UC-NRLF
$B 17 537
LB
3063
S3
YC 04195

GIFT
MAY 15 1913

Scientific Grading of College Students

RAYMOND W. SIES
PROFESSOR OF SCHOOL ADMINISTRATION
SCHOOL OF EDUCATION, UNIVERSITY OF PITTSBURGH

An address before the Conference of College Presidents of Pennsylvania, Feb. 29, 1912, held on the occasion of the celebration of the one hundred twenty-fifth anniversary of the University of Pittsburgh.

Reprinted from
UNIVERSITY OF PITTSBURGH BULLETIN
Vol. 8, No. 21

LB3063
S3

UNIV. OF
CALIFORNIA

SCIENTIFIC GRADING OF COLLEGE STUDENTS

RAYMOND W. SIES
PROFESSOR OF SCHOOL ADMINISTRATION
UNIVERSITY OF PITTSBURGH

Mr. Chairman:

President William T. Foster of Reed College in his recent book, "The Administration of the College Curriculum," opens an important chapter bearing upon the subject treated in this paper with the following paragraph:—

"College honors are everywhere awarded on the naive assumption that grades in college courses are distributed on a scientific basis. For many important administrative purposes we assume that an A in one course is equivalent to an A in another course; that the 80 per cent of one instructor indicates an achievement equal to the 80 per cent of another instructor. Accordingly we estimate the fitness of candidates for admission, determine eligibility for athletics, assign annually hundreds of thousands of dollars in scholarships and fellowships, award Commencement honors, elect men to Phi Beta Kappa, and confer degrees wholly or in large part on the evidence secured by merely counting the number of A's, the number of B's, and so forth, that each student has to his credit. The question is pertinent to what extent our assumption of the equivalency of grades is warranted by the facts."*

---

* Op. cit., pp. 250-251.

264332

UNIV. OF CALIFORNIA

If the answer to this question propounded by Dr. Foster is, as he intimates in his opening statement, that our assumption is without foundation, the further question immediately suggests itself whether anything can be done to bring a degree of order out of the relatively chaotic condition of affairs described. The present paper has to do with the problems raised by these questions.

Fundamental to the entire discussion are the scientific principles in accordance with which mental abilities or achievements are actually distributed among college students. It is a matter of common experience that human beings, even those of the same sex and approximate age, vary greatly among themselves. The variation is far greater in mental than in physical traits. Men differ much more widely in intellect and temperament than in height, chest capacity, or cephalic index. The same variation prevails among plant and animal species and in general throughout nature. This variation, however, does not represent so much chaos and confusion in nature (inclusive of man); it does not represent an unpredictable and unmanageable phenomenon which the scientist recognizes as merely a source of constant disturbance and irritation. For the most part the facts are quite the contrary. In fact, the variations in question are controlled in a large measure by natural law, and the scientist has equipped himself with the principles involved.

It has been found by many tests that physical traits of animals and men of a given species tend to be distributed in accordance with the distribution of a recurrent variable quantity resulting from the chance action in different combinations of a very great number of equal independent causes, each just as likely to occur in any given case as not. The ordinary graphic representation of such a distribution, known as a normal distribution, is the normal probability curve. In figure 1 we have an illustration of this curve. The abscissas represent the different magnitudes of the variable and the ordinates the frequency of occurrence of these magnitudes. It will be noted that the magnitudes cluster closely about a mean or norm and that on each side of this central tendency,

