CHAPTER 9

VALIDITY STUDIES
9.1. Concept of validity—

The concept of validity refers to the extent to which a measuring instrument measures what it is designed to measure i.e. how well it fulfills the function for which it is used. Different authors have discussed the concept of validity in different ways and have given different definitions. But the widely accepted basic definitions tend to follow the same general pattern. According to Garret (1967, pp. 364) "a test is valid when the performances which it measures correspond to the same performances as otherwise independently measured or objectively defined." This definition suggests that to determine the validity of a test, one must compare the performances of the test with some other independently measured performances. Gulliksen (1950, pp. 88) states "the validity of a test is the correlation of the test with some criterion." In the same way Cureton (1966, pp. 623) defined validity "in terms of the correlation between the actual test scores and the "true" criterion scores! The index of validity, thus, shows the extent to which a particular test measures what it intends to measure, when compared with an accepted criterion.

The proposed test has been designed to measure the adjustment problems of the adolescent girls in different areas and therefore it should be made sure whether it actually do so. In the field of personality testing, however,
the problem of establishing validity is more difficult than with other psychological tests. This is particularly due to the non availability of a satisfactory criterion. Intelligence tests are generally correlated with school marks, educational achievement test scores with other similar standards. But with personality tests, objective standards of scholastic success cannot be used and this makes the test constructor compelled to depend upon some other less satisfactory criteria. Case studies made by competent psychologists may be useful but a sufficient number of such cases are not generally available. Psychiatrists diagnosis also cannot provide much help as they are mostly concerned with abnormal cases. The test constructors of personality, therefore, often have to rely upon rating and observation by parents, teachers and other competent persons. The different methods of estimating validity and the criteria used for validating the present test are discussed below.

9.2. Methods of estimating validity—

There are several methods for determining the validity of a psychological test. But fundamentally, all methods for estimating validity are "concerned with the relationship between performance on the test and other independently observable facts about the behaviour characteristic under consideration" (Anastasi, '68, pp. 99). In the "Standards
for Educational and Psychological tests and manuals of American Psychological Association (1966), all the procedures for estimating validity have been classified into three distinct categories — criterion, content and construct related values (Cronbach, '71, pp. 444).

(1) **Criterion-related validity** — Criterion-related validity is also known as predictive validity. It "indicates the effectiveness of a test in predicting an individual's behaviour in specified situations" (Anastasi, '68, pp. 105). For determining test validity by this procedure performance on a test is checked against an external variable which is the criterion, i.e., a direct independent measure of that which the test is designed to predict. Thus the correlation between the raw test score and the raw criterion score is the estimate of validity. The evaluation of test validity by this method is primarily an empirical and statistical evaluation. Therefore this aspect of validity sometimes been called as empirical or statistical validity. But the most difficult problem which the test constructor often faces in this procedure is the selection of a suitable criterion. Thorndike and Hagen (1970, pp. 168) have mentioned four desired qualities in a criterion measure in order of importance. These are relevance, freedom from bias, reliability and availability. Validity coefficient depends to a great extent on the quality of the criterion employed. The test constructor, therefore, should be careful enough in selecting the
criterion employed for the test.

(ii) **Content validity** — Content validity means "the extent to which an examination is adequately samples the area of knowledge or skill with which it is concerned" (Pidgeon & Yates, '68, pp. 67). In this procedure the emphasis is placed on the test content, i.e., whether the test represents a fairly well defined universe or whether it covers a representative sample of the behaviour domain to be measured. Content validity is determined by evaluating the relevance of the test items individually and also as a whole. Every item should represent the knowledge or skill, which the test intends to measure and as a whole the items should constitute "a representative sample of the variable to be tested" (Freeman, '65, pp. 91). This type of validity is most appropriately applied in the selection of items in educational achievement tests and is most often estimated by expert judgement. This procedure is considered most satisfactory when the sampling of item is wide and judicious and adequate standardization groups are utilized. The test analyst examines carefully the content and objectives of the test for which it is designed and then prepares the outline based on the test itself indicating the content represented. After that the two outlines are compared and the discrepancies between them are noted. Thus the judgement concerning content validity is based on the extent of agreement between the test and the instructional plan.
(iii) **Construct validity**— The construct validity of a test is the extent to which a test tells something about the meaningful characteristic of the individual. It "is an analysis of the meaning of test scores in terms of psychological constructs" (Cronbach and Meehl, 1955, pp. 281-302). Neither criterion related nor content validity have as their fundamental goal the understanding of the trait that a test measures but construct validity focuses on this problem. It "depends upon the degree to which the items individually and collectively sample the range or class of activities or traits, as defined by the mental process or the personality trait being tested" (Freeman, '65, pp. 94). In the technical recommendations of American Psychological Association, construct validity has been explained as "The test user wishes to infer the degree to which the individual possess some hypothetical trait or quality (construct) presumed to be reflected in the test performance" (Womer, '68, pp. 53). Thus construct validity tries to answer two fundamental questions — what psychological construct being measured by the test and how well it does so? In construct validation there is no single criterion against which the test is to be validated — many criteria are required to confirm what the test does and does not measure. It is generally evaluated by the accumulation of evidence of the test, i.e., by examining whether its variables and scores do or donot correlate with the items included in the test, test scores remain stable under different condition, the test is fairly
homogeneous or any other data that may influence the test scores (Brown, '70, pp. 143). Thus, a clearer definition of the construct measured by the test emerges from evaluating and refining the evidence of the test.

