Psychological Testing

1997

ANNE ANASTASI
Department of Psychology, Fordham University

SUSANA URBINA
Department of Psychology, University of North Florida

correlation of zero.

Having established a significant correlation between test scores and criterion, however, we need to evaluate the size of the correlation in the light of the uses to be made of the test. If we wish to predict an individual's exact criterion score, such as the grade-point average a student will receive in college, the validity coefficient may be interpreted in terms of the *standard error of estimate* (SE_{est}), which is analogous to the error of measurement discussed in connection with reliability. It will be recalled that the error of measurement indicates the margin of error to be expected in an individual's score as a result of the unreliability of the test. Similarly, the error of estimate shows the margin of error to be expected in the individual's predicted criterion score, as a result of the imperfect validity of the test.

The error of estimate is found by the following formula:

p.142

$$SE_{est} = SD_y\sqrt{1 - r_{xy}{}^2}$$

validity coefficient and SD_y is the standard de-

The error of estimate is found by the following formula:

$$SE_{est} = SD_y\sqrt{1 - r_{xy}^2}$$

p.142

in which r_{xy}^2 is the square of the validity coefficient and SD_y is the standard deviation of the criterion scores. It will be noted that if the validity were perfect ($r_{xy} = 1.00$), the error of estimate would be zero. On the other hand, with a test having zero validity, the error of estimate is as large as the standard deviation of the criterion distribution ($SE_{est} = SD_y\sqrt{1 - 0} = SD_y$). Under these conditions, the prediction is no better than a guess; and the range of prediction error is as wide as the entire distribution of criterion scores. Between these two extremes are to be found the errors of estimate corresponding to tests of varying validity.

Reference to the formula for SE_{est} will show that term $\sqrt{1 - r_{xy}^2}$ serves to indicate the size of the error *relative to the error that would result from a mere guess* (i.e., with zero validity). In other words, if $\sqrt{1 - r_{xy}^2}$ is equal to 1.00, the error of estimate is as large as it would be if we were to guess the individual's criterion score. The predictive improvement attributable to the use of the test would thus be nil.

If the validity coefficient is .80, then $\sqrt{1 - r_{xy}^2}$ is equal to .60, and the error is 60% as large as it would be by chance. To put it differently, the use of such a test enables us to predict the individual's criterion performance with a margin of error that is 40% smaller than it would be if we were to guess.

It would thus appear that even with a validity of .80, which is unusually high, the error of predicted scores is considerable. If the primary function of psychological tests were to predict each individual's exact position in the criterion distribution, the outlook would be quite discouraging. When examined in the light of the error of estimate, most tests do not appear very efficient. In most testing situations, however, it is not necessary to predict the specific criterion performance of individual cases, but rather to determine which individuals will exceed a certain minimum standard of performance, or cutoff point, in the criterion. What are the chances that Mary Greene will graduate from medical school, that Tom Higgins will pass a course in calculus, or that Beverly Bruce will succeed as an astronaut? Which applicants are likely to be satisfactory clerks, insurance agents, or machine operators? Such information is useful not only for group selection but also for individual career planning. For example, it is advantageous for a student to know that she has a good chance of passing all