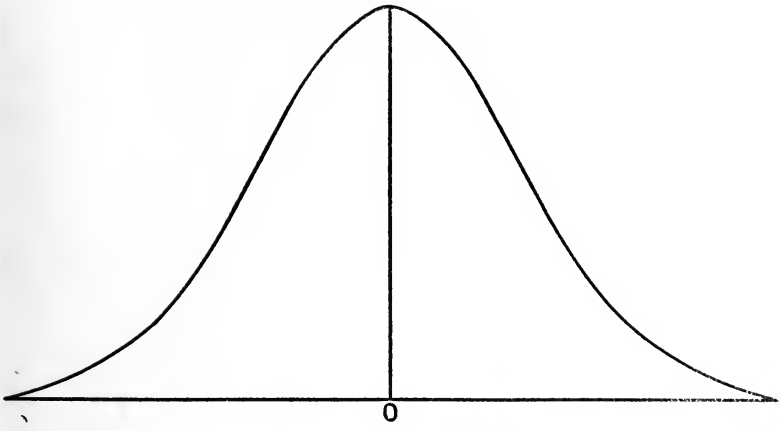
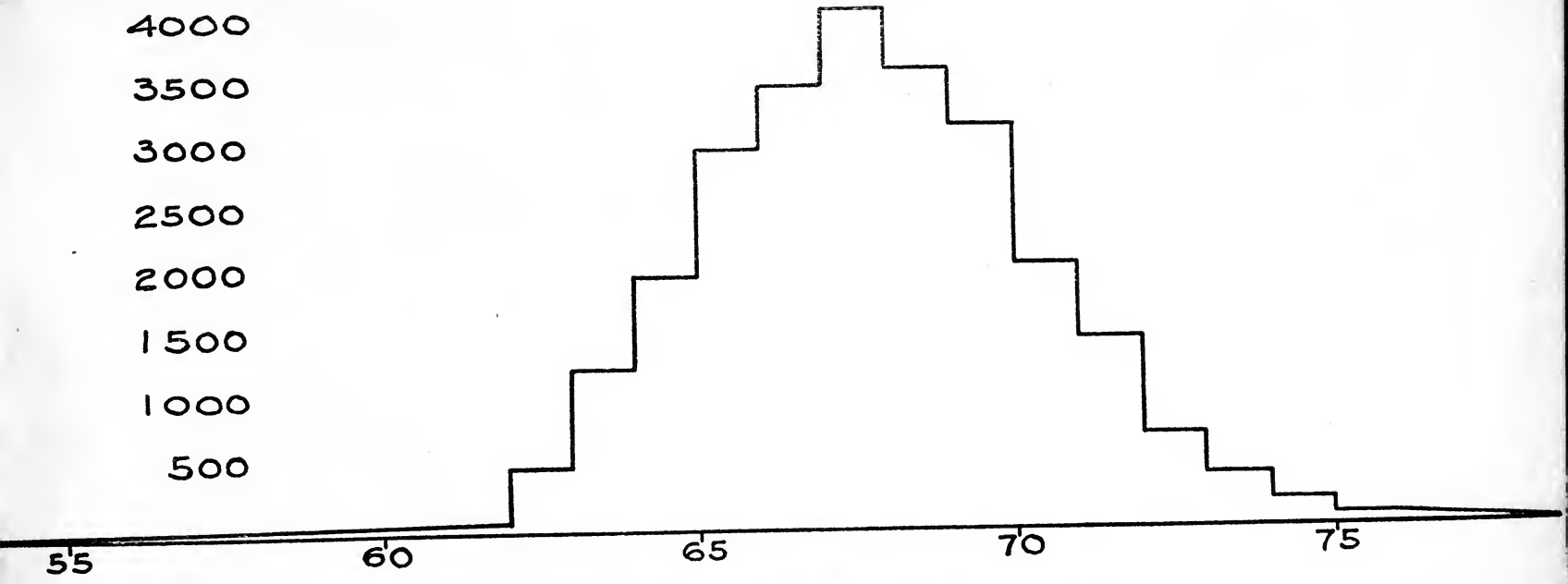


Fig. 1

indicated by the apex of the curve, the frequencies gradually decrease to zero. The relation between height and width varies in different curves, some are taller and some flatter than the one before us; but they all belong to the same species and are governed by the same mathematical equation. This equation in its simplest form is given in the figure. The area between the curve and the zero abscissa is known as the probability integral. The perfectly smooth and regular curve would appear only if the causes were infinite in number. Consequently, the actual curves that investigators have secured have been only approximations to the mathematical forms. Figure 2 presents a curve giving the distribution of the heights in inches of 25,878 recruits in the United States army, as an example of the distribution of physical traits. It is taken from one of Pearson's works.* The abscissas represent heights, the ordinates represent frequency of heights. Such a curve is technically known as a curve of frequency, and the area between it and the zero abscissa as a surface of frequency. The approximation of the curve and surface in question to the normal probability curve and the probability integral is apparent.

There is much evidence that mental traits tend to be distributed in the same way. In the first place modern

* Vide *The Chances of Death*, by Karl Pearson, Vol. I., pp. 276-277.



Height of 25878 Recruits

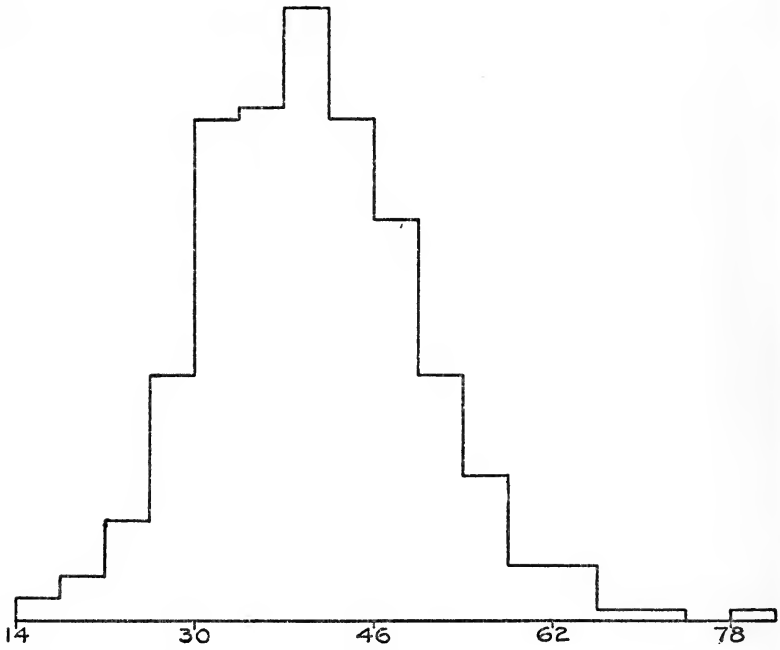
Fig. 2

physiological psychology teaches that mental traits have a physical basis in physical traits of the nervous system, and are correlated with them. There is no good reason to believe that variation in the nervous system is essentially different from variation in the muscular or other system. It would be anomalous if the parts and elements of the nervous system were distributed in accordance with a law different from the one governing the distribution of other physical traits. Granting then that variation in the nervous system is normal, we may argue on psychophysical grounds that the distribution of mental traits tends to be of the normal type. We may reasonably assume that mental traits follow the same law of variation as the corresponding physical traits. Furthermore, mental phenomena are natural phenomena, and their character as such is presumptive evidence of considerable weight in favor of the position here taken.

