CHAPTER XI

VALIDITY

It is the purpose of this chapter to emphasize several measures of the validity of the present test. Validity is one of the most important of the characteristics of a standardised test. McCall\(^1\) says that it is the most fundamental trait. Validity may be defined as the correspondence between the ability measured by a test and the ability as elsewhere objectively defined and measured. "When a test really measures what it purports to measure and consistently measures it throughout the entire range of the test, it is a valid test".\(^2\)

An assembly of items tested for internal consistency and possessing a high degree of reliability is not a difficult task to accomplish. But this will not constitute a valid test as nothing is known about what it measures. There are two common criticisms about the validity of a test,

---

1 W. A. McCall, "How to Measure in Education," Page 129
   New York: The Macmillan Company

2 H. E. Garrett, "Statistics in Psychology and Education," Page 34
   New York: Longmans, Green and Co.
viz., (i) either it is of a doubtful nature or (ii) it is, at times, spuriously high. It is from these two points that the validity of a test should be evaluated.

**TYPES OF VALIDITY:**

There are four types of validity depending upon the type of criterion used to validate a test.

a. **Content validity**: It is also known as "Logical validity". It is specially applicable to achievement tests or tests of proficiency. These tests are designed to measure how well the individual has mastered a specific skill or course of study. "The validity of this type is evaluated by showing how well the content of the test samples the class of situations or subject matter about which conclusions are to be drawn".¹

b. **Predictive validity**: Predictive validity of a test is determined by correlating the test scores with a future criterion measure. A test with a high predictive validity will be useful in predicting the subsequent behaviour of an individual on the basis of his scores on the test. It requires years of longitudinal studies of individuals before predictive validity of a test is established.

¹ "Technical Recommendations," Page 13
Psychological Bulletin (Supplement), Vol.51, No.2, Part 2
c. Concurrent validity: It shows how well the test correlates with the present criterion and not a future one. The validity of a test studied on the basis of identifiable criterion groups belongs to this category.

d. Construct validity: The correlation of test scores with those of other similar tests gives a measure of construct validity of a test. Factorial analysis also yields data for construct validity. "It is evaluated by investigating what psychological qualities a test measures, i.e., by demonstrating that certain explanatory constructs account to some degree for performance on the test." ¹

STUDIES OF THE VALIDITY OF SOME ALLIED TESTS OF SOCIAL INTELLIGENCE:

1. Gilliland and Burke have prepared a questionnaire directed towards measuring 'sociability'. Its validity, as determined by correlation with the average of ratings for sociability given by members of a psychology class is .60; .50 and .43 in three different classes.²

2. Stauter and Hunting³ have prepared another

¹ "Psychological Bulletin," op. cit., Page 14


questionnaire to measure the same trait viz., Sociability. This questionnaire is validated by comparing those who received mention for non-academic and non-athletic activity in the school annual with those who had not. Differences between the mean scores of two groups were from five to eight times their standard error. Thus the questionnaire is sufficiently discriminating between two contrasted criterion groups.

3. R. Strang has prepared a Test of Social Usage for high school students. The test covers those aspects of personal social contacts which are generally included in manners and etiquette. It is found that scores in the test increase steadily from the seventh to twelfth grade. The test also differentiates those students whose parents fall into different occupational classes. The test does not show consistent relationship to school activities. It correlates to the extent of .40 to .50 with abstract intelligence.

4. Reed and Weidemann have prepared a "Social Situational Test". Nowhere the authors mention about the validity of this test. A study of this test shows a correlation of .16 with an intelligence test.

5. A test that is widely studied both by its authors and other investigators for its validity is *The George Washington Social Intelligence Test*. The authors of this test are F. Moss and T. Hunt. Three types of evidence have been presented by Hunt to establish the validity of this test.

a) Hunt\(^1\) found marked occupational differences in the test scores. Executives, salesmen and teachers made high scores while clerks and unskilled labourers made low scores.

b) Hunt\(^2\) also found a correspondence between the scores on the test and the number of co-curricular activities of college entrants as given in the following table:

<table>
<thead>
<tr>
<th>Activities</th>
<th>Median score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 or more</td>
<td>116</td>
</tr>
<tr>
<td>3</td>
<td>112</td>
</tr>
<tr>
<td>2</td>
<td>110</td>
</tr>
<tr>
<td>1</td>
<td>106</td>
</tr>
<tr>
<td>0</td>
<td>99</td>
</tr>
</tbody>
</table>

c) Hunt\(^3\) also established the validity of the test by correlating the scores on the test with rating by a superior executive for employees in a large company. The correlation coefficient was found to be .61

As against the evidence put forward by Hunt, the George Washington Social Intelligence Test is found to yield a correlation of only .14 to .17 with Gilliland's *Sociability questionnaire* showing that the two measures are not of the same thing.

