CHAPTER XI

THE ESTIMATION OF TEST VALIDITY

In the previous chapter, the reliability of the present test was estimated by applying different methods and finally it was concluded that the reliability of the present test is 0.80 and that it is quite satisfactory as far as our purpose is concerned. This test is, then, a reliable one.

But a test constructor or a test-user will not rest content by knowing only the reliability of the test. He would like to know something more about the test. He will like to inquire, "Does this test measure what it is meant to measure?" Unless he gets a satisfactory answer to this question, he will not recommend anybody to use the test for any definite purpose.

In our every day life we need different types of measures for measuring lengths and weights. For this, the units of measurements and the instruments measuring these units are standardised. This is done by some central authority - in most of the cases - the nation's Government. The Government prepares such standard scales and puts in the market for common use. By a rule of the Government the people are required to use only these scales, the validity of which is testified by the Government. Similarly, the validity of a test also should have been testified by the test constructor.

There are a number of tests which measure everything else except what they are meant to measure. Such tests are of no use as they misguide equally both the tester and the testee. The tests whose validity has not been established or is doubtful should not be put in the market.

A question may arise, at this stage, then, 'what is a valid test?'

A simple answer will be this: A valid test is one, which measures what it purports to measure.

Or in other words the validity of a test depends on the efficiency with which it measures what it attempts to measure. Ross,[1] while defining validity says,

> One kind of validity concerns the degree to which the test or other measuring instrument measures what it claims to. In a word, validity means truthfulness.

Gulliksen[2] defines it in a little different manner. He says, ".....the validity of a test is the correlation of the test with some criterion".

Validity, then, refers to the truthfulness of the test and is always its most important characteristic. No matter what other merits the test may possess, if it lacks validity, it is worthless.

1 Ross,C.C., "Measurement in To-day's Schools", Prentice-Hall, Inc.,New York, 1955, p.107.

2 Gulliksen, Harold, "Theory of Mental Tests", John Wiley & Sons, Inc., New York, 1950.p.88.

Validity is a relative term. There is no such thing as general validity. The validity of the measuring instrument must always be considered in relation to the purpose it is to serve. Validity is always specific, in relation to some definite situation. In a test of arithmetic computation, for example, the validity of the test may be high for determining the present status of pupils in skills of arithmetic computation. Its validity may be moderate for judging aptitudes of pupils for business arithmetic. Its validity may be low for predicting success in the mathematical aspects of a subsequent course of study in physics. It can thus be said that a test is not just valid; it is valid for something.
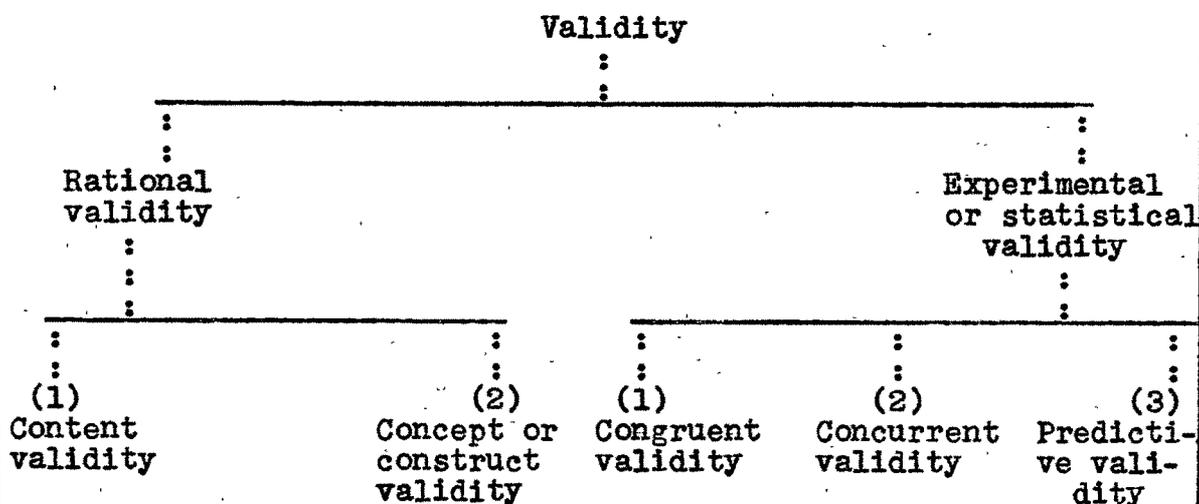
The validity of a test is determined experimentally by finding the correlation between the test and some independent criterion. A criterion may be an objective measure of performance, or a qualitative measure such as a judgment of the character or excellence of work done. Validity information indicates the degree to which the test is capable of accomplishing certain aims. Tests are used for several types of judgment, and for each type of judgment a somewhat different type of validating evidence is required. It should also be noted that independent standards are hard to get in mental measurement, and therefore, the validity of a mental test can never be estimated as accurately as can the validity of a physical instrument.