Statistical evidence in favor of the tendency of mental traits to be distributed in the manner under consideration is not so plentiful nor so definite on the whole as that concerning physical traits, owing, the scientist believes, to the difficulty of adequately measuring most mental traits with our present facilities. However, such evidence is by no means entirely absent. Some specific mental traits can be measured with a fair degree of accuracy, and measurements of such traits in groups of properly selected individuals have been made and the distributions worked out. When conducted by proper methods these investigations have definitely shown that the particular traits in question tend to be normally distributed. In figure 3 is presented the distribution of a relatively simple mental trait in 312 boys twelve years old. The trait in question is efficiency in perceiving A's on a page of capital letters appearing in indiscriminate order. This distribution is one of a number of the same class given by Dr. Thorndike.*

It is proper to add that it is an exceptionally good illustration of approximation to the normal distribution in the field of mental measurements, especially in view

*Vide *Introduction to the Theory of Mental and Social Measurements*, by E. L. Thorndike, pp. 46, 49, 83.



Efficiency in Perception of 312 Boys

Fig. 3

of the relatively small number of cases. However, the weight of evidence from tests of mental traits susceptible of fairly accurate measurements is very strongly in favor of the proposition that such traits are approximately normally distributed. In general, results leading to a different conclusion may be charged to defective methods, such as the mixing of types and the testing of selected groups. Causes affecting the form of distribution other than the natural forces of variation must obviously be eliminated. The individuals tested must correspond in age or maturity and in the degree of training previously received in the trait tested. They must also represent a random selection. Obviously the distribution of a mental trait in a group of geniuses or a group of idiots would not be normal.

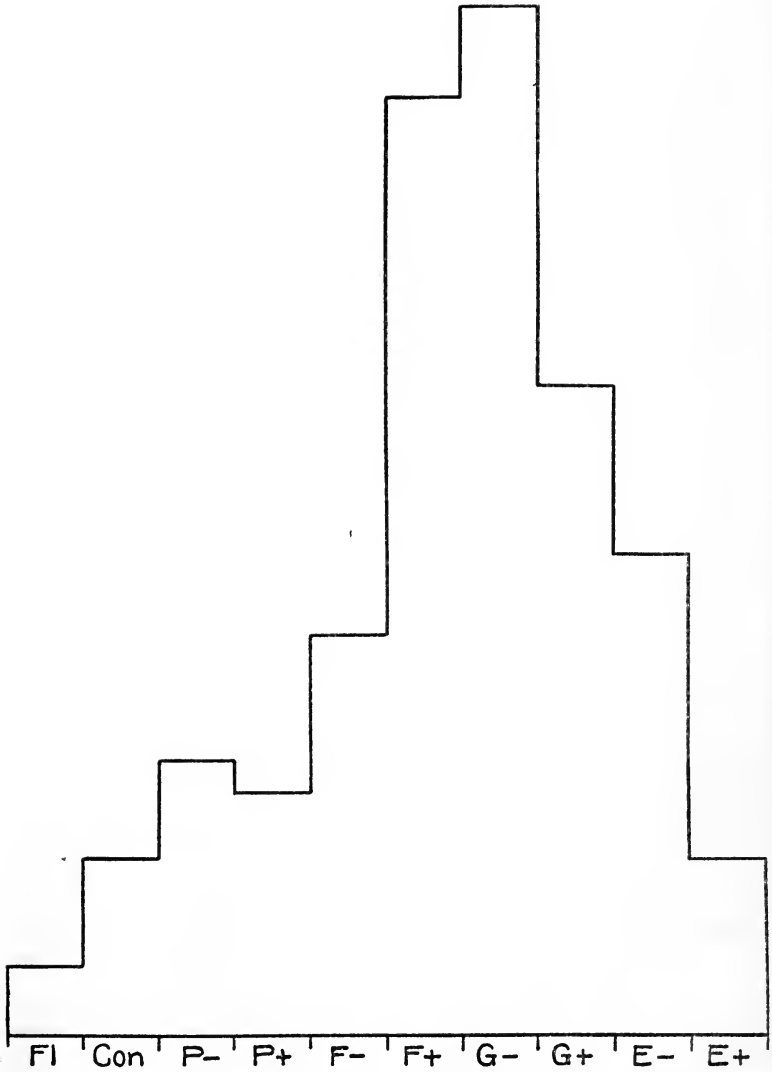
The preceding considerations with others that might be advanced constitute, in the minds of scientific men, sufficient evidence to warrant the assumption that mental traits not yet susceptible of fairly accurate measurement

tend to be distributed in the form represented by the probability curve. Experience thus far warrants the belief with some confidence that when the technique of quantitative methods in the mental and social sciences is sufficiently developed the distribution of such traits will be found to be not materially different in form from that of the traits we have been considering in the preceding pages. In this same connection Dr. Thorndike writes as follows:

"Distributions aproximating it [the probability integral] do occur very commonly in mental traits of original nature. And one will probably never be far misled by supposing that, in respect to the amount of original endowments in any trait, individuals of the same sex, race and age are distributed approximately according to the probability surface. The evidence from measurements points toward such approximation. Moreover, what is known of the physical basis of intellect and character leads to the expectation that many somewhat nearly equal factors are at work to determine the amount of any instinct or capacity possessed by men."*

