tion took place in something less than that time.

The favorite diet of the animals under observation was, without question, freshly killed mice. Shull, estimating four of these shrews to the acre, figured that on a farm of one hundred acres, they would, in a year, devour 38,400. Realizing the vast amount of damage these rodents are capable of producing in agriculture and considering also the almost exclusively carnivorous habits of the *Blarina brevicauda,* one must admit a great economic value for this shrew.

H. L. Babcock

Dedham, Mass.

The Limit of Uniformity in the Grading of College Students by Different Teachers[1]

In the University of Missouri our grades have, since five years ago, been defined by the frequencies of their permitted occurrence: according to our definitions 25 per cent. are superior, 50 per cent. medium and 25 per cent. inferior grades.[2] We hoped thereby to diminish or even exterminate the divergence of marking then existing. We actually reduced this divergence; but only two thirds. We failed to exterminate it. One third of the former lack of uniformity persists, as may be seen from my previous report in Science, and we ask the question: Why does it persist?

It seems that the chief cause is the inability (call it unwillingness, if you wish, but nothing is gained by any name) of the teachers to differentiate between the performances justly to be expected of a freshman and a senior. For simplicity's sake I speak of two college classes only. Instead of recognizing the relatively superior work of certain freshmen among the freshmen, the teacher compares their work with the work of seniors, and then, of course, finds it to be but weak. And, instead of recognizing that some of the seniors are much less accomplished than other seniors, the teacher compares the weaker senior's accomplishment with that of the freshman and finds it quite remarkable. The result is a widely spread tendency of teachers to report an excess of inferior grades in freshman classes and an excess of superior grades in senior classes. This seems to explain that persistent fraction of the lack of uniformity which we could not eradicate.

Here is the example of an individual teacher in history whose total distribution of grades is approximately that prescribed by the university:

	25% Sup.		50% M.	25% Inf.	
	%E	%S	%M	%I	%F
Underclassmen	1	16	51	25	7
Upperclassmen	6	30	40	20	4

Is there any remedy? It seems simple. Let the teacher differentiate more between the work of freshmen and that of seniors. Assign to the freshman such tasks as are appropriate to the condition of the student who has just left the high school, and to the senior tasks which approach in difficulty, in the requirement of initiative, of resourcefulness, the tasks which the research work of the graduate school keeps ready for the senior as soon as he has his diploma.

But this remedy is not as simple and easy of application as it looks, for the average college teacher seems to be incapable of making the differentiation required. Instead of comparing, rather, freshmen with high-school pupils and seniors with graduate students, he compares freshmen with seniors in the performance of an identical task given to both. However, we must have patience with the teacher. His own task is not small. There are three influences from which he can not easily free himself. (1) Freshmen and seniors, after all, belong socially to one group, that of college students, and neither to the group of high school pupils nor to that of members of the graduate school. (2) He is in mental contact with both freshmen and seniors all the time, but usually no longer with high school pupils and not, probably, with graduate

[1] Read before Section L—Education—American Association for the Advancement of Science, Atlanta, December, 1913.

[2] Compare two former papers: "The Grading of Students," Science, 28, pp. 243–250, 1908; "Experiences with the Grading System of the University of Missouri," Science, 33, pp. 661–667, 1911.

students either. (3) He probably has, frequently, in the same class both freshmen and seniors taking together exactly the same course, and then he can hardly be blamed for comparing their work, even though in the abstract he ought not to compare it. If we want to solve the problem, we have to free the teacher who, usually, is incapable of freeing himself, from these three influences. And that looks like an almost hopeless problem. But, meanwhile, let us not forget that two thirds of the lack of uniformity in grading among teachers can be removed, and that this can be done easily and simply by proper definitions of the grades, for example, by those definitions which we have used in Missouri.

I have now practically said what I wanted to say. If I continue, it is for the illustration of special points rather than for the statement of any additional principle. Let me recall the remark that the tasks to be assigned to seniors, or to members of both upper classes, ought to approach in the requirement of initiative, of resourcefulness, of originality the tasks which the research work of the graduate school places upon its students. I here wish to make it clear that the average college teacher may be expected to offer stubborn resistance to such a demand. For the illustration of the fact that the work assigned to upper classmen generally approaches, in the lack of any requirement of resourcefulness, the work of the high school rather than that of the graduate school, let me refer to data which, at the first glance, seem to be unrelated to the question, but which nevertheless illustrate it well. I am thinking of the high marks obtained by the women students in coeducational institutions. In the University of Missouri we find for the first semester 1912/13 the following record:

Grade Hours		Per Cent. Superior	Per Cent. Medium	Per Cent. Inferior
22,000	Men	23	53	24
7,000	Women	29	55	16

