

“THE MARKING SYSTEM IN THEORY”

By EDWIN G. BORING
Cornell University

The last ten years have seen the publication, in educational and scientific journals, of more than a dozen papers dealing with proposed reforms of the systems of marking now generally followed in colleges and universities. The writers advocate, for the most part, an assignment of marks in accordance with some theoretical principle. The awakening of teachers to the need for a real understanding of the basis upon which grades are assigned is, indeed, most desirable; and it appears, from the instances in which one or other of the new systems has already been adopted, as if reforms were about to be realized. It is, however, all the more important that every new proposal should be subjected to the most careful scrutiny. If the aim of our reform is to avoid injustice, we must be eagerly zealous to avoid any new injustice inherent in the reform itself.

One of the most recent and definite proposals for the alteration of the present scheme of grading has been put forward by I. E. Finkelstein in a book entitled, *The Marking System in Theory and Practice*.¹ The work presents a large body of facts, methodically arranged, which furnish a powerful argument for reform. Its final plea is, however, for the establishment of a system of marking, in which the quality graded shall be assumed to be distributed in accordance with a frequency curve that is skewed² toward the upper end of the marking

¹ *Educational Psychology Monographs*, Baltimore, 1913.

² Finkelstein's notion of 'skewness' is not altogether clear. If he intends his mention of a "skewed curve with the mode to the right of the middle of the abscissa" (p. 16) as a definition, we are probably to understand that a curve is skewed in the direction in which the mode is displaced. Without any previous mention of skewness at all, he presents a figure in which the curve is said to be "obviously skewed to the left" (p. 12); but obviousness is scarcely a rigid criterion. What, for example, is to be said of the curve figured in this article? In it the mode is "in the middle of the abscissa";—that is to say, the mode and the median coincide. For Finkelstein the curve would not be skewed. For E. L. Thorndike (*Theory of Mental and Social Measurements*, 1913, p. 77), however, this curve would be skewed, since, by the "approved" formula, skewness depends upon the deviation of the median from the average.

scale. As the assumption is open to criticism, and as the whole question of marking is of importance at the present time, a critical examination seems warranted. Finkelstein advances four arguments in behalf of his special contention (pp. 11-16). We propose to take them up in order.

(1) It is this author's first task to show that "the elimination of the academically unfit by the mill of the public school system does not produce a distribution like that of *Fig. B* [which is skewed toward the left, *i. e.*, toward the lesser degrees of ability], but leaves us, after all, the form of distribution shown in *Fig. A* [which is the normal frequency surface]" (p. 14). Finkelstein thinks that, if there were an elimination of those lowest in ability, the natural result would be a lopping off of the left portion of the curve, resulting in a skew in that direction.

He first quotes Judd³ to show that "the elimination is not so severe . . . as is often supposed." "A study of the large high schools in the city of Chicago in its relation to the University of Chicago," he quotes, "shows that students go to college from every level of scholarship above the passing mark." But it should be noted that the boundary is 'the passing mark;' *some*, therefore, are eliminated. In the passage quoted, Judd is arguing, not that colleges admit inferior individuals, but merely that they admit mediocre ones, that there is not a high degree of selection: "It is not likely to be true that any institution will draw sharp lines of college admission or will attempt to maintain standards which exclude the mediocre student."

After Judd, Meyer is cited.⁴ "College teachers usually assert that the curve of distribution is not the normal curve, but a skewed curve . . . They are usually ready to explain this by referring to the elimination of poor scholars in the high schools and lower schools. I have considerable doubts

Further, it is important to note that Finkelstein uses 'skewness' in exactly the opposite sense to Thorndike (*cf. ibid.*, p. 33; or the earlier edition, 1904, p. 31, which was available to Finkelstein). Finkelstein's curve, which is "obviously skewed to the left," would be said by Thorndike to be "skewed toward" the right, and the formula for skewness is consistent with this latter view. In order to avoid confusion, *we shall throughout this discussion use the term 'skewness' in the sense in which it appears to be used by Finkelstein, viz., the displacement of the mode to the one side or the other of the median.* Finkelstein's reversal of an accepted usage should, however, not be forgotten.

³ Judd, C. H., *On the Comparisons of Grading Systems in High Schools and Colleges*, *School Rev.*, 18, 1910, p. 469. The reference in Finkelstein to vol. 19 is incorrect.