We now come directly to the matter of distribution of the ability or achievements of college students in college classes, which constitutes one of the complex mental traits which we have no satisfactory means of measuring at the present time. Is this ability or type of achievement normally distributed? The grounds for assuming approximation to the normal distribution are just as strong in this case as in the case of other mental traits not susceptible of fairly exact measurement. Yet in the actual grading of college students by college teachers practically all conceivable forms of distribution are manifested. Approximation to the normal distribution is the exception. This diversity, of course, is due very largely to the absence of definite units and the inadequacy of measurement, but it is probably due more largely to the widely varying personal factor in the teachers. Each teacher has his own particular views and habits in grading. When, however, instead of the grades of individual teachers we distribute collectively the grades of a large number of teachers in different subjects, the influence of the personal fac-

*_Educational Psychology_, by E. L. Thorndike, 2nd Ed., pp. 165-166.



Grades of 12000 Students
University of Wisconsin

Fig. 4

tor is practically eliminated, since the peculiarities in the grading of one teacher are neutralized by those of another. Such a distribution may therefore be expected to exhibit a very rough approximation to the normal.

In figure 4 are exhibited about 12,000 general averages of students in the University of Wisconsin in recent years.* These averages were computed from the grades of a large number of teachers. The figure exhibits a distinctly imperfect, but yet quite recognizable, approximation to the probability integral. The relatively large divergence therefrom may fairly be charged very largely to the present unavoidable imperfections of the scale or scales wherewith ability of achievement in the classroom must be measured. President Foster in the book cited at the beginning has presented a study of nearly 12,000 different annual grades in many different subjects secured by students in Harvard College during the years 1903-4 and 1904-5.† The distribution of these grades on a scale with five divisions is similar in form to that of the general averages of the Wisconsin students.

How widely the grading of individual teachers diverges from the standard form may be judged from table I for which Professor Max Meyer of the University of Missouri is responsible. This table shows the distribution of the four grades, A, B, C, and F, in use at the University of Missouri, by forty individual teachers in that institution during a period of five years preceding 1908. For the most part the grades were given for courses in the College of Arts and Science, and therefore to a large extent the different teachers had the same students. To avoid undue publicity the teachers are designated by subjects. On the basis of the grades assigned the students graded by each teacher were divided into three groups designated superior, medium, and inferior. The first group comprised the twenty-five per cent of students ranking highest, the last group the twenty-five per cent ranking lowest, and the middle group the remaining fifty per cent ranking between the other groups. The last column indicates the total number of students graded by each teacher. The distri-

*Vide *School and University Grades, by Walter F. Dearborn, Bulletin of the University of Wisconsin, High School Series*, No. 9, pp. 43, 44, 46.

†Vide *Administration of the College Curriculum*, by W. T. Foster, pp. 251 ff., 276 ff.

bution of grades is indicated by the percentages of the total number of students receiving the different grades in the several groups. The teachers are listed in the order of their liberality of grading. Thus the first teacher assigned the grade A not only to all his superior students, but to considerably more than half of his medium students as well. He also assigned the grade B to more than half his inferior students, and few were failed. The last teacher, on the other hand, assigned the grade A to only one per cent of all his students, and over one half of his superior students received the grade C. All his inferior students

Table I

Grading of Students at the University of Missouri.

Teachers	25 Per Cent Superior Students			50 Per Cent Medium Students				25 Per Cent Inferior Students			Total Number of Students
	A	B	C	A	B	C	F	B	C	F	
Philosophy	25			30	20			13	10	2	623
Latin, I	25			27	23			19	6		130
Sociology	25			27	23			7	13	5	958
Mathematics, I	25			15	31	4			12	13	208
Economics	25			14	36			1	19	5	461
Greek	25			14	26	10			14	11	287
Latin, II	25			11	39			1	19	5	577
French	25			11	29	10			15	10	295
Political Science	25			9	30	11			16	9	592
Mathematics, II	25			7	29	14			9	15	145
German, I	25			5	39	6			14	11	586
Psychology, I	25			5	36	9			15	10	907
German, II	25			1	38	11			14	11	941
Elocution	20	5			50			6	19		917
Geology	22	3			45	5			17	8	293
History, I	14	11			42	8			19	6	779
Zoology, I	21	4			41	9			19	6	479
Psychology, II	19	6			41	9			20	5	238
History of Art	25				40	10			20	5	685
Bacteriology	20	5			40	10			21	4	263
Freehand Drawing	18	7			40	10			15	10	506
Chemistry, I	23	2			38	12			19	6	205
English, I	21	4			37	13			17	8	964
Astronomy	13	12			37	13			20	5	225
History, II	11	14			37	13			20	5	806
Zoology, II	24	1			36	14			17	8	250
German, III	22	3			34	16			12	13	441
Chemistry, II	9	16			32	18			25		21
Education	18	7			31	19			16	9	266
Mathematics, III	19	6			30	20			6	19	182
Mathematics, IV	25				29	21			15	10	380
Physiology	20	5			28	22			18	7	426
Anatomy	19	6			28	22			14	11	544
Mathematics, V	16	9			25	25			10	15	209
Engineering, I	13	12			24	26			16	9	813
Mechanical Drawing	18	7			22	28			13	12	558
Mechanics	18	7			19	31			11	14	495
Engineering, II	16	9			17	33			13	12	826
English, II	9	16			12	35	3			25	1098
Chemistry, III	1	11	13			47	3			25	1903

were failed and some of his medium students. The table shows all degrees of leniency and severity between these extremes. By reference to the last column we find that every teacher almost without exception had a sufficient number of students during the five years to warrant the belief that the distribution of actual ability or achievement in his classes was approximately normal.* The chaos and gross injustice of such grading are evident and require no comment.