---

1 Quoted by R. L. Thorndike and Saul Stein in, "The Psychological Bulletin," 1937, 5, 279-280
2 Ibid., 279-280
3 Ibid., 279-280
Again a number of investigations have tried to investigate the relationship of this test to tests of abstract intelligence. The correlation coefficients of this test with abstract intelligence tests vary from .25 to .69. These results show the doubtful nature of the validity of this test.

**VALIDITY OF THE PRESENT TEST:**

The validity of the present test has been studied in a number of ways.

**Content validity:**

As already mentioned, content validity depends upon the sample of items contained in the test. The present test is aimed to measure the social intelligence of individuals. This ability has been known to contribute to success in some specific professions viz., insurance agents, salesmen etc. The sub-tests composing the present test are all constructed on the basis of experts' opinion regarding the social skills required in people having a pronounced ability to deal with others. In chapter V, it was seen that before selecting the sub-tests, persons highly successful in occupations requiring ability to deal with others were approached to list a few skills that contributed to their success in professions. Twelve different abilities emerged from their opinions. Out of
these twelve abilities, six abilities that were ranked higher than others were selected for preparing sub-tests. The items selected in some of the sub-tests represent those situations where social skill comes into play. Again, the universe of items in the present test resembles the types of items usually found in already available measures of social intelligence. The nature of the items and the procedure of selecting the sub-tests are adequate to establish the content validity of the present test.

Concurrent validity:

Eventhough, a reference has been made to the content validity of the present test, it is not its chief feature. The main feature of the present test is the variety of ways in which its validity has been studied.

1. As already shown earlier, a laborious process was adopted to select an objective and reliable criterion to validate the items of the present test. The two contrasted criterion groups were formed after obtaining objective and reliable data on a mass scale. Each item of the test has been selected only after ascertaining that it discriminated between the two criterion groups. The validity of the total test also has been established on these two groups. When the final form of the test was ready, it was administered to the pupils in the two
criterion groups. Eighty-nine pupils in group A and ninety-three pupils in group B could be traced. The results of this study are given in the table below:

**TABLE 36**

**VARIETY DATA OBTAINED FROM TWO CRITERION GROUPS**

<table>
<thead>
<tr>
<th>Criterion groups</th>
<th>N</th>
<th>Mean Score</th>
<th>S.E. of Mean</th>
<th>Difference (D)</th>
<th>S.E. of C.R.</th>
<th>C.R.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The criterion group with high social intelligence</td>
<td>89</td>
<td>$M_1 = 10.60$</td>
<td>6.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>63.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The criterion group with low social intelligence</td>
<td>93</td>
<td>$M_2 = 12.7$</td>
<td>1.32</td>
<td>17.37</td>
<td>1.73</td>
<td>10.4</td>
</tr>
</tbody>
</table>

The difference between the mean performance of the two groups on the test is about ten times its standard error. This shows that the difference is highly significant. This study no doubt establishes the concurrent validity of the present test, yet it raises some doubts also. Writing about the high validity of tests, A. Anastasi\(^1\) writes:

---

\(^1\) A. Anastasi, "Psychological Testing," Page 166

Any validity coefficient computed on the same sample which was used for item selection purposes will capitalize on chance errors within that particular sample and will consequently be spuriously high. In fact, a high validity coefficient could result under such circumstances, even when the test has no validity at all in predicting the particular criterion.

As the present validation study has been carried out on the same criterion groups which were used for item selection purposes, there is a reason to believe that the result may be of a doubtful value. A cross-validation study was, therefore undertaken to further establish the true validity of the test.

**Cross-validation:** Cross-validation requires that the validity of a test should be determined on a different sample of persons from that on which the items were selected. The cross-validation of the test has been carried out in more than one study in the present case.

(i) In this study, two criterion groups were obtained from amongst the hostel students. Forty students were selected from the Boys' Hostels at Baroda and Vidyanagar. These students were such as were found making the best social adaptation of any in their group. These were the students who were active in the community life of the hostel and clearly having a high degree of sociability. The group was selected on the basis of judgments of residential students and the house-monitors. As against this group of students, an unselected group of students of the same strength was also taken. Students of both the groups were administered the test. The data is given in table 37.
TABLE 37
DATA OF TWO CRITERION GROUPS
OF HOSTEL STUDENTS

<table>
<thead>
<tr>
<th>Groups</th>
<th>No.</th>
<th>Mean Score</th>
<th>S.E.</th>
<th>Difference of mean M₁ &amp; M₂</th>
<th>S.E.</th>
<th>C.R. (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sociable students</td>
<td>40</td>
<td>M₁ = 9.16</td>
<td>1.44</td>
<td>12.9</td>
<td>3.15</td>
<td>3.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>69.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unselected group of</td>
<td>40</td>
<td>M₂ = 19.3</td>
<td>2.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hostel students</td>
<td></td>
<td>56.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The difference is about four times its standard error. This means that the difference is quite significant and it proves the sound validity of the test.