⁴ Meyer, M., *Experience with the Grading System of the University of Missouri*, *Science*, N. S. 33, 1911, p. 666.

as to this elimination. Is the work done in high school really so much like that done in college that there is a large previous elimination of poor college students?" There is an omission indicated by Finkelstein in his quotation. The passage from Meyer reads: "College teachers usually assert that the curve of distribution is not the normal curve, but a skewed curve like that of the continuous lines." The 'continuous lines' refer to a figure showing a curve skewed to the right. Finkelstein quotes Meyer to show that the curve is not skewed to the left; Meyer is actually maintaining that the skew is not to the right. Perhaps the end is the same in either case, *viz.*, that the curve is normal; but the error shows within what confusion we are working when the same sort of elimination can be held to skew the curve either the one way or the other.⁵ How, in such circumstances, can we expect to come to a reliable conclusion? And, in fact, Meyer does not pretend to do this (although Finkelstein does not cite the relevant passage); for he writes, on the same page, "I do not admit, then, that anybody has proved thus far that the distribution of scholarly accomplishments in college is like the normal curve or like a skewed curve either way."

Finkelstein, however, seeks to go farther (p. 13) than the bare statements of Judd and Meyer that selection is not severe at the time of entrance into college. He quotes Huey⁶ to show that "retardations occur continuously on up the years of growth to maturity," and suggests that "the remnant" that "arrives at college," after "its arrested members have dropped out by elimination," is still represented by the normal curve, because "the poorest of the surviving members represent a minimal ability just sufficient to cross the dead line of university entrance." The poorest drop out; but the distribution among college students remains unaltered, because others are retarded to take the places of those eliminated. So is nature become the handmaid of the academic statistician! It is not made clear why she is not equally concerned by the

⁵ Some recent figures published by D. Starch (*Reliability and Distribution of Grades, Science*, N. S. 38, 1913, p. 635) seem to indicate that a continuous progressive elimination to a point beyond the median might result in a curve skewed to the right. W. T. Foster (*Scientific Distribution of Grades at Reed College, Science*, N. S. 35, 1912, p. 887) assumes a displacement to the left,—the form that Finkelstein appears to attribute erroneously to Meyer.

The actual significance of Starch's curves does not concern us here, because we are arguing on purely theoretical grounds, whereas his curves are based on actual findings.

⁶ Huey, E. B., *Retardation and Mental Examination of Retarded Children, Jour. Psycho-Asthenics*, 15, 1910, p. 41 (not page 43, as Finkelstein states).

embarrassment of the general biometrician at the discovery that the distribution of a general sample of the adult population must now, on the basis of the same argument, be skewed to the left,—for the general sample must include the normally distributed sample at college *plus* a large mass of the previously eliminated. Finkelstein himself, in speaking of "native ability," says that "in any population its distribution is that known as the curve of error, the probability curve, or Gauss's curve" (p. 11). In this connection we do not need to differ from Huey. Throughout his paper, Huey places in opposition the ideas of growth and retardation. Both are continuous. The mean of mental ability advances with age, but retardations continually occur to a greater or less extent; they affect, more or less, the whole sample of population, and thus they increase in range while the mean advances. There is no reason, however, to assume that the curve for the whole sample is anything but the normal surface. Huey might just as well have spoken only of retardation, in which case the genius would have been described merely as the least retarded; or he might have considered only growth. The point is, in the first place, that Huey says nothing to indicate that a sample of the total population is not distributed normally; and, in the second place, that, if it is so distributed, then the separation from it of a large selected portion—such a selection as Finkelstein freely admits—would be very unlikely indeed to leave a symmetrical curve.⁷

⁷ For the sake of accuracy it should be noted that, according as selection tends to lop off the one end of the curve only, or to affect the whole group (in an increasing absolute amount) to a point somewhere beyond the median, the curve will be skewed to the left or right respectively. The first condition is what one would naturally expect, but it is conceivable that the lower two-thirds of the population, for example, should contribute, in proportion to their numbers, more or less to the general elimination. If both conditions are possible, then the intermediate case of the symmetrical curve is also possible. It is, however, improbable; and, in any case, Finkelstein should have shown all the possibilities, rather than have left the reader with the impression that the evidence preponderates in favor of the normal form of distribution.