It should be understood that there is nothing strange or peculiar about the state of affairs here described which formerly existed in the University of Missouri. It is simply typical of the situation at practically all colleges and universities throughout the country. The natural remedy for this condition of things with its confusion and injustice is plainly the normal or scientific distribution of students' grades which requires teachers to dispense grades according to the natural distribution of ability or achievement. A few progressive institutions, as indicated below, have recently begun breaking the way toward better things by the adoption of this device. Among these the university just mentioned must be given the leading place.

There are some persons without doubt to whom the preceding considerations will not appear convincing. Despite the above argument of numbers they will be inclined to insist that the striking and well-nigh universal differences in the distribution of grades by different teachers are to be explained primarily by the fact that some teachers and some subjects consistently draw a better grade of students than others. In considering this objection it is necessary to distinguish carefully between required and elective courses. In the first place, if the objection is sound, if variations among teachers in the distribution of grades are primarily due to differences in the quality of the students, we should expect a degree of uniformity in the distribution of the grades assigned by different teachers of required courses. This relative uniformity, however, does not seem to exist. Our experience does not bear out the assumption that it exists, and inves-

* Vide *The grading of students*, by Max Meyer, Science, n. s., Vol. 28 pp. 243 ff.

tigators have not recorded any observed differences, so far as the writer is aware, between the discrepancies among teachers in grading students in required courses and those in grading students taking elective courses.

Again, if the objection under consideration is sound, we should expect that teachers of elective courses who habitually grade higher than the average teacher, do really have superior students, and those who habitually grade correspondingly low really have inferior students. The facts, however, so far as they have been definitely ascertained, are rather the reverse. In table II are presented the results of an investigation made by Dean Ferry of Williams College upon this specific problem. The administration of the curriculum at Williams College is more than ordinarily favorable for the prosecution of such a study. In the freshman and sophomore years the studies are either prescribed or elective within narrow limits. In the junior and senior years there is practically free election. Dean Ferry computed the average standings secured during their freshman and sophomore years by the several students enrolled in all the junior and senior elective courses of the college offered during the years 1906-1907, 1907-8, and 1908-9. On the basis of these

Table II

Quality of Grading and of Students at Williams College

Teachers	Quality of Grading	Quality of Students	Teachers	Quality of Grading	Quality of Students
1	120	− 7	16	41	2
2	114	−30	17	41	9
3	95	−17	18	40	−14
4	89	−22	19	34	15
5	89	− 6	20	32	7
6	73	−40	21	27	77
7	66	−33	22	23	39
8	63	5	23	20	13
9	59	− 8	24	6	− 2
10	58	6	25	3	24
11	56	− 1	26	0	113
12	52	−21	27	0	113
13	50	17	28	−11	− 4
14	49	20	29	−21	39
15	42	1	30	−23	41

earlier records he improvised a measure or index of the quality of students in the several junior and senior courses. On the basis of the quality of students thus determined he further improvised a measure or index of the quality or standard of grading of each of the thirty teachers giving junior and senior elective courses. The latter measure was so formulated that a high index indicates liberality in grading and a low index indicates corresponding severity. Time and space do not permit here an explanation of the derivation of these indexes. It must suffice for the present to say that the writer has satisfied himself that, though not mathematically exact, the indexes are fairly reliable. Those desiring a detailed account are referred to Dean Ferry's report for 1910-11, in which the investigation is described.* Table II is copied from this report with changes in the order of arrangement. The data for the different teachers are here arranged in the order of the quality of grading. The indexes for teachers who were most liberal in grading head the columns, those for the teachers most severe in grading close the table. It will be observed that, whereas the indexes in column two regularly decrease, the corresponding indexes in column three irregularly increase. This means that on the whole the teachers who graded highest had the poorest students and those who graded lowest had the best students. The correlation between standard of grading and quality of students is distinctly negative. By various methods of calculation the coefficient of correlation has a high negative value. What is the explanation of these facts? The chief explanation is not far to seek. The negative correlation between quality of grading and quality of students results very largely from the well known proclivity of inferior and indolent students to seek snap courses and the tendency of able and earnest students to seek the more substantial courses. A snap course may be defined as one where for a given effort or achievement a relatively high grade may be gained, owing to low standards of qualitative or quantitative requirements, or both. To undertake to demonstrate here that such courses are studiously sought by some students and avoided to some extent by others would be a clear case of carrying

* Vide *Williams College Bulletin*, Series 8, No.5 (June, 1911), pp. 27-32.

coals to Newcastle. Finally in this connection it may be said that Dean Ferry's results and the conclusions therefrom are reinforced by results secured by President Foster from a study of the undergraduate history of the 4,311 men who graduated from Harvard College during the years 1886 to 1900 inclusive. President Foster divided the members of each class into two groups, those graduating with distinction and those graduating without distinction. He found that for every class, under the system of free election then in vogue at Harvard, the latter group had taken a greater proportion of their work in snap courses than the former. In the case of most of the classes the difference in question between the groups was definite and marked.* The correlation between quality of grading and quality of students as the latter was ascertained is obviously positive in the case of Dr. Foster's investigations. However, this may be due entirely to the fact that the quality of the students was necessarily determined by their standings in elective courses rather than in required courses as at Williams College. Doubtless many students gained distinction by selecting snap courses, while others lost that honor by selecting heavy courses. Had the quality of the students been determined on the basis of the work in required courses, the correlation might easily have been negative in this case also. That negative correlation, or even absence of correlation, between the quality of grading and the quality of the students of college teachers is sterling evidence of the need of scientific grading in colleges and universities goes without saying.

