

On the Principle and Methods of assigning Marks for Bodily Efficiency. By
FRANCIS GALTON, F.R.S.

THE principle and methods by which marks in examinations can justly be assigned is of no small interest to anthropologists, on account of their bearing on the art of describing the various degrees of human faculty among individuals and races. What will now be said is of general application, though it is written especially with reference to examinations in physical powers.

The question to be solved is of this kind. Suppose that one man can just distinguish a minute test object at the distance of 25 inches, another at that of 35, and again another at 45 inches, how should we mark them? We should be very rash if we marked them in the proportion of 25, 35, and 45, or even if, for some good reason, we had selected 25 as the lowest limit from which marks should begin to count, we should mark them as 0, 10, and 20.

Two separate considerations are concerned in the just determination of a scale of marks, namely, absolute performance and relative rank, which are apt to be confused in unknown and varying proportions.

Absolute performance is such as is expressed by the 25, 35, and 45 inches just spoken of. It is perfectly correct in some cases to mark, or let us say to pay, for this and this alone, upon the principle of piece-work, namely, that the pay ought to be proportionate to the work accomplished, or to the expected output in after life.

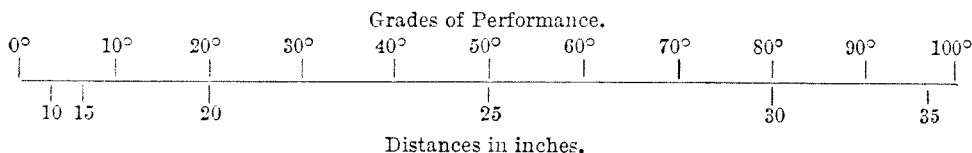
Relative rank is, however, on the whole, a more important consideration than the absolute amount of performance by which that rank is obtained. It has an importance of its own, because the conditions of life are those of continual competition, in which the man who is relatively strong will always achieve success, while the relatively weak will fail. The absolute difference between their powers matters little. The strongest even by a trifle will win the prize as completely as if he had been strongest by a large excess. Undertakings where many have failed are accomplished at last by one who usually is very little superior to his predecessors, but it is to just that small increment of absolute superiority that his success is due. Therefore it is clear that relative rank has at least as strong a claim for recognition as absolute performance, if not a much stronger one. They have each to be taken into separate consideration, and each to be separately marked. The precise meaning intended to be conveyed by the phrase 'relative rank' will be better understood further on.

Recurring to the example of keenness of eyesight, let the test object be words printed in diamond type, and the persons tested be Englishmen of the middle-classes, between the ages of 23-26, then the performance of reading diamond type at 25 inches happens to be strictly mediocre. Fifty per cent. of the many persons who were tested performed better than this, and fifty per cent. performed worse. The 35-inch performance was exceeded by only $2\frac{1}{2}$ per cent. of the persons tested, and as to the 45-inch performance it has not in my experience been reached at all. I have had 12,000 persons altogether tested in this way, of both sexes and of various ages, but not one of them has succeeded in reading diamond type at the distance of 45 inches. It is very rare to find one who can do so at 40 inches. Wherever superiority in eyesight is eminently desirable, it would

be absurd to make the marks for the three supposed cases to run proportionately either to 25, 35, and 45, or to 0, 10, and 20. The achievement of 45 inches would deserve much higher recognition. Relative rank and absolute performance should not be confused together.

I use the term relative rank in a large sense, with reference to all persons who have been, or are likely to become, candidates, and not to the small number of them who may happen to be present at a particular examination. Statistical tables concerning the class of persons in question have to be compiled from past examinations, and the rank of the individual has to be determined amidst these. I have often described how this is to be done ('Natural Inheritance,' p. 38: Macmillan & Co., 1859), but the form of a diagram that I now submit is a new and, I think, the simplest of all for the use of an examiner. It tells at a glance the rank held by a man among his fellows in respect to any single and separate faculty. The class from which it is constructed might have been of any length, subject to the condition that the distance between the limits *within which* it extends shall be always divided into centesimal grades; that is to say, running from 0° to 100°. This diagram refers to keenness of eyesight, but the method is of general application. I lay on the table several similar diagrams for various qualities, such as are hung up in my Anthropometric Laboratory in London.

*Keeness of Eyesight, measured by the greatest distance in inches at which
Diamond Type can be read.*



This same method admits of being extended in more than one way. That for which there is most call is where the rank of the quality immediately in question has to be considered in reference to some other quality. Thus it is of little use to know the breathing capacity of the man unless we also know his stature or his weight. Lungs capacious enough to enable a small man to labour violently without panting would be wholly insufficient for the ordinary purposes of a giant, just as an excellent little boiler for a small steam-engine would be ineffective with a large one. The diagram appropriate to the case we are considering could not be compressed into a single line, but requires many. Successive lines in the same page would refer to the successive weights of, say, 100 lbs., 120 lbs., 140 lbs., and so on, and a diagram of breathing capacities for each of these weights would be constructed, but in pencil, just on the principle of that, shown above, for keenness of eyesight. The grades along the top of the page would refer equally to all the lines below. Then bold lines have to be drawn from above downward to connect all the pencilled entries of the same value, just as iso-bars, iso-therms, and other contour lines are drawn (to which the general name of *isograms* might well be given). This completes the figure of which I submit a specimen to the meeting. It hardly needs further description, either how to make or to use it.