A word here may not be amiss concerning the conclusions of other writers with respect to the curve. College teachers, according to Meyer (*loc. cit.*) are inclined to accept a curve skewed to the right. Foster (*loc. cit.*) bases the Reed College system on one skewed to the left. J. McK. Cattell (*Examinations, Grades, and Credits, Pop. Sci. Mo.*, 66, 1905, p. 372) selects the normal surface as a sort of compromise between displacement to the left, which he thinks is what might be expected, and displacement to the right, which is what he actually finds. Cattell says that the latter form is due to the tendency of the teacher to mark up (p. 375), but does not say why he is unwilling to explain in this way the whole difference between the theoretical and the actual forms. Meyer (*loc. cit.*), lack-

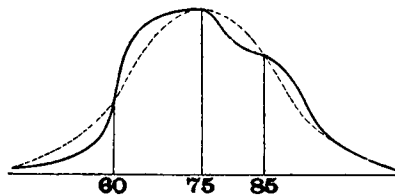
(2) Finkelstein's second point is that "marks should indicate, not ability, but accomplishment" (p. 14). He continues; "Is accomplishment distributed like ability? We have argued that accomplishment is the result of ability, *plus* previous preparation and zeal. When like zeal or effort is exhibited by two pupils, the one dull and the other brilliant, it seems probable that the gifted pupil reaps a *proportionately* larger result. It is as if effort were multiplied into ability rather than added to it. It is difficult to prove this assertion, but if it be admitted, the conclusion is evident that, if all the students put forth the same amount of effort, the curve of accomplishment would be skewed to the right." This is the argument in full. There has been no attempt to increase its obscurity by condensing it in transcription. Apparently what Finkelstein means is this: Ability is distributed normally. Zeal is assumed tentatively to be equal for all. It is effective for accomplishment in proportion to the amount of ability. Thus the normally distributed 'ability' is multiplied by the constant 'zeal' to give the curve of accomplishment. The result is, of course, not a skewed curve, but another normal curve. The form of a curve is never altered by multiplication by a constant.⁸

(3) In the third place, Finkelstein maintains that "when a marking system is arranged to measure accomplishment,

ing evidence, takes the normal surface as the most probable form. So also does W. F. Dearborn (*School and University Grades*, Bull. Univ. Wisconsin, 1910, No. 368). A. G. Smith (*A Rational College Marking System*, *Jour. Educ. Psych.*, 2, 1911, p. 388) adopts the normal surface, because, "although only a small proportion of high-school graduates take a college course, it is fair to assume that the majority of such graduates have mental ability sufficient to take such a course, and a large percentage of the students who do not finish the high school would, if given a fair chance and encouragement, be capable of development and completing a college course." A majority,—but, like Finkelstein, Smith forgets the minority. And all these writers seem to expect the elimination to occur, like a sudden visitation of Providence, between graduation from high school and entrance into college. Why can not they learn from Huey that "retardations occur continuously on up the years of growth to maturity,"—in infancy, in the primary, in the grammar grades, in high school, in the college, in the graduate school? Surely the imbecile is eliminated from the college ranks, though every high-school graduate take his Ph.D.!

⁸ Finkelstein's meaning may possibly be that accomplishment is dependent *both* directly upon ability *and also upon* effective zeal, which is, in turn, dependent upon ability. Thus accomplishment would be the sum or product of these two normally distributed factors. But since the one is ability, and the other is directly dependent upon ability, the curves could not be displaced relatively to each other and the resulting surface could not be asymmetrical.

another factor undoubtedly enters to skew the curve further to the right. This factor is the incentive offered to certain groups of students by the critical points of the system. Thus, at Cornell University, where a mark of 60 or above is necessary to pass, there evidently exists a type of student of inferior or average ability who aims for this mark. The possibility of exemption from final examination by obtaining a mark of 85 in preliminary examinations offers also a powerful incentive to students of average ability to push their accomplishment to the extent of their ability. Again, an average student, by persistent effort, may hope to obtain a mark of 80 or over, and thus be eligible for consideration for Phi Beta Kappa and other special honors. On the other hand, the brilliant student, who probably profits most by effort, cannot attain a mark above 100, and it is not difficult for him to obtain a mark of 85 or over. The net result, then, of the imposition of a marking system is to crowd the grades of accomplishment forward to the right, and thus, again, to tend to form a skewed, and not a bell-shaped curve" (pp. 15f.).



Hypothetical curve of distribution of marks on percentile scale. Median and mode at 75%; hence (in Finkelstein's terms) the curve is not skewed. Departure from normal surface (dotted lines) is explained by additional 'zeal,' which may raise values normally below 60 to amounts between 60 and 75, and values normally between 75 and 85 to amounts above 85.

It is true that the effect of additional zeal or effort on the part of special groups would result in a curve that is not bell-shaped; but it is not necessarily true that the curve would be skewed in the sense that the areas on either side of the mode are unequal, that is to say, that the mode and the median do not coincide. In the figure is shown a curve in which a number of individuals, falling below 60 in the normal distribution, have been displaced to a position above 60, and in which the

same sort of alteration has been made at 85. The curve is irregular, but it is not skewed.⁹ If it were the case that the additional effort of the students, who would naturally fall below 60, brought them above 75, or if the struggle to pass 85 extended in any large measure to those who would otherwise receive less than 75, a skew might occur. In such a case, however, the increase of effort would extend to a very much larger number of students, and would be effective to a very much greater extent, than Finkelstein seems to imply.