The need for a scientific distribution of the grades of college students having been presented, our next question is whether such a thing is feasible, mathematically and practically. If, as we have assumed, the abilities or achievements of college students are distributed in the form of the probability surface, there can be no question regarding the mathematical feasibility of the undertaking. All that is necessary is to scale off equal distances for the different grades to be given on the zero abscissa between the extremes of the curve and to compute the percentage

* Vide *Administration of the College Curriculum*, by W. T. Foster, pp 217 ff., 302-303.

of the surface included between the ordinates at the end of each division. These percentages will be identical for all curves having the same number of divisions on the abscissa and may be readily calculated by use of tables. These specifications having been fixed and the computations made, each teacher simply arranges his students in order of merit and assigns the different grades to approximately the proper percentages in order. A plan of grad-

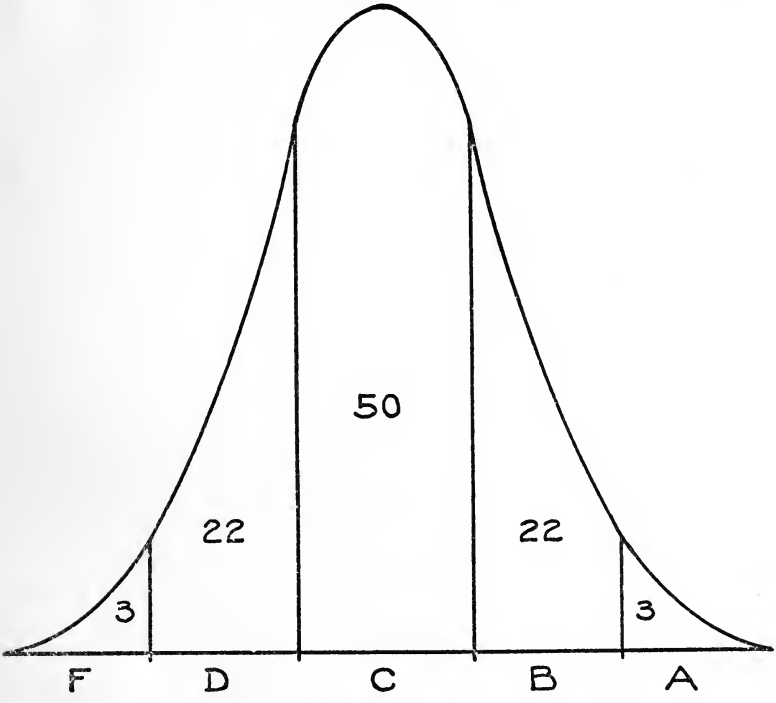


Fig. 5

ing of the type just described which has frequently been recommended for adoption by men interested in scientific grading is shown in figure 5. The five divisions on the abscissa are all approximately equal, save that the middle one corresponding to the grade C is somewhat longer in order that the middle group of students may comprise exactly fifty per cent of the entire number. In accordance with this scheme teachers assign the grade A to the three per cent of their students ranking highest in the same

class or subject, the grade F to the three per cent ranking lowest therein, the grades B and D to the twenty-two per cent ranking next highest and lowest respectively, and the grade C to those in the middle group of fifty per cent. Teachers with small classes do not necessarily follow the scheme each term or year, but rather in a cumulative way through a series of terms of years. The one criticism the writer has to offer here upon this plan is that three per cent seems too small a percentage of failures to properly motivate a considerable class of students well known to college teachers. The scale, however, could readily be adjusted to meet this objection. It may be suggested further that where there is no desire to indicate different degrees of failure the length of the portion of the scale representing failure may without violence be fixed somewhat arbitrarily without reference to the remainder. Doubtless the form of the curve in figure 5 should be slightly modified to take account of the selective influence of the college, but since nothing definite is known regarding the amount of this influence it is commonly disregarded.

