scates (12) of Chicago University reported on 1707 students who had entered the University of Chicago directly

from high schools of Chicago during the five years 1916 to 1921 inclusive. Only grades made in non-vocational subjects were included. He found a correlation of .61 between high school grades and freshman college grades. MacPhail (7) in reporting on 37 colleges found a mean correlation of from .40 to .45 between high school and college grades. He cited a number of other results from similar studies, and finally concluded that intelligence examinations would give a better forecast of success in college than would high school grades, but that the latter would give a better indication of the quality of work that men will do in college than does the comprehensive examination. Odell (8) of Urbana University of Illinois reporting on 368 colleges in which 11,500 cases were included found a correlation of .55 between high school and college freshmen grades. He concluded that "combining the evidence from all the studies along this line with which the writer is familiar the statement seems warranted that, in general, a score on any one of the best intelligence tests and the proper high school mark have about equal value in predicting probable freshmen marks".

The S. P. E. E. (15) in reporting on the results obtained from over 6,000 cases in 43 institutions states that "the Thorndike psychological examination correlates .29 with first year engineering scholarship and that for the

same data a coefficient of .29 was found between high school mean and first year engineering scholarship". He also states that it is significant that the test and high school marks should correlate to the same degree with first year scholarship.

From the foregoing one may reasonably infer that investigators have more generally found correlations of .40 to .60 between high school grades and college grades. There is no report of homogeneity having been established in any of these studies. In this study a correlation of .571 for the heterogeneous group compares favorably then with the findings of other investigators. Therefore, a correlation of .743 for the homogeneous group becomes very significant. The correlation of .371 for the heterogeneous group has a predictive index of .19. The correlation for the homogeneous group has a predictive index of .33. By this comparison we see that the correlation for the homogeneous group has almost twice as high a predictive value as the correlation for the heterogeneous group.

According to Terman (6) at least 10 per cent of those who enter college are totally unfit to profit by a college education. Charters (3) states that of these 10 per cent only 38 per cent complete the first year in college. This improved prediction, resulting from the standardization of

grading, would therefore be of great value in advising such students. Other studies have shown that the class rank of a student in addition to his high school grades forms a more reliable basis for forecasting success in college. Again, as has been shown earlier in this study, high school grades combined with some standard intelligence test will materially increase the correlation coefficient and consequently will give a better prediction of the quality and quantity of work that a student will do in college.

In the light of the results found and comparisons made in this study we are lead to the conclusion that the standardization of a grading system would be of great value for the prediction of success in college.

ACKNOWLEDGMENT

This study was undertaken at the suggestion of Dr. J. C. Peterson, Professor of Psychology at Kansas State College of Agriculture and Applied Science. I am indeed greatly indebted to him for his kindly criticisms and valuable counsel in the organization of the material for this thesis.

A word of appreciation is also due Principal F. V. Bergman of Manhattan High School and to the Registrar of the College for the privilege of having access to the records necessary for this study.

REFERENCES

1. Ayers, L. P.
1920. A Shorter Method for Computing the Coefficient of Correlation. Journal of Educational Research, Volume I, March.

2. Bentley, Bancroft
1922. The Relative Standing of Students in Secondary Schools on Comprehensive Examinations and in College. School Review, Volume XXX, Number 2, February.

3. Charters, W. W.
1929. Predicting Success in College. Educational Research Bulletin, Number 9, May; Ohio State University.

4. Colvin, S. S.
1921. Educational Guidance and Tests in Colleges. Journal of Applied Psychology, Volume V, March.

5. Dearborn, W. F.
1909. The Relative Standing of Pupils in the High School and the University. University of Wisconsin, Bulletin Number 312.

6. Lincoln, E. A.
1917. Relative Standing of Pupils in High School, in Early College, and in College Entrance Examinations. School and Society, Volume V, 417-420.

7. McPhail, A. H.
1924. The Intelligence of College Students. Warwick and Jork, Inc., Baltimore.

8. Odell, Charles W.
1927. Predicting the Scholastic Success of College Freshmen. Urbana University of Illinois.

9. Peterson, J. C.
1920. Report upon the Engineering Intelligence Tests. Kansas State College of Agriculture and Applied Science.

10. Roberts, A. C.
1922. Objective Measures of Intelligence in Relation to High School and College Administration. Educational Administration and Supervision, Number 9, December; University of Washington.

11. Rugg, H. O.
1917. Statistical Methods Applied to Education. Houghton Mifflin Company, Chicago.

12. Scates, D. E.
1924. A Study of High School and First-year University Grades. The School Review, Volume XXXII, p. 182.

13. Strickland, V. L.
1927. Relative Levels of Intelligence in College and High School. Kansas State Agricultural College Bulletin, Volume X, November.

14. Terman, L. M.
1921. Intelligence Tests in Colleges and Universities. School and Society, Volume XIII, Number 9, December.

15. Thorndike, E. L.
1923. Intelligence Examinations for College Entrance. Journal of Educational Research, Volume I, May.

Other References

Holzinger, C. J.
1928. Statistical Methods for Students in Education. Ginn and Company, Chicago.

Pintner, Rudolph
1931. Intelligence Testing, Methods and Results. Henry Holt and Company, New York.

Proctor, W. M.
1920. Psychological Tests as a Means of Measuring Probable School Success of High School Students. Journal of Educational Research, Volume I, April.

Wood, Ben D.
1923. Measurement in Higher Education. World Book Company, New York.