

and of perfecting the tests for their truth. The latter may involve a better working out of the technique of the use of some of the known tests, or the discovery of new tests (as of experimental verification in modern times), or a better apportionment of the tests to the fields of investigation for which they are best adapted. The business of the philosopher would seem to be concerned especially with this last task. How remarkable and fruitful a revolution would take place in philosophy if philosophers would concern themselves, not in showing that in a certain remote region of thought a certain criterion of truth will not work (and thereby implying that it ought to be rejected completely), but rather in working out the most effective correlation between the various criteria and the types of problems to which they are to be applied!

DENTON L. GEYER.

THE RICE INSTITUTE.

COEFFICIENTS OF DIAGNOSTIC VALUE

ONE of the reasons for the interest in the relation between performance in mental tests and specific or general abilities is the possibility of utilizing the former in the prediction of the latter. All that is needed to establish the possibility of such diagnosis is an unequivocal demonstration that the relation between performance in tests and the ability in question is close enough to permit diagnosis of such accuracy as the concrete situation requires. The necessity of measuring the observed relation between the variables is at once evident; and the most commonly used index of the closeness of the relation is the Bravais-Pearson coefficient of correlation (or some coefficient devised to give approximately the same result with the expenditure of less arithmetical effort).

Now, although it is clear that a close relation between performance in tests and the ability in question is associated with the possibility of accurate diagnosis, and although any increase in the closeness of the relationship is accompanied by an increase in the accuracy of diagnosis, it does not follow that an index which is used in describing the relationship can be taken *directly* as an index of the diagnostic value of mental tests. In fact, the coefficient of correlation, a widely used index of relationship, gives a decidedly misleading notion of diagnostic value. It is my purpose to show why the coefficient of correlation is misleading, and to suggest a coefficient which may be used with less confusion in stating the value of mental tests in practical situations.

The coefficient of correlation is a generalized statement of linear

regression. And "all the statistician means by regression is this: If all the organs *A* of a certain size have associated with them an array of *B* organs having a definite mean value, then this mean value changes with change of *A*. The distribution of the means of *B* arrays for given values of *A*, whether expressed by curve or table, is in its most general form the phenomenon which Mr. Galton has termed regression."¹ Thus, on the average, men who are *x* inches tall will have sons who are shorter than the sons of men who are *x* + 3 inches in height. On the average, children who are eight years old will read a certain fifty words in less time than will children who are six years old. In each of these cases two variables are involved, the *A* and the *B* of Mr. Pearson's statement. In the first illustration, *A* is the height of fathers, *B* the height of sons; in the second illustration, *A* is the age of children, *B* is the time required to read a given fifty words.

Regression is said to be linear if equal increments of one variable are always associated with equal increments of the other. Concretely, if, on the average, the sons of men *x* + 3 inches tall are 1 inch taller than are the sons of men *x* inches tall, no matter what the value of *x* may be, the regression is linear.

The regression coefficient in linear relations is the ratio of an increment in one variable to the associated increment in the other. In the example above, the regression coefficients are one third and three. If each variable is measured as a deviation from its mean, a general statement of the relation is the following:

$$x = b \cdot y$$

where *b* is the regression coefficient. Verbally this equation states: Individuals who differ from the average of variable *Y* by an amount *y*, on the average will differ from the average of variable *X* by an amount *b* times *y*, or *x*.

The scales in which the variables are measured are accidental facts in a majority of cases, facts which affect the relation between the variables not at all. In order to make all scales comparable, the variables may be measured in terms such that the standard deviation of each variable is 1. *When such scales are used, the regression coefficient becomes the coefficient of correlation.*

The coefficient of correlation, therefore, states the ratio of increments in one variable to corresponding increments in the other variable, when the regression is linear and when the standard deviations of the variables are unity (or equal). Thus the correlation coefficient is a natural statement of the regression phenomenon.

¹ K. Pearson, "On the Fundamental Conceptions of Biology," *Biometrika*, Vol. I., p. 323.

From the point of view of diagnosis, the correlation coefficient gives a means of determining preferred values in X for observed values in Y through the use of the regression equation. But the only reason for having preferred values at all is to increase the *accuracy* with which the prediction can be made; and the correlation coefficient is not a direct statement of the accuracy of prediction. It is simply a generalized statement of regression. Clearly the coefficient of correlation should not be used directly as an index of the accuracy with which diagnosis can be made, unless regression and accuracy of prediction can be shown to vary in a one to one relation.

Since the value of one variable is to be predicted from known values of another, it seems logical to measure the accuracy of diagnosis in terms of the amount of error that will be made in prediction by the linear regression equation.² Suppose for one individual the regression equation gives a value x_1 ; suppose that the true value is X_1 . The error is $X_1 - x_1$. Call this error e_1 . For a second individual, the error will be e_2 , and for a third e_3 . The sum of the squares of these errors divided by the total number of individuals is known as the mean square error of estimation. The square root of this quantity, the root mean square error of estimation, may be regarded as the standard deviation of the true values of the individuals around their predicted values. The root mean square error is sometimes called the *standard error of estimation*; it is an index of the dispersion of true values around predicted values, and might be taken directly as an index of the accuracy of diagnosis.

Again, as in the case of the coefficient of correlation, it is desirable to obtain a coefficient that is independent of the scales in which the variables are measured. Again, this independence is gained by expressing the measurements of the variables in such a way that the standard deviation of each variable is unity.

