A Ten-Year Study of Iowa Placement Tests

H. H. ARMSBY

If, as the statisticians tell us, our population curve has commenced to flatten out, and if our college enrollment is due to reach its maximum in the very near future, it seems to me the problems of admissions and guidance will become increasingly important. If their prophesies are correct, and I believe they are, they mean an increasing responsibility on college officials to make every possible effort to get each student into the sort of work for which he or she is best suited-in other words to make the most efficient use they can of the limited amount of human material coming to their educational factories.

At the 1932 meeting of the American Association of Collegiate Registrars I presented a paper under the title "Can Success or Failure in Engineering Colleges be Predicted in Advance?" That paper contained a preliminary report of a study which was started in 1925 in which scores made by ten entering Freshman classes on a set of placement tests were to be compared not simply with the work of their first semester, but with their entire college performance. The ten classes, consisting of about 1300 students, have taken the tests; and seven classes, containing about 800 students, had completed their four college years up to the time the tabulations to be presented in this paper were prepared. The numbers are sufficiently large to be significant, and the results of the study should be of particular interest to Registrars, who more and more are assuming the duties of admissions and guidance officers. Although this study was conducted in an engineering college, the methods of work and the results and conclusions are in general applicable to any college.

In the fall of 1925 the Missouri School of Mines began the use of the Iowa Placement Tests as a basis for sectioning freshman classes according to the ability of the students. We have continued their use ever since and are pretty well "sold" on the sectioning ideaour faculty members in general believing that we have made distinct educational gains, at least in the high sections. We have used the Chemistry Aptitude, English Aptitude and Training,

Mathematics Aptitude and Training tests of the Iowa series, and a Drawing Aptitude test prepared by Dr. C. V. Mann, our Professor of Engineering Drawing and Descriptive Geometry.

At the time we began the use of placement tests, I conceived the idea of carrying much further than one semester a study of the students who took the tests, with the idea of attempting to discover whether or not there was any relationship between the results of the tests and general success in college beyond the Freshman year. In other words, it was my idea to ascertain whether or not the tests really were selecting the superior and inferior students, and if so which tests gave the most valid predictions. This study has been carried forward throughout the succeeding years, and in this paper I will give a very brief and partial report of some of my findings.

First, I wish to outline briefly my data and method of work. I compiled an abstract of the record of every student who has taken these tests at the School of Mines, a total of about 1300 boys. The record shows the score on each test, the number of semesters the boy has been in college, and his total scholastic record. For each test I made a tabulation of all the scores for the entire eleven years, and have then divided this range of scores into tenths; and in my study the figure I use for a boy's score is his tenth rather than his actual score on the test. Then under each separate test I grouped the members of each class according to these tenths. In other words, I combined the records of all those members of the class of 1929 who in the Chemistry Aptitude test made scores which placed them in the first tenth of all the students taking this test in the eleven year period, and I computed average figures for this small group; and the same for those ranking in the second tenth, the third tenth, and so on; and the same for each class and each test. For each group I listed the following information: the number in the group and the per cent of the whole class; the number graduated and the per cent of the group; the number of semesters completed; the number of credit hours scheduled; the number passed, and the grade points earned.[1] I also tabulated the number and percentage of each group graduating in eight semesters, the num

[1] In computing grade points each hour of E grade carries three grade points, each hour of S carries two, each hour of M carries one, I grades carry no grade points, and each hour of F carries one negative grade point. The student's average grade is found by dividing his total grade points by the total number of credit hours he has passed.

45 ber and percentage who had no failures, the number and percentage dropped from school for failures, and the number and percentage receiving honors on graduation. Then I calculated for each group the percentage of their work passed and an average grade for the group. These figures have been tabulated for each class under each test, and a grand total for each test made for all the students in the seven classes which had graduated up to the time I started the compilation (about 800 students). I plan to complete the study for the remaining four classes.