KINDS OF VALIDITY

Since validity is not an absolute characteristic,

several kinds of validity may be identified,depending upon the purpose for which the evaluative method is to be used. According to a report prepared by a joint committee of the American Psychological Association, American Educational Research Association, and National Council on Measurements used in Education, four types of validity have been distinguished, namely Content validity, Concurrent validity, Predictive validity and Construct validity. 'Factorial validity' may be added to this list.

Or according to Thorndike and Hagen,[1] there are two main types of evidence bearing on the validity of a test, rational and empirical. They sub-divide these two main types of validity in the following way:

Validity

Rational validity

Experimental or statistical validity

(1) Content validity

(2) Concept or construct validity

(1) Congruent validity

(2) Concurrent validity

(3) Predicti- ve vali- dity

Each of the different types of validity is discussed

1 Thorndike, R.L. and Hagen, Elizabeth, "Measurement and Evaluation in Psychology and Education", John Wiley & Sons, Inc., New York, 1958, p. 109.

in brief in the following pages.

(1) CONTENT VALIDITY

Content validity is concerned with the sampling of a specified universe of content. That is to say an analysis of the topics and areas included in the test is made. The analysis is done to show how well the content of the test samples the class of situations or subject matter about which conclusions are to be drawn. Barr, Davis and Johnson[1] say,

> Logical (content) validity is obtained
> when an investigator defines and describes
> the abilities, traits, concepts, or skills
> that he expects to be measured by an instru-
> ment of research, analyses them to identify
> the elements needed in a measuring instrument
> and designs the instrument with the demands
> of the situation as his criteria.

(2) CONCEPT OR CONSTRUCT VALIDITY

When the analysis of the activities and processes that correspond to a particular concept is made, the concept or construct validity is tried to judge through the analysis. To have a concept or construct validity, the test items must be specific, concrete and precise. They must consist of definite limited tasks. As Thorndike and Hagen[2] put it, the problem of preparing a test that has concept or construct validity is that of bridging the gap from broad general concept to specific

---

1 Barr, A.S., Davis, R.A., & Johnson, P.O., "Educational
        Research and Appraisal", J. B. Lippincott Company,
        Chicago: Philadelphia: New York, 1953, p. 93.

2 Thorndike, R.L., and Hagen, Elizabeth, Op.Cit., p. 112.

tangible tasks or test items.

(3) CONGRUENT VALIDITY

This type of validity refers to evidence of validity obtained by correlating a test with an existing similar measure of the same function. The validity of a new test obtained by this method is satisfactory only if the validity of the test with which it is compared, is testified.

(4) CONCURRENT VALIDITY

Concurrent validity is concerned with the relation of test scores to an accepted contemporary criterion of performance on the variable which the test is intended to measure.

(5) PREDICTIVE VALIDITY

Predictive validity is concerned with the relation of test scores to measures on a criterion based on performance at some later time. It can, thus, be said that concurrent validity refers to the relation between test scores and criterion scores which can be obtained at the same time, while predictive validity refers to the relation between test scores and criterion scores which cannot be secured until some time later.

