ITEM SELECTION BY MEANS OF A MAXIMIZING FUNCTION

PAUL HORST
The Procter and Gamble Company
Cincinnati, Ohio

A new item selection technique is presented which takes into account the intercorrelations of the items as well as their correlations with the criterion. The technique is regarded as superior to comparable techniques in that it is considered to achieve greater economy of time, greater objectivity of procedure, higher validity, and higher reliability. The mathematical theory underlying the method is developed. An approximate solution of the mathematical equations is suggested. An approximation procedure for the complete item selection technique is presented, based on the mathematical solution, but much simpler in procedure. The clerical operations involved in the approximation procedure are outlined and illustrated on a sample worksheet.

Numerous methods purporting to increase the validity of objective tests have been published. Many of these are based on techniques of individual item analysis. Perhaps the most comprehensive report of existing item analysis techniques is that of Long and Sandiford.*

It has long been known that the validity of a test is a function, not only of the correlation of each item with the criterion, but also of all the possible intercorrelations of the test items. Notwithstanding this knowledge, surprisingly little has been done in the way of developing item selection techniques which take into account the intercorrelations of the individual items. Professor H. A. Toops of Ohio State University has developed an item selection method, known as the “L-method”, which takes into account item intercorrelations, but so far as I know, this method is not readily available in published form. Furthermore, I understand that the method involves a great deal of labor.

Another technique,† known as “The Method of Successive Residuals”, is also based on a consideration of item intercorrelations, as well as item-criterion correlations. While this method has yielded results definitely superior to those ignoring item intercorrelations, it also is time consuming.

†Horst, Paul, "Item Analysis by the Method of Successive Residuals", Journal of Experimental Education II (March 1934) pp. 254-263.
Recently, however, I have developed a method which is based on both inter-item and criterion correlations, which I regard as much superior to the "Method of Successive Residuals" for the following reasons:

1. Less Time Consuming. It requires only from one-third to one-half the time required by the other method.
2. Greater Objectivity. For the "successive residuals" technique to be practicable it is necessary to run the analysis on separate groups of from 40 to 50 items at one time. The grouping of these items is largely arbitrary.
3. Higher Validity. In general the new method yields slightly higher validity coefficients.
4. Higher Reliability. This point I have not verified experimentally, but the mathematical rationale underlying the latter method suggests definitely that higher reliability may be expected than by the "Method of Successive Residuals".

I. Theoretical Solution

Suppose that we have a test battery of \( n \) items, and that we have responses on all these items from a population of \( N \) cases. We also have an external criterion measure \( Y \). Let us consider the problem of selecting \( m \) items from the total group of \( n \) items so that when the cases are scored on only these \( m \) items the scores will give the maximum correlation with the criterion measures \( Y \). We let \( X_1, X_2, \ldots, X_n \) be the scores on the individual items. If we adopt the unit weight procedure for scoring the items, the \( X \) values will be either zero or unity.

The score of individual \( e \) on the total battery of items is then

\[
S_e = X_{e1} + X_{e2} + \cdots + X_{en}
\]

or simply the total number of items answered correctly.

The correlation between the total test scores and the criterion is given by:

\[
\rho_{yx} = \frac{N \sum Y S - \sum Y \sum S}{\sqrt{N \sigma_y^2 N \sum S^2 - (\sum S)^2}}
\]

(2)

If we let

\( M_y = \) mean of criterion scores

\( M_x = \) mean of total test scores