We may be assisted in attempting to answer the question of the practicality of a scientific or normal distribution of students' grades by a brief description of a system that has been in operation at the University of Missouri since 1908. The system in question was introduced by action of the faculty, and its administration is in charge of a special committee of the faculty. It is definitely based upon the assumption that the distribution of ability or achievement in college classes is approximately normal. Every teacher is expected to rank the students in his classes in order of merit and then to assign the grades E and S (excellent and superior) to the twenty-five per cent ranking highest, the grades I and F (inferior and failure) to the twenty-five per cent ranking lowest, and the grade M (medium), to the remaining fifty per cent between. At present the distribution of the grades E and S and I and F among the groups of students ranking highest and lowest respectively is left to the individual teachers. The committee on grading after the close of each semester publishes a statistical table showing the character of the

grading of each teacher for the semester and since the inauguration of the present system. This table is circulated among the faculty. Teachers whose grading deviates markedly from the standards established are called to account by the committee and asked to justify their failure to conform. The grading of teachers of small classes is expected to conform to the standards only when taken through a series of semesters or years. This new system has very largely eliminated the diversity of practice in grading at Missouri, which is shown in table I. After a test of more than two years the system was declared to be fairly successful by Professor Meyer and to be becoming more fully so.* A similar but less elaborate system was introduced at the University of Iowa in 1910. Still another similar system is being introduced at Reed College during the present academic year. Other institutions will not be slow to follow. Given a faculty on the whole really desirous of improvement and with conditions carefully controlled, as at the University of Missouri, and there is no good reason why the distribution of students' grades on scientific principles should not be practically feasible. The great difficulty is to overcome the inertia of past custom and individual freedom. The actual distribution of the grades according to the new plan is easy. The gain in accuracy, standardization, and justice are worth the effort involved in a change to the new system many times over.

Let it not be imagined, however, that the scientific distribution of grades will eliminate all the difficulties and discrepancies encountered in the administration of the grading of college students. The introduction of such a plan of grading in any institution must be considered a big stride forward, but after its adoption there will still remain possibilities of improvement in the field under consideration. Thus the wide variations among small college classes in quantitative standards of requirements and accomplishment will not be very greatly affected by the new plan. It is not too much to say that some college teachers now require three and four times the amount of

* Vide *Experiences with the grading system of the University of Missouri*, by Max Meyer, Science, n. s. Vol. 33, pp. 661 ff.

a student's time and energy to earn a given amount of credit toward graduation that some other teachers in the same institution demand or accept. Snap courses and the opposite variety do abound, and the latter are just as unjust and just as conflicting with correct standards as the former. This is a sad state of affairs for academic standards, but something besides correct distribution of grades is needed to adequately remedy it. The requirement that a teacher giving a snap course or the opposite distribute his grades in the form of the probability integral will of itself by no means raise or lower the standards to a proper level, though without doubt it will ordinarily contribute something to that end. Obviously the new distribution of grades could be made mechanically with no change whatever in the work of the students. Fear of failure is not the only factor contributing to the maintenance of class standards. The nature and amount of the teacher's assignments and his efficiency in the classroom as a teacher are two additional factors of equal and greater importance respectively. However, despite the relative incapacity of the scientific distribution of grades to standardize amount of work in college classes, such distribution does render the grading in the variant courses under consideration distinctly more accurate and just, or rather, less inaccurate and unjust. This comes from the fact that in general teachers of snap courses are required to lower their grades and teachers of the opposite type of courses are correspondingly required to raise their grades, which is all well and good so far as it goes.

An extended discussion of other means and methods of quantitatively standardizing requirements and achievement in college courses might be undertaken here, but such a discussion would carry us far beyond the limits of the present subject. The problem of such standardization is both important and extremely difficult. The solution involves first of all full recognition by teachers of the desirability of greater uniformity in quantitative standards and concerted faculty action toward that end. This goes without saying. Another effective factor in the solution of this and other college problems would be competent supervision of college teaching. It is realized that

the mere suggestion of such a thing will be considered heresy by many, but no one familiar both with college teaching and with the effects of skillful supervision of teaching in elementary and secondary schools can doubt the potential efficacy of supervision in the college. Supervision of instruction is distinct from administration. The college administrator is found everywhere, but the supervisor nowhere. One of the important duties of a supervisor of college teaching would certainly be the standardization of requirements and accomplishment among the different classes.

Another very important advance in the grading of students which the normal distribution of grades can in no way bring about is what is known as credit for quality. This is a device gradually coming into favor for graduating the amount of credit received in a course according to the quality of work done as well as by the number of hours per week the class is in session. There is a very wide difference between the actual achievement of a student who secures the required number of credits for graduation with an average grade of D+ or C— and that of the student who gains the same number of credits with an average grade of A—, let us say. It is probably not too much to say on the basis of the usual values of these grades that the actual achievement of the latter in the thing the college stands for is at least twice that of the former; yet both receive the same degree and the same diploma. A student who is obliged to leave at the end of his junior year with an average grade of B+ or A— to his credit is certainly as much entitled to his degree on the score of actual attainment as the one who finishes in the regular way with an average grade somewhere below C. From these considerations both the justice and the standardizing value of credit for quality are manifest. In addition it has a distinct educational value corrective of unfortunate habits in many students in that it places a premium on thoroughness and penalizes superficiality.

The plan of credit for quality seems to have been first proposed by President Hyde of Bowdoin College in a magazine article ten years ago.* It was seriously advocated by Professor Cattell of Columbia University in

* Vide *Outlook*, Vol. 71, pp. 886-889.

another article a few years later.† In 1905 the plan was put into operation at the University of North Dakota, but it was abandoned in most respects about six years later. The scheme was adopted in 1905 at Columbia College, where it has since been in operation. In 1908 credit for quality was introduced at the University of Missouri in connection with the plan for the distribution of grades in the form of the probability curve. It is now being introduced at Reed College. The schemes at these different institutions are essentially alike. They provide that students doing superior work in a course shall receive more than the normal amount of credit toward graduation and, what is equally important, that those doing inferior work shall receive less than the normal amount of credit. Thus at Missouri where the five grades E, S, M, I, and F are now given, as already explained, the grade E carries thirty per cent additional credit, the grade S fifteen per cent additional credit, M carries the normal credit, I twenty per cent less than the normal, while F signifies no credit. The other schemes differ in various details from the one here described.