This type of validity is required in such measures as tests of intelligence or academic aptitudes used for predicting later scholastic success; in tests of aptitudes used for

predicting later success in professions such as Engineering,

Law, Medicine, Teaching, Music, Art, etc.

(6) FACTORIAL VALIDITY

The validity of a given test is defined by its factor

loadings - and these are given by the correlation of the test

with each other. A vocabulary test, for example, may correlate

0.85 with the verbal factor extracted from the entire test

battery. This coefficient becomes the test's factorial

validity.[1]

VALIDITY OF THE PRESENT TEST

The purpose of the present test is to measure aptitude

for teaching and thereby to forecast whether the testee will be

a successful teacher. The purpose of the test itself, thus,

suggests that the predictive validity of the test should be

estimated. The predictive validity of the test is, then, the

main aspect of this test. Besides this the content validity

of the test is obtained and to establish the congruent validity

of the test, an attempt is made to carry out a factor analysis

of the test data. (This is discussed in chapter XIII of this

treatise).

CONTENT VALIDITY OF THE PRESENT TEST

In chapter VI, on test construction, the whole

1 Garrett, H. E., "Statistics in Psychology and Education",
       Longmans, Green & Co., New York, 1958, p.356.

procedure adopted for item selection, has been discussed in full details. Each of the five sub-tests in the whole battery was discussed with the experts. There was a consensus of opinion on the selection of the five factors which are supposed to condition success in teaching in secondary schools. Items included in sub-tests - each sub-test measuring one factor - were discussed with the experts to judge the fitness of their being included in the sub-test and to decide whether each item would measure what it is meant to measure. All the items in all the sub-tests were, thus, included after a careful scrutiny. Over and above this, all the items were validated against the two criterion groups. Only the valid items were selected while the items that were not found to be valid were summarily rejected. The content validity of the test was, thus obtained to the full satisfaction of the investigator, his guiding professor and all other experts who were consulted.

Concept or construct validity, as is enunciated by Thorndike and Hagen, is not needed to be obtained, as no concept is to be measured here.

Congruent validity, as is enunciated by Thorndike and Hagen (or construct validity as the joint committees' report term it) is not possible to be estimated as no aptitude test of the same nature is available in this country.

Concurrent validity also, is not possible to be estimated, as the present test is designed to forecast a

performance in teaching at some later time, while the con-
current validity is concerned with the contemporary criterion
of performance.

PREDICTIVE VALIDITY OF THE PRESENT TEST

The definition of the predictive validity reveals
that to estimate this type of validity the test scores are
correlated with the measures on a criterion based on perform-
ance at some later time. It is, therefore, clear that some
external criterion is required to estimate predictive validity
of a test.

THE CRITERION

It is discussed in the earlier pages of this chapter
that the validity of a test is the closeness of agreement
between the scores and some other objective measure of that
which the test is designed to measure.   This other measure is
called the criterion.   The coefficient of validity of a test
is the coefficient of correlation between test scores and
criterion scores.

Good objective criteria of occupational success are
extremely difficult to define and to measure.   The fact that
tests of vocational aptitude so frequently have low coefficients
of validity is not alone due to the difficulty of devising
reliable tests of the aptitudes it is desired to measure.
Adequate and reliable criteria of success against which to

validate the tests are no less difficult to define and to
measure.

The criterion, against which the present test is
validated is the total marks - total of part I and part II
marks - obtained by a testee on the University examination.

The topics on the job analysis and the selection of
the criterion are discussed in full details in chapter VI of
this treatise.   The repetition of the discussion is out of
place here.   But a mention should be made here that the
criterion was selected after a much careful scrutiny and that
all the experts agreed that the criterion is satisfactorily
reliable and valid.   It should, of course, be noted that no
criterion, howsoever, carefully scrutinised and selected, is a
perfect one or is even to optimum satisfaction of the experts.
Remembering that it is extremely difficult to fix up criteria
to judge success in teaching, the present criterion also cannot
be a perfect one; but the investigator is confident that it is
reliable and valid to the extent it is possible to reach in the
present circumstances and limited facilities for doing research
work in this country.