I then prepared studies of combinations of tests by a method which schools of Education would probably denounce as very unscientific, but which strikes me as being very simple and as possessing all the accuracy warranted by the data. Instead of going through the long and laborious process of computing multiple correlations, standard deviations, etc., to combine the results of two or more tests, I have combined them by the simple expedient of taking the arithmetical average of each student's tenths on the two or more tests to be combined. For instance, if John Smith ranks in the third tenth on the English Aptitude test and in the fifth tenth on the English Training test, I call his rank on the combined tests 4. I don't believe that our original data contains enough precision to warrant all the elaborate mathematical computations which some of our friends in the School of Education would have us go through.

No account has been taken in this study of such factors as sickness, financial difficulties, family troubles, transfer to other schools, etc. I have simply compared Placement examination scores with actual total performance in college, making no attempt to make allowances for special circumstances. In other words, I am dealing with facts and not with what might have happened.

I shall not try to present in this paper all the details of this study. The amount of data is very large and to include all of it would make a paper of unpardonable length. I therefore include only one of my tables of data, Table I, which contains the data for all of the graduating classes under the combined tests for Chemistry Aptitude, Drawing Aptitude, Mathematics Aptitude, and Mathematics Training.

It will be noticed that the ten groups are not equal in size. This is due to a number of reasons, the chief one in all probability being my method of averaging tests; but regardless of the size of the groups, the data gives ten points for each of the curves which

are shown in Plate I. I include this table merely to show how the data is assembled and to give a general impression of what the figures look like. It will be observed that in all the columns except one the numbers decrease as we go down the columns and that the difference between groups 1 and 10 is very large and significant. The one column for which this is not true is the one showing the percentages dropped for failure, and in this column . the figures in

TABLE I

COMBINED TESTS, GROUP "E"

CHEM. APT., DRAW. APT., MATH. APT., MATH. TRAIN. ALL GRADUATING CLASSES

SCHOLASTIC RECORD — PER CENT Or GROUP

aCF DF '~Li m

6c y ~ b Am a ~ ~ ~ W

°°m m a c7 E.' W s°. v

N z H czw

o GO G~° G

o za m kq c7 b~ c c~ B A a

```
1 46 6 59 6.4 132 129 229 98
1.78 52 54 0 44 2 99 12 42 4.8 98 96 143 98 1.64 35 49 4 21 3 83 10
49 5.8 118 114 142 97 1.25 42 46 5 13 4 96 12 39 5.1 102 97 107 95
1.10 31 38 14 5 5 108 14 47 5.4 110 104 112 95 1.08 39 26 12 5 6
87 11 37 5.1 95 90 90 95 1.00 27 30 17 3 7 96 12 24 3.7 75 67 64 89
0.96 20 16 27 2 8 104 13 20 3.4 65 59 49 91 0.83 16 13 23 0 9 70 9
7 2.2 40 31 17 77 0.55 7 6 47 1 10 11 1 0 0.9 14 5 -7 36 -1.00 0 0 55
0

800 100 35 4.5 90 85 98 94 1.15 29 29 15 9
```

crease as we go down instead of decreasing, which is, of course, what would be expected. The average student in group 1 stayed in school seven times as long as his classmate in group 10. He attempted almost ten times as many credit hours, and passed 26 times as many, having passed 98 per cent of his attempted work, as compared to only 36 per cent of attempted work passed by the boy in group 10. The boy in group 1 earned an average grade only a little short of an S (B on the conventional A, B, C, D grading system), while the boy in group 10 made an average grade of failure. Slightly more than 50 per cent of the boys in group 1 graduated on time and none of them were dropped for failure, while in group 10 the

47 figures are almost exactly reversed. Nearly half of the members of group 1 graduated with honors as compared to none of the boys in group 10.

Plate I shows graphically the data contained in Table I. These curves were plotted by laying off on the base line the ten groups, starting with group number 10 at zero, plotting the ordinates on the center line of each group, and drawing a smooth curve to average the ten points. I might say that very little averaging was required and that every curve shown here very closely fits its ten points.

