



A Statistical Study of the Distribution of College Grades

J. Crosby Chapman & Myra E. Hills

To cite this article: J. Crosby Chapman & Myra E. Hills (1916) A Statistical Study of the Distribution of College Grades, *The Pedagogical Seminary*, 23:2, 204-210, DOI: [10.1080/08919402.1916.10534707](https://doi.org/10.1080/08919402.1916.10534707)

To link to this article: <http://dx.doi.org/10.1080/08919402.1916.10534707>



Published online: 30 Aug 2012.



Submit your article to this journal [↗](#)



Article views: 5



View related articles [↗](#)

A STATISTICAL STUDY OF THE DISTRIBUTION OF COLLEGE GRADES

By J. CROSBY CHAPMAN and MYRA E. HILLS,
Western Reserve University

In the majority of colleges, the method of assigning grades to students' work in order to secure a fairly uniform standard in the various departments is confined to collecting in the central office the grades given in the different courses. The results thus obtained may be investigated and compared by those members of the faculty wishing to do so, or may be called to the attention of certain instructors when it seems advisable that their system of grading should be modified.

The assumption would seem to be that, with this slight check on the extremes of grading, the marks of the various departments and of the various instructors in the departments, are of equal value; that is, the grade of E from any member of department A is equivalent to an E given by any member of departments C, D, E, etc., that is

$$E_A = E_B = E_C = E_D = \text{etc.}$$

Whether this is true in a typical institution, it is the aim of this paper to investigate.

The study is based on the records of students of the Freshman classes of the years 1910-11, 1911-12, 1912-13, 1913-14, in the College for Women, Western Reserve University. It deals with Freshmen because many of the subjects studied during the first year are compulsory and on that account the principle of selection does not operate. Whereas in the higher electives where the personnel of the class is to a certain extent determined by preference, a greater percentage of the higher grades would naturally be expected than in the large unselected groups following a prescribed course; in these latter classes one might reasonably expect a normal distribution of ability.

The number of grades given by each instructor in each department for each of the four years was counted and this material arranged so as to give:—

1. The percentage distribution of grades by departments.
2. The percentage distribution of grades by different instructors in the different departments.

3. The percentage distribution of grades by years, enabling a comparison to be made of the four classes with regard to their power of gaining the grades of E and G.

In the case of each subject, percentages have been reckoned on the basis of the total number of marks given in that subject, so that where the work covered two semesters, as it did in some departments, one individual has received two marks. The total number of grades assigned, on which the percentages are based, is recorded in each case.

DISTRIBUTION OF GRADES BY DEPARTMENTS

Table I shows the percentages of grades in each department taken as a whole, over a period of four years, on the basis of the total number of marks given in that department.

TABLE I

Department	Percentages Receiving Grades of						Number of grades awarded
	E	G	F	P	D	X	
A	13.6	32.2	39.4	10.8	2.4	1.6	242
B	5.6	30.8	42.0	12.3	4.8	4.5	772
C	11.7	43.6	30.0	9.0	5.6	0.1	427
D	25.2	36.0	24.6	10.7	2.8	0.7	690
E	13.3	47.4	16.9	13.2	6.2	3.0	781
F	8.8	24.7	30.7	25.2	7.4	3.2	215
Ideal distribution	5	20	50	20	5		

The wide divergences of the various departments both from one another and from the ideal distribution which is discussed later, and which is recorded at the bottom of the table, are apparent.

The natural conclusions would be, *a*, that the grade of E varied in value among the various departments; or *b*, that the instruction was so much better in some departments as to carry a larger number of students into the higher grades; or *c*, that the mental ability necessary in the different subjects was distributed unevenly in the classes. This last hypothesis can be dismissed at once; the second supposition will later be shown to be extremely unlikely, which leaves as the only

feasible explanation for the differences in the grading of the various departments, the fact that the grade E in one department bears little if any relation to an E given in any other