It is also discussed in the same chapter, while
discussing the criterion, that different training colleges are
affiliated to different Universities and that the nucleus of
the syllabus followed in all the four different Universities -
M.S, University, Baroda, Gujarat University, Ahmedabad, Bombay

University, Bombay and Poona University, Poona - being the same there is little change in the number of papers in part I examination, in conducting Part II examination and in assigning internal marks.

Different University examinations, therefore, carry different total marks - total of part I marks and part II marks. Naturally, the raw scores obtained by different testeesfrom different colleges are not comparable. They had to be converted into standard scores before they were used to correlate them with the test scores. (The procedure of converting the raw scores into standard scores is discussed in chapter XIV on Norms).

There is no system of part II examination in Bombay University. Rule no. 69 in Pamphlet[1] No.3 of the same University reads as follows:

R.69: "There shall be no external examination in practical skill, in class management and class teaching of candidates."

And rule no. 71 in the same pamphlet reads as follows:

R. 71: "The practical work done by a candidate during his year's training will be classed by the Head of the

---

1 University of Bombay, Ordinances and Regulations relating to the examination for the Degree of Bachelor of Education (B.Ed.), Pamphlet No. 3, 1956, p.21.

Institution on the strength of the year's record by 5 grades
A, B, C, D, and E.   'A' being the best and 'E' the lowest
considered as failure."

Part I examination marks of testees from the S.T.T.
College, Bombay, and St.Xavier's Institute of Education,Bombay,
were available.   This naturally reduced the sample (N) for
correlating the criterion scores with the test scores from 530
to 358.

This situation necessitated the selection of two
sub-criteria for validating the test scores, these being (1)
part I marks and (2) part II marks.   The total sample for Part
I criterion was then of (N) 530 and that for part II criterion
of (N) 358).

The raw scores on total test obtained by all the
testees were also converted into the standard scores.   Thus
for calculating the Pearson product-moment coefficient of
correlation, the standard scores on criterion and the standard
scores on the aptitude test of all testees were used.

CALCULATION OF THE VALIDITY COEFFICIENT
BETWEEN THE CRITERION SCORES AND TEST SCORES

The two sets of scores were arranged in the form of
a scatter diagrom and the product-moment coefficient of
correlation was calculated.

TABLE NO. 56

SHOWING A SCATTER-DIAGRAM BETWEEN THE CRITERION AND THE TEST SCORES

Test-Score (Y-variable)	Criterion score (X-variable)													fy
	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119	120-129	130-129	140-149	150-159	160-169	
150-159	-	-	-	-	-	-	1	-	-	-	-	-	1	2
140-149	-	-	-	-	-	-	4	3	2	1	-	-	-	10
130-139	-	-	-	-	2	-	3	2	5	3	-	2	-	17
120-129	-	-	-	1	2	2	7	4	4	2	2	1	-	25
110-119	-	-	1	-	4	13	10	11	6	5	1	-	1	52
100-109	-	-	1	2	11	17	12	14	3	3	4	-	-	67
90- 99	-	1	8	5	20	15	16	16	4	-	-	-	-	85
80- 89	-	1	2	10	10	19	10	2	2	-	-	-	-	56
70- 79	2	1	-	3	6	8	2	-	-	-	-	-	-	22
60- 69	-	-	-	2	4	1	2	2	-	-	-	-	-	11
50- 59	1	-	2	3	1	2	1	-	-	-	-	-	-	10
40- 49	-	-	-	-	-	1	-	-	-	-	-	-	-	1
fx	3	3	14	26	60	78	68	54	26	14	7	3	2	358

The product-moment coefficient of correlation, 'r', was calculated according to the usual procedure from the above scatter-diagram.   It was found to be 0.502.   Probable error of this 'r', also, was calculated.   It was found to be ± 0.026.   The predictive validity of the present test is, therefore, 0.502 ± 0.026.   This validity coefficient is interpreted at a later stage in this chapter.