It can easily be seen from these curves that this combination of tests certainly did pick out the very good students and the very

TABLE II

COMBINED TESTS, GROUP "E"

CHEM. APT., DRAW. APT., MATH. APT., MATH. TRAIN. ALL GRADUATING CLASSES

Summary of High and Low Groups

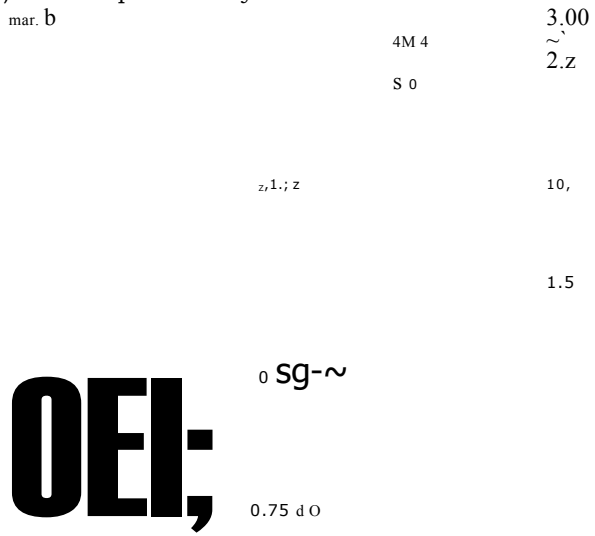
GROUP	SEMESTERS COMPLETED	AVERAGE GRADE	PER CENT Or GROUP GRADUATED	GRADUATED ON TIME	DROPPED FOR FAILURE	RECEIVING HONORS
Highest 10%	5.9	1.76	58	51	0	40
Highest 20%	5.8	1.72	53	45	2	25
Average	5					
Lowest 20%	2.8	0.71	14	12	38	1
Lowest 10%	2.1	0.56	8	7	45	0

poor ones. A comparison between the values of the lowest 10 or 20 per cent of the class and the highest 10 or 20 per cent shows that on the basis of any of the criteria considered the boy in group 1 is a very much better scholastic risk than the average of his class, and that the boy in group 10 is a very much poorer risk than the average.

To aid in interpreting these curves I have prepared the next table showing in convenient form the ordinates of the various curves at the center line of the lowest 10 per cent, the lowest 20 per cent, the highest 10 and 20 per cent, and the average of the entire class. This table (Table II) shows that the student who was in the top tenth of his class on this combination of tests had

while 17 per cent of all the 800 were dropped, and almost three times that proportion of the low tenth boys.

Two per cent of all the 800 boys graduated with honors-none of whom came out of the low tenth. But 40 per cent (or 20 times the average) of the top tenth boys were on the honor lists.



GiEMMPT, URftw.AVT.,MRT1I.EIpT.MaTM.TRAttt ALL GLZHDUffinG GLA55E5.

V

Evidently this combination of tests really did select the very good and the very poor students. It should be borne in mind that this study compares total college performance with a set of tests lasting less than seven hours, given before the students attended any college class, and that we are making no allowances for sickness, financial troubles, etc.

49 Having progressed thus far, it seemed quite natural for a Registrar and an ex-engineer to try to find some way of **still** further condensing his results. I wanted to find some one number to express all these criteria-for a quick and easy comparison of students and of tests. I believe I have found such a number-or invented it, and I wish to explain it briefly.

I believe that there are three standards by which the achievement of different students or groups of students may be compared-by quantity of work, by quality of work, and by speed. My Scholastic Index consists of three factors, one for each of these three standards, for I believe all three must be considered. Fortunately we have easily available very good measures of all three.

The factor measuring quantity of work consists of the total credit hours passed by the student divided by the number of hours required for graduation, so that for the student who exactly meets graduation requirements, the factor is 1; for the student who exceeds requirements, the factor is greater than 1, and for the student who falls short of graduation requirements, the factor is less than 1.

The second factor in my Index (quality) is the student's average grade divided by the average grade required for graduation-so that for the student who exactly meets graduation requirements this factor becomes 1; for the superior student it is greater than 1; and for the inferior student it is less than 1.

_

The third factor-time-is 8 divided by the number of semesters completed by the student. Certainly the boy who can graduate in 7 semesters deserves a higher rating than the boy who requires the standard eight semesters; and the boy who requires nine or ten semesters to graduate deserves a lower rating. In the case of the student who remains in college only two or three semesters this factor means that we are assuming for the purpose of this study that had he continued in school he would have continued to do the same calibre of work that he did while he was in school. This is probably as reasonable an assumption as we could make, and certainly is better than merely omitting him from the study or omitting the time factor.