(ii) A second cross-validation study was undertaken with secondary school pupils. The socio-metric technique was used to secure the criterion groups. Two questions were used in this study. The questions were:

(1) "With whom would you like to be friendly?"

(2) "With whom would you like to play?"

This study was undertaken in ten different schools. These schools were different from those included for selecting criterion groups for item-validation. 62 peers and 73 isolates were sampled out. Although the samples are small, the results are interesting. The results are tabulated in table 38.
TABLE 38
DATA OF GROUPS OBTAINED ON THE BASIS OF SOCIOMETRIC STUDIES

<table>
<thead>
<tr>
<th>No.</th>
<th>Mean</th>
<th>S.E. Difference of Mean</th>
<th>S.E. of C.R.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peers</td>
<td>62</td>
<td>66.3</td>
<td>14.8</td>
</tr>
<tr>
<td>Isolates</td>
<td>73</td>
<td>50.7</td>
<td>15.2</td>
</tr>
</tbody>
</table>

It is seen that the difference between the mean scores of the two groups is very wide. It is more than six times its standard error. If social intelligence, sociability or social sensitiveness is a factor contributing to popularity in a group, it can be stated very confidently from this study that the present test has a sound validity.

2. Another study was undertaken to ascertain the validity of the test with a group of insurance agents. These agents were rated by their field officers on a five point scale with respect to their ability to deal with people. The criterion adopted was the amount of work put in by an agent during the last two years. This is a very objective criterion whose reliability cannot be questioned. On the basis of work put in, the field officers rated eighty-two agents under them. The five points on the scale were 'very superior', 'superior', 'average', 'inferior' and 'very inferior'. Those agents were already administered the test while the norms were being fixed.
Pearson's product-moment correlation was calculated between the ratings and the test scores. The product-moment coefficient of correlation was found to be .52 with a P.E. of .055.

3. The last evidence regarding the validity of the present test is gained from an examination of the mean scores of S.S.C. class pupils, general adults, salesmen, insurance agents and supervisors. The test was administered to an unselected sample of 420 adults drawn at random from Baroda. The mean scores of these various groups on the test are given in table 40.

**TABLE 40**

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Groups</th>
<th>N</th>
<th>Mean</th>
<th>Difference</th>
<th>S.E. of Difference</th>
<th>C.R.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>S.S.C. class pupils</td>
<td>4434</td>
<td>54.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Adults</td>
<td>420</td>
<td>62.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Supervisors</td>
<td>220</td>
<td>67.6</td>
<td>5.3</td>
<td>.967</td>
<td>5.4</td>
</tr>
<tr>
<td>4.</td>
<td>Insurance agents</td>
<td>267</td>
<td>68.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Salesmen</td>
<td>253</td>
<td>70.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each professional group has a mean score on the test which is significantly higher than that of the normal unselected adult sample. These professions employ persons who are expected to have a good ability to deal with people. The high mean scores of these groups on the present test as compared to the scores of
S.S.C. class pupils and the adults provide a supporting evidence regarding the validity of the present test.

4. There is one more aspect of the validity of the present test. This aspect relates to the validity of the sample. In chapter VIII, the method of selecting the sample for standardisation of the present test has been given in detail. In chapter X, the normality of the distribution of test scores has been discussed and checked. The chi-square test has been applied to check the 'goodness of fit' of the distribution. The deviation of the obtained distribution from normality is not at all significant. It can, therefore, be concluded with confidence that the sample of pupils tested is representative of the population for which the tests are designed.

**Construct validity:**

As there is no test of sociability or social intelligence available in Gujarat, the scores on the present test have not been correlated with scores on other tests of social intelligence. To establish the construct validity of the test, an attempt has been made to carry out a factor analysis of the test. The procedure and the conclusions are discussed in the next chapter.
Selected References


13. Long, Sandiford and Others, "The Validation of Test Items," University of Toronto.

15. Ross, C. C., "Measurement in To-day's Schools," ch. 4
   New York: Prentice Hall, Inc.

16. Thorndike, R. L., Stein, S., "An Education of the
    Attempts to Measure Social Intelligence,"
    The Psychological Bulletin, Vol. 34, No. 5,

17. Thorndike, R. L., "Personnel Selection," ch. 3-9
    New York: John Wiley & Sons, Inc.