CALCULATION OF THE COEFFICIENT OF
CORRELATION BETWEEN PART I - SCORES
AND   TEST   SCORES

The two sets of scores were arranged in the form of a scatter-diagram as shown in the table given on the next page.

CALCULATION OF THE COEFFICIENT OF
CORRELATION BETWEEN PART II -
SCORES AND TEST SCORES

The two sets of scores were arranged in the form of a scatter-diagram as shown in the table given on the next page.

The above obtained results are tabulated below.

TABLE NO. 59

VALIDITY COEFFICIENTS BETWEEN CRITERION SCORES
AND   TEST   SCORES

Test	Criterion 'C' (N = 358)	Part I- Score (N = 530)	Part-II Score (N = 358)
Total Test -score	0.502 ± 0.026	0.448 ± 0.023	0.428 ± 0.029

SCATTER-DIAGRAM BETWEEN PART-I SCORES AND TEST SCORES

Test-Score (Y-variable)	Part-I – score (X-variable)													
	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119	120-129	130-139	140-149	150-159	160-169	fy
150-159	-	-	-	-	1	-	-	-	-	-	-	-	1	2
140-149	-	-	-	-	-	4	3	1	4	2	1	-	-	15
130-139	-	-	-	-	1	2	3	-	7	2	2	1	-	18
120-129	-	-	-	2	1	10	12	8	8	5	1	1	-	48
110-119	-	1	-	3	7	21	15	17	10	5	2	1	-	82
100-109	-	-	2	8	14	24	22	13	11	6	2	1	-	103
90-99	1	3	3	15	19	15	33	13	8	2	-	-	-	112
80-89	-	-	4	10	16	15	22	2	3	-	-	-	-	73
70-79	1	2	6	5	8	13	5	2	-	-	-	-	-	42
60-69	-	-	1	4	8	3	2	2	2	-	-	-	-	22
50-59	-	2	2	2	3	1	2	-	-	-	-	-	-	12
40-49	-	-	-	-	1	-	-	-	-	-	-	-	-	1
fx	4	6	18	49	78	109	119	58	53	23	8	3	2	530

Product-moment, 'r', = 0.448   P.E.r = ± 0.023

TABLE NO. 58

SCATTER-DIAGRAM BETWEEN PART II - SCORES AND TEST-SCORES

Part-II - score (X-variable)

Test-score (Y-variable)	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119	120-129	130-139	140-149	150-159	160-169	fy
150-159	-	-	-	-	-	-	-	-	1	-	1	-	-	2
140-149	-	-	-	-	-	3	3	1	2	1	-	-	-	10
130-139	-	-	-	-	2	2	4	3	2	3	1	-	-	17
120-129	-	-	-	2	5	4	6	3	2	2	1	-	-	25
110-119	-	-	-	4	10	9	10	9	4	4	2	-	-	52
100-109	-	-	4	10	10	19	14	4	4	2	-	-	-	67
90- 99	-	1	4	10	19	20	16	9	3	4	3	1	1	86
80- 89	-	-	-	10	14	14	13	2	1	-	1	-	-	55
70- 79	1	1	1	1	1	3	9	2	3	-	1	-	-	22
60- 69	-	-	2	2	2	3	2	-	-	-	-	-	-	11
50- 59	-	-	1	2	2	2	2	-	1	-	-	-	-	10
40- 49	-	-	-	-	-	1	-	-	-	-	-	-	-	1
fx	1	2	11	35	69	71	69	41	28	19	7	4	1	358

Product-moment, 'r' = 0.428

P.E.r = ± 0.029

The coefficients of correlation between criterion score and each of the five sub-tests scores were calculated. The coefficients obtained, are shown in the following table.

TABLE NO. 60

COEFFICIENTS OF CORRELATION BETWEEN CRITERION
SCORE AND EACH OF THE SUB-TESTS SCORE

Sub-test No.	Criterion 'C' (N = 358)	Part I Score (N = 530)	Part II Score (N = 358)
I	0.380 ± 0.031	0.355 ± 0.026	0.265 ± 0.033
II	0.383 ± 0.030	0.277 ± 0.027	0.386 ± 0.030
III	0.100 ± 0.035	0.126 ± 0.029	0.100 ± 0.035
IV	0.316 ± 0.032	0.233 ± 0.028	0.279 ± 0.033
V	0.237 ± 0.034	0.313 ± 0.026	0.179 ± 0.034

THE INTERPRETATION OF VALIDITY COEFFICIENT

Is the validity coefficient of the test satisfactory? Is the test a good predictor of aptitude for teaching?