Expressed as a formula, we then have; Quality

Factor= A = Average Grade

$$\text{Quantity Factor} = B = \frac{\text{Credit Hours Passed}}{\text{Graduation Requirements}}$$

$$\text{Time Factor} = C = \frac{8}{\text{No. of Semesters Completed}}$$

The Scholastic Index is the product of all three factors, or

Scholastic Index =AXBXC

The Scholastic Index has a value of 1 if the student has just exactly satisfied graduation requirements; a value of greater than 1 if he exceeds them; and a value of less than 1 if he falls below the requirements.

Fortunately for me, this somewhat complicated-looking formula reduces to a very simple expression. The product of the numerators of the first two terms (Average Grade times Credit Hours Passed) is the number of Grade Points earned. The product of the denominators of these two terms (Graduation Average times Hours Required for Graduation) is the number of Grade Points required for graduation, which in our school is 120. So our expression becomes

$$\text{Scholastic Index} = \frac{\text{Grade Points}}{120} \cdot \frac{8}{\text{Sem's Completed}}$$

$$= \frac{\text{Grade Points}}{15 \times \text{Sem's Completed}}$$

Plate III shows graphically the Scholastic Indices for each separate test and for each combination of tests studied. I computed the Scholastic Index for each student and for each of the groups I mentioned, and these graphs show the average values of the Scholastic Index for each percentile of the entire 800 boys who entered in the 7 classes which comprise this part of my study. The curves were constructed in the same manner as the previous ones, and here again little averaging had to be done-each curve actually passes through nearly every one of its ten

would h.
from .•
about .
20th .• o
..o.•:1 . I

COMBINED there is very little difference between any of the curves, but at the two ends there is a considerable difference. We cannot expect such tests to differentiate very sharply among the large group of boys in the middle of a class, but we do expect them to pick out the very ... ,

IV.4-00. SINGLE TESTS

to 4<> PLATE

good students at the top and the very poor ones at the bottom; and middle, the more sharply the curve deviates from the average line through the .• .•

The English tests give the poorest predictions, and the combina-

drawing test is a very close second to it. Among the combined tests the combination of chemistry, drawing, and the two mathematics tests gives the best results, being a little better than the combination of all tests. (This is why I use this combination in this paper.) The combination of drawing aptitude and the two mathematics tests is a very close second, and it, too, is better than the combination of all tests.

Table III shows the Scholastic Indices for the upper and lower percentiles of all single tests and combinations of tests. It is obvious that the Mathematics Training Test is the best single test. Of the

TABLE III

SWOLASTIC INDICES
ALL GRADUATING
CLASSICS

	SINGLE TERTs				COIOINED TESTS						
	DRAW.	ENGL. APT.	ENGL. APT.	MATH. ,TIN.	MATH. APT.	TEsTs TESTS III &IV	Trims TESTS V & VI	TESTS I & II	ALL TESTS a &b	TESTS I & II &b	CHEx. II &b	TESTS
Gaoue	Arr.				TRAIN.							
I II III IV V VI a b o d e f g Highest 10~J_0'		2.00	2.20	1.78	1.86	2.05	2.20	1.76	2.13	2.25	2.10	2.45
2.30 **2.20** H"Ighest20%		1.83	2.00	1.70	1.75	1.88	2.00	1.72	1.95	2.03	1.92	2.12
2.03 **2.05**												
Average		**1.35**	1.35	1.40	1.32	1.27	1.25	1.35	1.23	1.26	1.30	1.30
1.30 **1.30**												
Lowest 20,%$_y$		**0.90**	**0.97**	**0.97**	**1.06**	**0.83**	0.65	**1.00**	**0.70**	0.88	**0.72**	**0.76**
0.79	**0.77**											
Lowsstiw	**0.80**	**0.90**	**0.90**	1.05	0.72	**0.45**	**0.95**	0.25	0.70	**0.45**	**0.45**	**0.45**
0.46												

combined tests, combination "e" is much the best on the upper end, while d, f, and g are equal to it and b is slightly superior to it on the lower end. The English tests seems to be the least predictive, in fact including them in our combinations seems to lower the predictive value. Perhaps the explanation is that the very best students in English have a tendency to be more interested in other fields than in Engineering.