There can be no doubt that under certain right conditions credit for quality is an extremely valuable feature in the grading of college students. In fact, it is an essential factor in thoroughly scientific grading. Though the correct scale of credit for quality has not been determined, the experience of the future may be trusted to yield an approximation to it. The right conditions referred to are the scientific distribution of grades and the quantitative standardization of requirements, so far as possible. The first of these is far the more important. Without scientific distribution of grades credit for quality is apt to produce very disastrous effects by greatly multiplying the evils of snap courses. Not only are high grades awarded for relatively inferior or little work, but these unearned grades are weighted with additional credit. Students are given two strong motives instead of one for taking such courses, and teachers seeking popularity and large classes are doubly tempted to bait the students. If the negative correlation between the quality of the stud-

† Vide *Pop. Sci. Mo.*, Vol. 66, pp. 375-378.

ents and the quality of grading of teachers offering elective courses which was found to prevail at Williams College is a general phenomenon throughout the country, the general introduction of credit for quality without the safeguard of scientific distribution of grades would be nothing short of an educational calamity. Required distribution of grades in the form of the probability curve, however, by confining within proper limits the proportion of each grade awarded, would not only render credit for quality safe, but would also render it highly effective in accomplishing the results desired from it. The scheme of credit for quality in operation for a number of years at the University of North Dakota was abandoned because of precisely the evils we have been considering.* It is distinctly in point to add that at this institution the scheme was not safeguarded by a correct distribution of grades. At the University of Missouri and Reed College credit for quality is properly safeguarded in this manner. The scientific distribution of grades, however, does not eliminate snap courses, as has been indicated. Low standards are not necessarily changed by it. Snap courses retain some of their power to work ill in connection with a scheme of credit for quality in spite of normal distribution of grades. A few moments of reflection will make this clear. Therefore a maximum of uniformity in class standards, correct of course, is another condition which should be realized as a basis of credit for quality, as well as for other good and sufficient reasons.

Scientific distribution of grades, standardization of requirements, and credit for quality all have more or less important bearings upon the troublesome problem of properly restricting the number of hours of work which students are permitted to carry. If left to themselves a large proportion of students would register for course after course almost without limit. This tendency under present conditions seems to require some artificial check. It is recognized that without some such limitation under our present clumsy systems of testing and grading achievements in college classes, many students by skilfully choosing courses and complying only with the minimum re-

* Cf. *Administration of the College Curriculum*, by W. T. Foster, p. 247.

quirements in courses selected could easily fulfil the formal requirements for graduation in a scandalously short time. Hence by use of the device under consideration we legislate in effect that students must remain in college the usual time whether they employ the additional time to good advantage or not. Some students thus required to remain in college longer than they otherwise would do improve the extra time sufficiently to warrant the delay in completing the requirements; others do not. Another object in establishing the limitation in question which is much more sound fundamentally than that just considered is to shield the thoughtless and over-confident students from superficiality and failure. On the whole arbitrary limitations of the number of hours of work students may carry has been far from a satisfactory solution of the problem raised above. The significance of the problem in this connection arises from the fact that the various features of scientific grading named at the beginning of this paragraph would contribute very definitely to its solution. Obviously the stiffening of snap courses and the award of credit according to quality together would tend to regulate automatically the number of hours elected by students. Further, they would be a fair guarantee that all degrees would be really earned. These statements should be self-evident. In the degree to which scientific grading could be perfected the need of artificial limitations of the amount of work carried by students would disappear.

And now very briefly in conclusion let us connect the special problems considered in these pages with the larger problem of which they are only a part. Somewhat radical innovations in methods and system of grading college students have been suggested, but certainly in a spirit of helpfulness rather than of captious criticism. It is believed that nothing has been proposed which can not, should not, and will not in due time be carried out. The movement is now definitely under way in progressive institutions. It is but one aspect of the growing response to the widespread demands of the present day for some definiteness in educational standards in the interest altogether of increased efficiency and economy in the education of our children and youth and in the service of society through

education. These are the great objective points. We have frequently heard and read of late that the American college is under fire. So are other educational institutions and systems. The special reference in these pages to the grading of students in college does not mean that the college alone is in need of a better system of grading. The same remarks are also applicable to this phase of educational administration in secondary schools, normal schools, the university in general, and elsewhere. It is confidently believed that improvement in grading is one of the important lines along which future progress in educational standardization will occur.

RETURN TO the circulation desk of any
University of California Library
or to the
NORTHERN REGIONAL LIBRARY FACILITY
Bldg. 400, Richmond Field Station
University of California
Richmond, CA 94804-4698

ALL BOOKS MAY BE RECALLED AFTER 7 DAYS
2-month loans may be renewed by calling
(415) 642-6753
1-year loans may be recharged by bringing books
to NRLF
Renewals and recharges may be made 4 days
prior to due date

DUE AS STAMPED BELOW

JUL 17 1991

Returned to

JUN 21 1991

Santa Cruz Jitney

LD 21A–50m-3,'62
(C7097s10)476B

Gaylord
PAMPHLET BINDER
Syracuse, N. Y.
Stockton, Calif.