It is difficult to answer such questions. But one point can well be emphasized. Validity is always specific to a particular job.

Bingham[1] says,

Coefficients of correlation between
tests of aptitude and measures of subse-
quent educational or vocational achievement

1 Bingham, W.V.D., "Aptitudes and Aptitude Testing", Harper &
Brothers Publishers, New York, 1942, p.213.

are apt to be nearer to zero than one,
and the zone of possible error in
prediction is correspondingly wide.

Most vocational aptitude tests have lower validities.
This is to be expected.   A person's occupational achievement
is due to a great variety of causes in addition to his aptitudes.
Moreover, any quantitative measure of the degree of that
achievement is at beat a rough one and subject to error.   In
other words the criteria are not always adequate or highly
reliable.

The present one is an aptitude test for teaching.
Before drawing any conclusion regarding its validity coefficient
(0.502), it will be helpful and advisable to study validity
coefficients of other studies in this field.

(1) Bowers[1] reports a validity coefficient of 0.70
for his ATEST battery.

(2) The scores of 135 senior and graduate
students on the Coxe-Orleans Prognosis
Test of Teaching Ability were compared
by Dodd (391) with scholarship and with
supervisors' ratings of teaching
ability at the close of the practice-
teaching period. Correlations of
0.425 with supervisors' ratings and
0.636 with scholarship were obtained
when raw scores rather than weighted
scores were used.   Part II of the test
- a true-false test of professional
interest - had a larger correlation
(0.454) than the entire test with

----

1 Bowers, Henry, "Aptitude Test for Elementary School
        Teachers - In - Training", J. M. Dent & Sons,
        Limited, Toronto, 1948, p.8.

supervisors' ratings.[1]

(3) Cain, Michaelis & Eurich[2] say,

> Correlations vary considerably as
> noted in the following examples report-
> ed : in various studies: practice
> teaching, 0.06 to 0.81; mental ability,
> 0.17 to 0.61; age, 0.01 to 0.34;
> experience, 0.10 to 0.39; high-school
> marks, 0.08 to 0.77; personality, 0.25
> to 0.83; scholarship in college, 0.15
> to 0.72.   No single factor can be
> used successfully as a predictor.

(4) Barr[3] in a study of the validity of a number

of tests of teaching ability, employing a composite criterion,

secured coefficients of validity as follows: for the Knight

Aptitude Test, 0.43; the American Council Psychological

Examination, 0.36; the Torgerson Professional Information test,

0.35; and the Morris Trait Index L, 0.22.

(5) In the following table[4], some more validity

coefficients of studies in the same field are given.

---

1 Toops, Herbert and Kuder, G.F., "Measures of Aptitude",
        Review of Educational Research,Vol.V,June 1935,No.3,
        p.225.
2 Monroe,W.S.,"Encyclopedia of Educational Research",(revised
        edition),The MacMillan Company: New York,1956, p.889.
3 Monroe,W.S., Ibid., p. 1447.
4 Adapted in part from Sandiford,Peter & others."Forecasting
        Teaching Ability", Bulletin No.8, Department of
        Educational Research,University of Toronto,1937,p.25.

TABLE NO. 61

CORRELATIONS OF JUDGMENTS OF TEACHING
SUCCESS WITH VARIOUS TESTS

Test	Investigator	Criterion of teaching success	'r'
Professional information test	Boardman	Supervisor's ratings	0.26
Professional information test	Bossing	Teaching success in the field	0.19
George Washington Teaching Aptitude Test	Hunt	Teaching success in the field	0.3 & 0.5
Coxe-Orleans Prognosis test	Dodd	Teaching success in the field	0.41

If the validity coefficient of the present test is compared with those mentioned above, it can be said with some degree of confidence that it is much satisfactory and the test is a good predictor of aptitude for teaching.