I suppose that the purpose of college training is to prepare students to meet more proficiently all the varied demands which society later will make upon them,—as the common phrase is, to make better men and better women of them. According to the college records one should expect that women rather than men would be found to be the leaders of human society. As a matter of fact there are but few women among the leaders of mankind even in this decade of this century. I recognize, of course, that women are handicapped by three conditions, by legal discriminations, by the force of tradition, and especially by the obstacles resulting from motherhood. No one, however, would assert that, these obstacles being removed, the women would surpass the men in the leadership of society. There is, then, something wrong in such college records which bluntly state that college women are better prepared for leadership in human life than college men. What is wrong in these records is obviously the result of the teachers giving the wrong kind of a test. Instead of testing the initiative which the student should have been trained to put into action for the solution of a certain kind of problems, the teacher tests almost exclusively that kind of accomplishment which depends on the degree of faithfulness and regularity in the performance of assigned tasks. We need not be astonished that the average teacher does not and really can not give the former kind of test, the test of "initiative put into action." Educational science is still so undeveloped that in many subjects the teacher himself does not know how to give such a test. And then—he who tests initiative has to employ initiative himself in the act of testing. That requires an immensely greater effort on the part of the teacher than to test, in the traditional way, how faithfully the students have done their assigned work, and so we can hardly expect the teacher, already overworked, to put himself under the strain resulting from a more proper method of testing.

The same conditions apply to the testing of freshmen and seniors. The seniors, being only one step removed from graduate students, ought to possess a comparative degree of initiative. But their examinations are conceived more like those of college freshmen

than like those of beginners in graduate work. The teacher thus develops in himself the illusion that his average senior, however illogical this is, stands above the average of his own group, and that all the seniors deserve unusually high marks, that is, in comparison with freshmen. But let these seniors enter the graduate school, and some of them will be found, by the different kind of test there employed, to be almost incapable of doing any graduate work at all, because they are deficient in originality, initiative, resourcefulness, whatever you call it, in their chosen line.

This tendency to compare freshmen and seniors is so deep-seated that there is no hope of eradicating it by simply calling attention to it. As in college, so you find it in the high school. My former colleague in Missouri, Professor C. Alexander, found in an (unpublished) investigation of the grading of high schools, that the freshmen are reported partly as average scholars, partly as superior, and partly as inferior; but the seniors, there, too, are reported mostly as high-grade scholars. The low-grade scholars are said to have been eliminated. Now some of these high-grade scholars, obviously not the worst, enter the state university. One should think, then, that our college teachers in the freshmen classes would find it a difficult task to separate from this whipping cream any more plain milk. But the contrary is true. Our teachers complain constantly of the poor scholarship of these "selected" college freshmen.

All this shows, by the way, how unfounded the statement is which we hear again and again that the normal, *i. e.*, symmetrical, curve of distribution is inapplicable to college students because they are supposed to be a selected group. Only then would the symmetrical curve of distribution be inapplicable, if the college freshmen under consideration had been selected by freshmen tests from college freshmen, or if the college seniors had been selected, by tests appropriate to seniors, from the entire group of seniors. There is no reason why the symmetrical curve should be inapplicable to the entire group of freshmen, or to the entire group of seniors, or to the entire group of graduate students or to any group, provided only that the group is complete as a group. That the group came into existence by selection from a different group does not seem to matter when each new group is confronted with new kinds of tasks. There are those who say that it is easy to prove, by examination tests of the ordinary, traditional type, that college students must be regarded as a selected class[3] in the sense that their distribution is not represented by a symmetrical, but by a skewed curve. I have already, a few years ago, called attention to the fact[4] that such examinations are unreliable. Simply make the examination difficult and set a time limit; the curve appears skewed one way, most of those tested crowding in the direction of low ability. Make the examination easy and abolish or greatly extend the time limit; the curve appears skewed the other way. I offer to prove at will by an examination left to my choice that any group of students is distributed either way. Just tell me in advance which way you want the curve skewed.

For the practical problems of college administration this question as to the exact nature of the curve of distribution is really of minor importance. If, however, we just have to make an assumption, it is safest to assume the symmetrical normal distribution. We have assumed in Missouri that the distribution is either normal or very nearly so and experienced no inconvenience. We have reduced the lack of uniformity between teachers to one third of its former amount simply by the adoption of scientifically justifiable definitions, and a reduction to that amount is worth while. But to eradicate the last third is a complex problem of the future, so complex that it may never be completely solved. As has been indicated, it seems to involve problems of our whole educational system and even of the broader social organization of the nation.

Max Meyer

[3] Compare the two tendencies, conflicting with each other, according to Cattell, *Popular Science Monthly*, 1905, p. 372.

[4] Science, 33, p. 667, 1911.