When the quality that has to be marked depends upon more than one other quality, as breathing capacity may have to be marked with reference

both to weight and to stature, the simplest plan is to make a separate diagram for each inch or second inch of stature, which is quite near enough. I have, however, contrived to make a single page serve for the whole process by using a sliding strip of paper. I submit it for inspection, but do not care to describe it.

A strong reason for giving prominence to relative rank is that it affords the only feasible measure for many qualities, so that differences in absolute performance have to be inferred from it, according to a principle now familiar to most anthropologists, by using the well-known table of the Probability Integral. A small table based on the latter, but of a totally different form, that I have lately more than once published (*Op. cit.*, p. 205), is very convenient for this sort of work. The following is a brief extract from it:—

Grades of Rank from 0° to 100°, together with the Deviations¹ from the Mean Values at those Grades.

Grades	:	:	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°
Deviations	:	:	- infinity	-1.9	-1.3	-0.8	-0.4	0.0	+0.4	+0.8	+1.3	+1.9
Grades	:	:	91°	92°	93°	94°	95°	96°	97°	98°	99°	100°
Deviations	:	:	+2.0	+2.1	+2.2	+2.3	+2.4	+2.6	+2.8	+3.1	+3.5	+ infinity

¹ The unit by which the deviations are measured is half the difference between the performances of the persons who respectively occupy the grades 25° and 75°.

Some of the consequences of marking separately the relative rank and the absolute performance are seen by the table below. Here the relative rank is in each case supposed to count between the grades of 50° and 100°. Then, if it alone is considered, a man who stands at the grade of 99° in a class that ranges within the limits of 0° and 100°, will be seen to get very nearly the full amount of 10 marks, whereas if absolute performance is alone considered, he would get no more than 7 marks, the full number of 10 being never actually reached, but only closely approached at some such high grade as 99.99...°. The figures in the table would have run very differently if the marks for relative rank had begun after 90° and not after 50°. Still more so, if the lower limit had been 99°, and more still if it had been 99.9°. It seems to me most reasonable, on the whole, that they should usually begin after 50°, as in the following table:—

Proportion of marks assigned to		Rank, 0° to 100°				
Relative	Absolute	55°	75°	95°	99°	99.99...°
All	None	1.0	5.0	9.0	9.8	10.0
$\frac{1}{2}$	$\frac{1}{2}$	0.7	3.5	6.9	8.4	10.0
$\frac{1}{3}$	$\frac{1}{3}$	0.6	3.0	6.2	8.0	10.0
$\frac{1}{4}$	$\frac{1}{4}$	0.5	2.7	5.8	7.7	10.0
None	All	0.4	2.0	4.8	7.0	10.0

The general conclusion to which these remarks lead is, that before arranging scales of marks, the first step is to measure a large number of persons who are of the same class as the expected candidates; this has

already been done to a considerable extent at Cambridge, at Marlborough College and elsewhere. Thence to make tables, and to deduce diagrams from them like that referring to keenness of eyesight, in some cases, and like that referring to breathing capacity in relation to weight, in others. These will exactly determine the qualities of the men to be dealt with, in a statistical sense. It is now the part of those who have to fix the scales of marks to determine the grade at which rank shall begin to count, and to arrange the weight to be given respectively to relative rank and to absolute performance in each sort of examination. This and a few other obvious preliminaries having been settled, the construction of consistent scales of marks would follow almost as a matter of course.

Experiments at Eton College on the Degree of Concordance between different Examiners in assigning marks for Physical Qualifications. By A. A. SOMERVILLE.

AN experiment was made at Eton in July last, with the object of obtaining information upon the following points: (1) whether it is possible to frame a system of marking for physical excellence, based partly upon Mr. Galton's system and partly upon medical examination; (2) whether marks assigned by medical examiners would be as safe a test of excellence as those assigned, *e.g.* by examiners in English essay. The experiment was conducted as follows. A list of points was drawn up with the help of two able medical men. These points were: (1) breathing capacity, as tested by the spirometer; (2) hearing; (3) eyesight, tested by Snellen's type; (4) strength, tested by the grip dynamometer; (5) endurance, tested as follows,—after the maximum reading of the dynamometer had been obtained and registered for strength, it was again (as nearly as possible) obtained, and the number of seconds during which the candidate could hold the needle of the dynamometer between this reading and the reading 10 below it was taken by a stop-watch; (6) relation of height to weight; (7) girth and shape of chest; (8) general muscular development; (9) health record, particular inquiries being made as to rheumatism, asthma, and scarlatina; (10) general aspect and condition.

The first five points depend solely upon measurement, and consequently the marks of the two doctors are the same for those points. The next point was marked, partly by impression, and partly by reference to a table of averages, but it might be made to depend altogether upon averages. The seventh and eighth points were marked partly by measurement of chest, arms, and legs, and partly by examination. The last two points depend altogether upon medical opinion, and involved a thorough medical examination. Ten marks were assigned for each point, and the examination was conducted independently by the two doctors in separate rooms. Thirty-two boys were examined: (1) 20 Army Class boys, including 10 successful candidates at the recent Sandhurst and Woolwich Further Examinations, 2 members of the Cricket XI., and 2 members of the Rowing Eight; (2) 6 other members of the XI.; (3) the remaining 6 members of the Eight. The following table gives the final results, average differences per cent. being calculated with reference to a maximum 50, as the marks for the first five points are the same for the two examiners. (N.B.—Letters are substituted for the names of the boys.)