Even then, obviously, other things being equal, the higher the correlation, the better.

According to Bingham[1], validity of prediction can be increased (1) by improving the adequacy and reliability of the measures of success used as criteria; (2) by choosing those traits for testing which are most indicative of aptitude for achieving the success specified; (3) by choosing the tests which most accurately measure those traits; and (4) by ascertain -ing the proper weight to be assigned to the scores in each

1 Bingham, W.V.D., Op.Cit., pp.219-220.

test, in order most closely to predict the criterion.

While constructing the test, utmost care was taken in selecting factors indicative of aptitude for teaching in constructing test items and finally in securing adequate and reliable criterion. The investigator, therefore, legitimately, feels that nothing substantial would be added to the test validity even if an attempt is made to apply the first three suggestions made by Bingham to increase it.

As regards the fourth suggestion, it may be said that the proper regression weights to be assigned to the test scores, are found out. (This topic is discussed at length in the next chapter).

There is one more suggestion from Gulliksen[1] to increase the test validity. He says, "Multiplying the length of a test by K increases the validity coefficient." He gives the following formula:

$$R_{k1} = \frac{r_{cx} \sqrt{K}}{\sqrt{1 + (K-1) r_{11}}}$$

Where R_{k1} = the augmented validity coefficient,

r_{cx} = validity coefficient of the test,

r_{11} = reliability coefficient of the test,

K = the number of times the test is increased in length.

---

1 Gulliksen, Harold, "Theory of Mental Test", John Wiley & Sons, Inc., New York, 1950, pp. 88-89.

If Gulliksen's formula is applied, it will be seen that the increase in the test length by three times, will add to the test validity only by 0.026. This does not raise the test validity substantially at all. It will be inadvisable and uneconomical to increase the test length three times for such a meagre gain in test validity.

The validity of the present test is, therefore, finally accepted to be 0.502 ± 0.026.

COEFFICIENTS OF CORRELATION BETWEEN
CRITERION SCORES AND TEST SCORES FOR
DIFFERENT UNIT-SAMPLES

A number of coefficients of correlation of the following types were calculated:

(1) r's between criterion scores and test-scores
for unit-samples, A,B,C,& F.

(2) r's between Part I scores and test scores for
all unit samples.

(3) r's between part II scores and test scores for
unit samples A, B, C, & F.

(4) r's between criterion scores and scores on
each of the sub-tests for unit-samples A,B,C,& F

(5) r's between Part I scores and scores on each of
the sub-tests for all unit samples.

(6) r's between Part II scores and scores on each of
the sub-tests for unit-samples A,B,C, & F.

All these coefficients of correlation are shown in the tables given on the next page, prepared unitwise.

These results show that validity coefficients obtained from different unit-samples agree in size to a great extent with the validity coefficient obtained for the test for the whole sample. The coefficients for sub-tests do not agree to as much extent, there being variations in sizes of r's at some places a little, while at some places, significantly large. Even then, on the whole, it can be concluded that r's obtained for different unit-samples are satisfactorily of the same size allowing some degree of divergence here and there.

Anyway these results help us arriving at the following tentative conclusions:

(1) The test as a whole is a satisfactory predictor of the aptitude for teaching.

(2) The validity coefficient of the test is fairly stable as the unit-samples also give the same sizes of validity coefficients.

(3) As the validity coefficients obtained for different unit-samples are fairly of equal sizes, the distribution of aptitude for teaching is normal.

(4) The results prove the representativeness of the total sample and also of the unit-samples.