TABLE II

De- part- ment	Instruc- tor	Years	Percentages Receiving Grades of						Number of grades awarded
			E	G	F	P	D	X	
A	a	1910-11	8.7	19.6	45.7	21.8	2.1	2.1	46
	b	1911-14	16.0	32.7	38.1	8.0	3.4	1.8	113
	c	1913-14	14.8	32.4	44.0	5.9	0	2.9	34
	d	1911-13	12.2	42.9	32.7	10.2	2.0	0	49
B	a	1910-14	8.5	34.5	34.5	14.0	3.6	4.9	165
	b	1910-12	2.7	33.2	42.0	12.4	5.4	4.3	186
	c	1912-14	4.1	29.4	44.2	14.3	3.3	4.7	147
	d	1912-14	14.3	7.7	59.1	6.7	5.6	6.6	91
	e	1912-13	4.7	37.2	41.8	11.6	4.7	0	43
	f	1910-14	2.9	36.2	38.0	12.2	6.4	4.3	140
C	a	1910-14	7.7	44.4	36.4	6.4	5.0	0.1	143
	b	1910-14	13.7	43.5	26.6	10.2	6.0	0	284
D	a	1910-14	28.6	39.7	21.4	8.5	1.4	0.8	494
	b	1910-13	14.6	26.1	38.2	11.4	8.9	0.8	123
	c	1912-14	20.5	27.4	24.6	26.1	1.4	0	73
E	a	1910-14	12.3	43.8	17.4	14.1	7.6	4.8	397
	b	1910-14	11.9	51.0	17.2	12.9	5.5	1.5	326
	c	1912-13	27.6	51.7	12.1	8.6	0	0	58
F	a	1910-14	8.8	24.7	30.7	25.2	7.4	3.2	215
Ideal distribution.....			5	20	50	20	5		

department. That is, we may say at once that the equation of identity

$$E_A = E_B = E_C = E_D = \text{etc.}$$

does not hold even approximately.

DISTRIBUTION OF GRADES BY DIFFERENT INSTRUCTORS IN
DIFFERENT DEPARTMENTS

In Table II the percentages of the grades given by different instructors in each department are presented. The percentages were reckoned on the basis of the total number of grades given by each instructor for the four years, or whatever portion of the four years that instructor was present in the department. The general reliability of the percentages as representing an average distribution, is indicated by the total number of grades assigned.

From these tables it will be seen that different instructors even in the same department, dealing with students of the same mental calibre, do not agree in their system of marking; for in extreme cases over a period of several years, a particular instructor will give 400 per cent. more E's than another instructor in the same department. This can be due to no other cause than a difference in a standard which is purely subjective; it is not a tenable position that such a situation can arise merely from a difference of pedagogical skill. In fact the experience of many would lead to the view that in general, the better the teacher, the lower the system of the grading.

DISTRIBUTION OF GRADES BY YEARS IN THE DEPARTMENTS

Table III shows the percentages of grades given in each department taken as a whole, in the separate years.

TABLE III

Department	1910-11						1911-12						1912-13						1913-14					
	Percentages Receiving Grades of																							
	E	G	F	P	D	X	E	G	F	P	D	X	E	G	F	P	D	X	E	G	F	P	D	X
A	9	19	46	22	2	2	15	28	45	6	3	3	14	35	28	22	1	0	12	40	35	10	3	0
B	2	41	31	15	7	4	4	33	41	12	5	5	6	32	40	14	4	4	8	42	37	5	4	4
C	11	37	37	7	8	0	14	50	22	9	5	0	11	34	34	13	8	0	9	51	29	7	3	1
D	20	36	27	10	4	3	28	40	22	7	2	1	21	37	23	15	4	0	30	30	27	12	1	0
E	14	41	21	15	6	3	11	48	13	17	8	3	16	52	14	10	4	4	12	47	21	12	6	2
F	2	18	37	39	0	4	7	16	34	31	12	0	11	29	27	22	11	0	14	36	25	11	5	9

It is interesting to note whether there was any unanimity of opinion among the departments as to the abilities of the various classes, as indicated by the relative number of high marks. This point can be investigated by a direct examination of the data, but the conclusions are shown more clearly by means of Table IV, which is constructed on the following basis:—the Freshman class which received from any particular department the highest percentages of the grades E and G, is given the rank of 1, the class with the next highest percentage the rank 2 and so on for the four classes. Thus the Freshman class of 1913-14 was awarded a larger percentage of the grades E and G than any of the three preceding classes, except in a single department where it was beaten by the class