Again I wish to emphasize the fact that Scholastic Index measures entire college performance-and that this whole paper is based, not on theory, but on recorded facts.

Table IV shows, for those interested in such measures, some correlations. I divided the individual Scholastic Indices into tenths just as I did the test scores, and computed the correlations between the tests and the Scholastic Indices, using ten divisions of each. These correlations are shown in Table IV and it will be observed that the values are very good when English is omitted

And for those who put little faith in correlations I have computed the figures shown in the last two columns. Perfect

3 on the tests would rank in the fourth tenth of the Scholastic Indices -and the same for each tenth. The figures in my last two columns
show the per cent of cases in which the prediction is in error by two tenths, or less, and by three tenths or less. The first value is perhaps about the range of one letter in the ordinary four-letter grading system-the second is about a letter and a half.

To me these agreements are quite remarkable, in view of the character of our data, and in view of the many unknown factors which are of necessity omitted.

TABLE IV

CORRELATIONS OF TESTS WITH SCHOLASTIC INDICES ALL CLASSES

TENTS	DESIGNATION	NUxBER Or CASES	CORRELATION (r)	PREDICTION 1N ERROR 2 DmsioNs OR LESS	PREDICTION IN ERROR 3 DIvisima OR LEss
Chemistry Aptitude	I	1246	0.50	64%	78%
Drawing Aptitude	1I	1078	0.50	64%	78%
English Aptitude	111	1320	0.39	59%	73%
English Training	I V	1319	0.39	57%	71%
Mathematics Aptitude	V	1320	0.53	67%	78%
Mathematics Training	VI	1333	0.57	68%	81%
Engl. Apt. and Engl. Train.	a	1310	0.43	61%	77%
Math. Apt. and Math. Train.	b	1318	0.60	70%	83%
Chem. Apt. and Draw. Apt.	c	1000	0.53	69%	82%
Chem. Apt., 2 Engl. Tests, 2 Math. Tests	d	1160	0.59	70%	83%
Chem. Apt., Draw. Apt., 2 Math. Tests	e	975	0.61	73%G 84%G	
Draw. Apt. and 2 Math. Tests	f	1062	0.58	73%	84%
All Tests	9	956	0.62	72%	85%

I make no claim for any high degree of mathematical accuracy at any point in this study. When it is remembered that the placement examinations occupy a total time of less than seven

> effective general intelligence test and enable us to pick out with striking accuracy, before entrance to college, the very good students and the very poor ones. It is not claimed that these examinations are perfect, or that they cannot be improved upon, but as they stand now they do a very satisfactory piece of work. Perhaps some other group of tests would give better predictions in other colleges, but surely this study demonstrates that very reliable predictions of scholastic success are entirely feasible.
>
> If such predictions can be coupled with tests of interest, such as the Strong Interest Test; and with a program of giving to high school students real first-hand information as to the work of the various professions and the aptitudes requisite to success in each; then, and only then, will we be in a position to make our educational factories run at their highest efficiency.

I see a vision of some future day when every high school boy and girl will have an opportunity to learn about engineering, or law, or journalism, or nursing, or any field in which he or she is interested, through lectures, printed material, and consultation with members of that profession and with other properly informed advisers.

I see the students who are interested in any profession taking a set of tests to determine whether they are adequately prepared to enter the college course leading to that profession with a reasonable hope of succeeding there; and being either encouraged to enter, urged to secure more preliminary training, or advised to enter some other field.

I see the various colleges, as a result of such a program, admitting only the well qualified students, and hence able to place their work on a much higher plane than is now possible.

And finally I see each student placed in the work for which he is best qualified, and hence utilizing his natural abilities where they can do the most good to himself, and hence to society.

Will my dream come true? It can if the professions, the colleges, and the high schools will all work together for its realization. It isn't going to "just happen." If it comes to pass it will be as the result of long, careful, sympathetic co-operation between the three groups. If it comes to pass we Registrars will of necessity play a very prominent part in the program. I believe the high schools are ready-are we?