TABLE No. 62

UNIT A - COEFFICIENTS OF CORRELATION

Sr. No.	Criterion Score	N	Total test score	Sub-Test I score	Sub-Test II score	Sub-Test III score	Sub-Test IV score	Sub-Test V score
1	Criterion 'C' Score	100	0.520 ± 0.049	0.336 ± 0.059	0.451 ± 0.054	0.233 ± 0.064	0.301 ± 0.061	0.224 ± 0.064
2	Part- I - score	100	0.530 ± 0.048	0.405 ± 0.056	0.394 ± 0.057	0.226 ± 0.064	0.289 ± 0.062	0.274 ± 0.062
3	Part - II - score	100	0.360 ± 0.059	0.160 ± 0.066	0.402 ± 0.056	0.177 ± 0.065	0.234 ± 0.064	0.190 ± 0.065

TABLE No. 63

UNIT B - COEFFICIENTS OF CORRELATION

Sr. No.	Criterion Score	N	Total Test score	Sub-Test I score	Sub-Test II score	Sub-Test III score	Sub-Test IV score	Sub-Test V score
1	Criterion 'C' Score	74	0.480 ± 0.060	0.433 ± 0.063	0.210 ± 0.075	0.088 ± 0.078	0.349 ± 0.069	0.284 ± 0.072
2	Part - I - score	74	0.350 ± 0.069	0.339 ± 0.069	0.137 ± 0.077	0.036 ± 0.078	0.306 ± 0.071	0.211 ± 0.075
3	Part - II - score	74	0.480 ± 0.060	0.355 ± 0.068	0.263 ± 0.073	0.122 ± 0.077	0.296 ± 0.071	0.316 ± 0.071

TABLE No. 64

UNIT C - COEFFICIENTS OF CORRELATION

Sr. No.	Criterion Score	N	Total Test score	Sub-Test I score	Sub-Test II score	Sub-Test III score	Sub-Test IV score	Sub-Test V score
1	Criterion 'C' score	66	0.479 ± 0.064	0.409 ± 0.069	0.239 ± 0.078	0.241 ± 0.078	0.336 ± 0.074	0.271 ± 0.077
2	Part - I - score	66	0.433 ± 0.067	0.277 ± 0.076	0.220 ± 0.079	0.295 ± 0.076	0.199 ± 0.079	0.330 ± 0.074
3	Part - II - score	66	0.330 ± 0.074	0.337 ± 0.074	0.183 ± 0.080	0.123 ± 0.082	0.335 ± 0.074	0.095 ± 0.082

TABLE NO. 65

UNIT D - COEFFICIENTS OF CORRELATION

Sr. No.	Criterion Score	N	Total test score	Sub-Test I score	Sub-Test II score	Sub-Test III score	Sub-Test IV score	Sub-Test V score
1	Part - I - score	94	0.520 ± 0.051	0.316 ± 0.063	0.192 ± 0.067	0.302 ± 0.063	0.215 ± 0.066	0.459 ± 0.055

TABLE NO. 66

UNIT E - COEFFICIENTS OF CORRELATION

Sr. No.	Criterion Score	N	Total test score	Sub-Test I score	Sub-Test II score	Sub-Test III score	Sub-Test IV score	Sub-Test V score
1	Part - I - score	78	0.484 ± 0.058	0.340 ± 0.068	0.255 ± 0.71	0.153 ± 0.075	0.095 ± 0.076	0.523 ± 0.056

TABLE NO. 67

UNIT F - COEFFICIENTS OF CORRELATION

Sr. No.	Criterion Score	N	Total test score	Sub-Test I score	Sub-Test II score	Sub-Test III score	Sub-Test IV score	Sub-Test V score
1	Criterion 'C' score	118	0.520 ± 0.045	0.374 ± 0.053	0.498 ± 0.046	-0.015 ± 0.062	0.304 ± 0.056	0.260 ± 0.058
2	Part - I- Score	118	0.450 ± 0.049	0.414 ± 0.051	0.370 ± 0.053	-0.085 ± 0.062	0.258 ± 0.058	0.263 ± 0.058
3	Part - II - Score	118	0.510 ± 0.046	0.260 ± 0.053	0.556 ± 0.043	0.060 ± 0.062	0.273 ± 0.057	0.223 ± 0.059

In the next chapter, we shall discuss regression weights and prediction formula to forecast success in teaching in secondary schools.