TABLE IV

Classes arranged in rank order with respect to power of gaining grades E and G in each department; 1 = best class, 4 = worst class

Department	1910-11	1911-12	1912-13	1913-14
A.....	4	3	2	1
B.....	2	4	3	1
C.....	3	2	4	1
D.....	4	1	3	2
E.....	4	3	2	1
F.....	4	3	2	1

of 1911-12. This may be due to a tendency to give an increasingly large number of higher grades for identical work, but it is more likely to be due to a superiority of work on the part of the Freshman class of 1913-14. If this latter hypothesis is true, it would seem to indicate that the grades given by the various departments do relatively from year to year, give an indication within that department of the excellence of the class.

CONCLUSIONS

An examination of the tables reveals the great differences in the distribution of the grades in the various departments. These differences can be accounted for only by assuming that there is great disagreement as to what standard of work is represented by a particular grade. Even in the same department different instructors, dealing with similar classes, fail to grade on a common basis.

Obviously as long as the present system continues, which admits of such great variability, there can be no safeguard against too easy or too severe marking. Two possible methods of producing increased uniformity are now in use:

1. The publication of the grade statistics showing the extent to which there is a reasonably uniform standard.

2. The employment of a uniform system of grading based on the assumption of the normal distribution of ability, in large classes over a period of several years.

The first method is effective to a certain degree; by its means any flagrant injustice is prevented, for an individual member of the faculty is informed if his distribution of grades varies considerably from that of the rest of the faculty.

The second method of obtaining uniformity is now in use in a number of the institutions of the country, among which number are:—Harvard, Yale, Michigan, Missouri, etc.* This method assumes, when dealing with large classes, where there is no principle of selection in operation, that the abilities of the members of the class vary according to the normal law of distribution, which is represented by the curve of which the equation is given:—

$$y = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

Every statistical study of the ability of college students would suggest that it is not unreasonable to assume such an approximate distribution of ability. If this is the case, the most obvious method of assigning grades would be as shown below:

E	G	F	P	D and X
5	20	50	20	5

This method has in fact been adopted, with slight modifications by the institutions named.

It may appear that such a method leaves little liberty to the individual instructor, fixed as it is by arithmetical consideration. Herein lies its excellence, for when large classes free from selection are under consideration, it may be sup-

* Uniformity of Grading in Colleges and Universities, by R. C. Brooks, *School and Society*, Vol. I, No. I, pp. 32-35.

posed that the class average forms a better estimate for marking than the variable and often chance standard of a particular instructor. It should be observed that this method of marking does not apply to small selected classes, but it may be noted that a correct standard deduced from the large elementary classes, by statistical methods, is the best guarantee of a reasonable scale of marking in the more advanced classes.

The question of college grading admits of long and often heated argument; great differences in the standards of departments produce discontent in the student body; so much so that at times, to anyone who is apt to think quantitatively, the whole system seems chaotic, and the only solution would appear to be to discontinue grades, and rely on percentages which do not admit of such ready comparison from department to department. If colleges insist on some method of mark grading, and for many reasons it is desirable to give such grades, the following course should be adopted. Each department that has to deal with large groups should be informed that, unless there is great reason to the contrary, the distribution of grades, over a period of one or more years, must fall within certain limits. These limits would be determined partly by the ideal distribution curve, and partly by the grading system in vogue.

Where the classes are small, and selection and elimination have operated, another course may be adopted with a view to obtaining more uniformity of grading. When the constitution of each class is determined, the office of the college inquires into the previous grades of each member of this class, and sends to the instructor, with no individual names, the distribution of grades of these students based on past records. Thus in a class of ten, if the instructor is informed of the numbers of individuals who, on their total previous record, have obtained the grades:—E, G, F, etc., respectively, he is in a position to know the type of class he is receiving, which gives the surest basis for a sound relative system of marking. If the grades given to these small classes, over a period of time, do not approximately correlate with the average of the grades received previously, there is strong evidence in favor of the assumption that the grading of that instructor is not comparable with what may be called the "mean college grading."