(4) Finally, the advocate of this system of marking would have us believe in the displacement of the ideal surface to the right, because "the actual distribution of over 20,000 marks in Cornell University is a curve skewed to the right" (p. 16). So, then, we are to reform our distribution by reference to the distribution itself. Could anything be less promising to a seeker after a 'scientific method of grading'? It is true that Finkelstein regards his curve as reliable because it is the "*consensus gentium*" of the faculty. But to this we may reply, in the first place, that if the *consensus gentium* be the just standard, then the individual student, whose final fate in college depends upon the average of a large number of marks, is not so badly off as we had supposed; and, in the second place, that a reformer has done but half his duty when he has corrected the unjust deviations from the mean without inquiring into the deviation from justice of the mean itself. If it is at all demonstrable that special factors might be operative to skew the actual curve to the right, then this curve can not be considered as a basis for the ideal distribution. And it is not difficult to suggest influences that might be operative.

(a) The curves of both Cattell¹⁰ and Finkelstein are skewed to the right; that presented by Starch¹¹ shows a far lesser skew; and, in fact, Starch assumes that it is symmetrical.¹² Dearborn¹³ finds the distribution practically normal. Since the curves do not agree, there must be some special factor operative, knowledge of which would explain the difference.

(b) A factor that is apparently effective in skewing Fink-

⁹ It is only in Finkelstein's use of the term that we can say that the curve is not skewed. By Thorndike's definition it is skewed, but it is skewed in a manner that in appearance somewhat resembles Finkelstein's skew, not to the right, but to the left!

¹⁰ *Op. cit.*, p. 374.

¹¹ *Loc. cit.*

¹² Critics, however, on the basis of the figure shown, might object to the statement that "the distribution of the marks of college freshmen . . . in fact does conform to the probability curve." (p. 364.)

¹³ *Loc. cit.*

elstein's curve to the right is inherent in the method of assigning final marks at Cornell University. Students having a grade of 85 at the end of the term may be exempt from the final examination. Those who take this examination have their grades raised or lowered, as the case may be, by its result. It thus happens that, while students below 85 at the end of the term have a chance to be raised above 85, the students above 85 have no chance of being lowered, because their work is already completed. Thus it happens that there is a weighting of the curve in favor of the student. In some courses, where the final examination is made easy, the effect would be considerable; and in those courses in which the grading by final examination is substituted for grading by term work, the differences would be very great indeed.¹⁴

(c) Finally, it may be suggested tentatively that, among other factors operating to skew the curve to the right, there is the fact that instructors do not regard all units on the marking scale as equivalent. It seems probable that a unit at the upper end is considered as more significant than a unit at the lower end. We hesitate more between giving 80 and giving 85 than we do between 20 and 25. Most instructors think nothing of giving a mark of 20% to an answer, which they would have marked zero had they continued to deduct for errors and omissions at the rate at which the first 10% was deducted. If, for example, Finkelstein's curve were replatted with an irregular scale of abscissae, so proportioned that from 85 downwards equal units of abscissa represent uniformly increasing numbers of percentile grades, the curve might very well be symmetrical. It is approximately symmetrical, if, as equivalent groups, there are selected the grades 0-44, 45-53, 54-61, 62-68, 69-74, 75-79, 80-84, 85-89, 90-94, 95-100.

It appears, then, that special factors may be present which determine the actual curves; and that, therefore, it is scarcely legitimate to utilize these actual curves as a basis for the determination of the ideal curve.

In general, Finkelstein's conclusion that "the pattern or theoretically ideal curve of high-school and college marks is, therefore, not the probability curve, but the skewed curve with the mode to the right of the middle of the abscissa"

¹⁴ Finkelstein speaks throughout as if the number of marks of 85 or above actually represent (at least in the College of Arts and Sciences) the number of exemptions that were or would have been given; and he speaks as if the percentage found at 85 or above in the general curve represented the percentage that should receive exemption in practice. Such a procedure, on his own assumption that the average curve is the ideal one, would, for the reasons shown, be incorrect.

(p. 16) cannot be justified. Just what harm his assumption might work can not be definitely stated. We can say only of *The Marking System in Theory and Practice* that much more seems to be known of practice than of theory; and that, if reforms are to depend upon both—as, indeed, it would seem that they must—a more careful induction is needed than has hitherto been presented.