Fortunately, the individual errors need not be determined, and the scales need not be actually changed, for the generalized standard error may be obtained from a knowledge of r . The equation³ is

$$\epsilon = (1 - r^2)^{\frac{1}{2}}$$

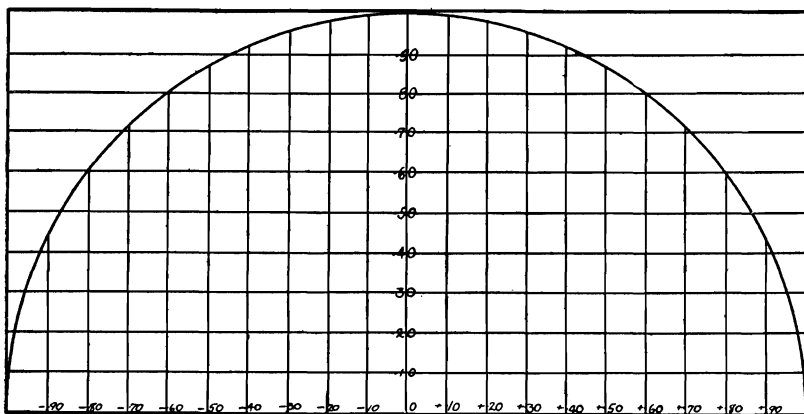
It is important to remember that the error coefficient which we shall use is the root mean square error of estimation when prediction is made from the linear regression equation, and when the standard deviation of each variable is unity.

² Discussion in this paper is limited to the case of linear regression. The significance of the linear equation in the general case is discussed by Mr. Yule in *Royal Society Proceedings*, Vol. 60, p. 477.

³ Yule, *Introduction to the Theory of Statistics*, p. 177. The derivation of the formula is independent of an assumption of Gaussian distribution in the variables.

ϵ is the suggested index of the diagnostic value of mental (or other) tests. It expresses directly the magnitude of the error made in predicting the values of one variable from the values of another. Furthermore, it is easily derived from the coefficient of correlation.⁴

The meaning of ϵ is easily understood. If there is perfect linear relationship between the variables, ϵ equals 0; that is, there is no error in estimating one variable from another. If the two variables are independent, ϵ equals 1. The standard error of estimation is



The Relation between r and ϵ . Horizontal axis: coefficient of correlation. Vertical axis: error coefficient.

here equal to the standard deviation of the variable to be predicted; the accuracy of diagnosis will be only that which results from chance success. The maximum value of ϵ is 1; this is a logical limiting value, since it represents the error that is made in the wisest possible guesswork, that is, in giving all cases in the variable to be predicted the mean value. ϵ takes all values between 0 and 1; .5 means that the standard error made in predicting B from A is half as great as the error that would be made in the wisest guessing at the values of B ; .25 means that the error is one quarter as great. These values, therefore, have a simple and easily understood meaning in the discussion of accuracy of diagnosis.

The coefficient of correlation is a misleading index of the diagnostic value of a test because it does not bear a one to one (or linear) relation to the standard error of estimation. The actual relation is as stated above. The graph of this relation is of considerable interest, and shows some striking facts. It will be seen that as r decreases from $+1$, or increases from -1 , the standard error increases

⁴Dr. W. V. Bingham suggests that the fact that ϵ does not measure "reliability" in the psychological sense should be mentioned.

with extreme rapidity. Concretely, when r is equal to $\pm .86$, the standard error is half as great as it ever will be; when r equals $\pm .50$, the standard error is seven eighths its maximum size. *The amount of error made in predicting from the regression equation on the basis of a correlation of $+.44$ is nine tenths as great as the error made in guessing on the basis of no information whatever.*

Does all this mean that a correlation of $+.45$ indicates that a test has no diagnostic value? Not at all; for values are relative things that depend upon the accuracy of diagnosis that particular situations demand. In the most objective terms, ϵ indicates that the root mean square error of estimation in predictions on the basis of the regression equation and a correlation of $+.45$ is nine tenths as great as the error that would be made in predicting by the same method on the basis of a correlation of $.00$. The investigator must decide for himself whether such reduction in the error of estimation makes the test valuable in the particular situation in which the test is to be used. The point of the present paper is to emphasize the necessity, in judging diagnostic value, of recognizing that the error of estimation is not reduced directly as the correlation is increased, of recognizing, for example, that a correlation of $+.60$ reduces the error of estimation only two tenths. I believe for these reasons that r is misleading, and that when diagnostic value is discussed, ϵ might be used as an index with advantage.

BEARDSLEY RUMML.

UNIVERSITY OF CHICAGO.

REVIEWS AND ABSTRACTS OF LITERATURE

Egotism in German Philosophy. GEORGE SANTAYANA. London and Toronto: J. M. Dent and Sons, Ltd.; New York: Chas. Scribner's Sons. Pp. 171.

To the many hard things that are nowadays being said against Germany Mr. Santayana adds the accusation of Protestantism. That the Germans have long since adopted Luther as a patron saint and claimed him as exclusively their own, is well known. But this has commonly been cited as evidence of their tendency to nationalize and appropriate the common European heritage. Most of the critics of Germany have been themselves Protestants who were unwilling that their faith should be identified with the creed of a militant nationalism. Mr. Santayana has no such scruple. Their philosophy, he says, "is Protestant theology rationalized" (p. 22); and it is clear that so saying he means to disparage both German philosophy and Prot-