RB-50-34

RESEARCH BULLETIN

ESTIMATING TOTAL TEST RELIABILITY FROM PARTS OF UNEQUAL LENGTH

Paul Horst

(To be submitted for publication)

Educational Testing Service

March 9, 1950

Estimating Total Test Reliability —

from Parts of Unequal Length

It sometimes happens that two part scores are available for a test when the parts are not of equal length. This can happen in a number of ways. For example, the two scores may be those from the front and back of an answer sheet respectively, or from a first and second page or group of pages of a test booklet. Or the scores may be from separately timed parts of a test where unequal time limits were used.

The Spearman-Brown split-half formula has long been used for estimating the reliability of a test from the correlation between its two comparable halves. This formula is given by:

$$R = \frac{2r}{1 + r} \qquad (1)$$

where R is the reliability of the test and r is the correlation between the two parts. However, from practical considerations of administration, scoring, format, or for other reasons, it may be more convenient to get part scores from parts of unequal length.

Assume that we have two such part scores and the two parts are comparable in all respects which make formula (1) applicable save for length. We can still estimate the reliability of the total test from the correlation between the two parts if we know what proportion of the total test is taken up by each part. The formula for this estimate is given by:

$$R = \frac{r\sqrt{r^2 + 4Pq(1 - r^2)} - r}{2Pq(1 - r^2)} \qquad (2)$$

where P and q are the proportions of total testing time taken up by each part respectively, r is the correlation between the two parts and R is the estimated reliability of the total test.

It is easy to show that formula (1) is a special case of formula (2). If the two parts are equal then P and q are both $\frac{1}{2}$. We substitute $\frac{1}{2}$ for P and q in formula (2) and get

$$R = \frac{r\left(\sqrt{r^2 + (4\frac{1}{2} \times \frac{1}{2})(1 - r^2)} - r\right)}{2\frac{1}{2} \times \frac{1}{2}(1 - r^2)}$$

or

$$R = \frac{r\sqrt{r^2 + 1 - r^2} - r}{\frac{1}{2}(1 - r)(1 + r)}$$

or

$$R = \frac{r(1 - r)}{\frac{1}{2}(1 - r)(1 + r)}$$

or

$$R = \frac{2r}{1 + r}$$

which is the same as formula (1).

To illustrate how the formula works, let us assume we have a hundred-item test which is divided into two comparable sets of items, one of which has 70 items and the other 30. Then P and q are .70 and .30 respectively. Assume the correlation between the two parts is .60. Substituting these values in formula (2) will give an estimate of the reliability of the entire test.

$$R = \frac{.60\left(\sqrt{(60)^2 + 4 \times .30 \times .70\ (1 - (.60)^2)} - .60\right)}{2 \times .30 \times .70 \times (1 - (.60)^2)}$$

$$= .775$$

If the test had been evenly divided with a correlation of .60 between the parts the estimated reliability of the total test would have been .75 or .025 less than with the thirty-seventy split.

The table gives the value of the estimated reliability for selected values of r and P. It does not matter which of the two proportions is taken for P. An examination of the table shows that for any value of r the estimated reliability increases as the deviation from a fifty-fifty split increases.

Furthermore the deviation of P from .50 has a greater effect for small values of r than for large. For example, when r is .10, the estimated reliability is .284 for P = .10 and .182 for P = .50 or a difference of .102. When r is .90, the estimated reliability is .972 for P = .10 and .946 for P = .50 or a difference of only .026.

The proof of formula (2) is as follows: assume we have available a number of comparable test units. Consider the reliability of a test made up of n of these units. It will be the correlation between the sum of n of the units and another set of n of the units. This is well known to be

$$R = \frac{n\rho}{1 + (n - 1)\rho} \qquad (3)$$

where R is the reliability of the test of n units and ρ is the correlation

between any pair of units. Formula (3) is, of course, the well known Spearman-Brown formula.

Next consider the correlation between the sum of a of the units and b of the units. This correlation is known to be

$$r = \frac{\sqrt{a\ b}\ \rho}{\sqrt{1 + (a - 1)\rho}\ \sqrt{1 + (b - 1)\rho}} \tag{4}$$

Assume now that $a + b = n$. By means of formulas (3) and (4) we can get the relationship between the estimated reliability of the test and the correlation between its two parts of length a and b respectively. We solve (3) for ρ which gives

$$\rho = \frac{R}{n + (1 - n)R} \tag{5}$$

Substituting the value of ρ given by (5) in (4) gives

$$r = \frac{\sqrt{a\ b}\ R}{\sqrt{n + (a - n)R}\ \sqrt{n + (b - n)R}} \tag{6}$$

Dividing the numerator and denominator of (6) by n and remembering that $\frac{a}{n} = P$ and $\frac{b}{n} = q$, we get

$$r = \frac{\sqrt{Pq}\ R}{\sqrt{1 - qR}\ \sqrt{1 - PR}} \tag{7}$$

Multiplying (7) by the denominator on the right, squaring both sides, transposing and arranging terms in powers of R we get

$$Pq(1 - r^2)R^2 + r^2R - r^2 = 0 \tag{8}$$

Solving (8) for R gives

$$R = \frac{-r^2 \pm \sqrt{r^4 + 4r^2Pq(1 - r^2)}}{2Pq(1 - r^2)}$$

Since R must be positive we take the plus sign in (8) and rewrite

$$R = \frac{r\left(\sqrt{r^2 + 4Pq(1 - r^2)} - r\right)}{2Pq(1 - r^2)}$$

which is the same as formula (2).

Estimated Total Test Reliabilities for Given Values of Part Correlations and Given Proportions of Part to Total Test

r	10%	20%	30%	40%	50%
.10	.284	.222	.197	.185	.182
.20	.487	.396	.357	.339	.333
.30	.634	.536	.490	.468	.461
.40	.741	.648	.601	.579	.571
.50	.819	.738	.695	.674	.667
.60	.878	.812	.775	.756	.750
.70	.918	.874	.844	.829	.822
.80	.949	.922	.902	.893	.889
.90	.972	.959	.954	.950	.946