9.3. Validating criteria used for the present test——

It has been mentioned earlier that establishment of validity in a personality test is somewhat more difficult than in other psychological tests. Still, every possible step was taken to make the Adjustment Inventory AB a properly valid one. About the content validity of the test, it has been explained in the earlier chapter (Chap. 4) that in the preliminary exploration, several girl students of the High Schools of Gauhati, for whom the test is designed were approached for frank statement of their personal problems in different areas, in the form of a short essay and problem papers. Moreover some experienced teachers and parents were also consulted in this connection. All the problems obtained through these sources were then pooled together in the form of items and were sorted and resorted to eliminate overlapping and to give a comprehensive coverage of problems faced by students. All the items were carefully edited in order to ensure freedom from ambiguity. From this one can say that the inventory is constructed with valid items, as it is mainly derived from what the students themselves reported as being their problems. Again in item analysis (Chap. 6), validity coefficients
determined for each item by biserial correlation method and only such items were included which yielded biserial correlation with both criteria (i) total score and (ii) total area score and having discrimination indices .20 or above .20.

Besides these, the following two criteria were used for estimating the validity of the present test — (a) Adjustment Inventory for school students developed by A.K.P. Sinha and R.P. Singh, published by National Psychological Corporation, Agra (1971). (b) Teachers' rating about the adjustment of their students.

(a) Correlation of the Adjustment Inventory AB with Sinha and Singh's Adjustment Inventory for school students—

In Hindi, A.K.P. Sinha and R.P. Singh have developed an Adjustment Inventory for school students for use with the Hindi knowing school students of India. The inventory intends to segregate well adjusted secondary school students (age group 14 to 18 years) from poorly adjusted students in three areas of adjustment: emotional, social and educational. This inventory has been translated into Assamese for the present purpose and administered to a representative sample of students for whom the test is ultimately designed. Scores obtained in this inventory are correlated with the scores obtained in the Adjustment Inventory AB. The following table shows the correlation coefficients of the test with Sinha and Singh's Inventory scores in different classes.
Table 9.1: Validity coefficients of the Adjustment Inventory AB estimated by correlating with Sinha and Singh's inventory scores.

<table>
<thead>
<tr>
<th>Classes</th>
<th>No. of students tested</th>
<th>Validity coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIII</td>
<td>82</td>
<td>.75</td>
</tr>
<tr>
<td>IX</td>
<td>70</td>
<td>.73</td>
</tr>
<tr>
<td>X</td>
<td>70</td>
<td>.75</td>
</tr>
</tbody>
</table>

(b) Correlation of the inventory with teachers’ rating—
The inventory has also been validated by correlating the inventory scores with teachers' ratings. A rating scale has been constructed for this purpose as shown in Appendix D No. ..... The teachers were asked to rate their students with whom they were more familiar with. They rated their students on a four point rating scale namely — very well adjusted, well adjusted, moderately adjusted, and ill adjusted in five different divisions — classroom adjustment with teachers, classroom adjustment with friends, outside classroom adjustment with teachers, outside classroom adjustment with friends and adjustment with people other than teachers and friends. The scale was scored as 1 for very well adjusted, 2 for well adjusted, 3 for moderately adjusted, 4 for ill adjusted and 5 for very ill adjusted. The total rating scores for each student was found out and correlated with the inventory scores. The results are shown in the table 9.2.
Table 9.2: Validity coefficients of the Adjustment Inventory AB estimated by correlating with teachers' rating scores.

<table>
<thead>
<tr>
<th>Classes</th>
<th>No. of students rated</th>
<th>Validity coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIII</td>
<td>40</td>
<td>.56</td>
</tr>
<tr>
<td>IX</td>
<td>40</td>
<td>.59</td>
</tr>
<tr>
<td>X</td>
<td>40</td>
<td>.58</td>
</tr>
</tbody>
</table>

The validity coefficients shown in the table 9.1 range from .73 to .75 and the table 9.2 range from .56 to .59 and may be considered adequate for the present test, since such coefficients in other comparable studies reported in the First Mental Measurement Handbook of India (1966) and Handbook of Psychological and Social Instruments (1974) are in conformity with the present test. Some of these are —

1. Vyaktitva Prakash Prashnavali, M.S.L. Saxena correlated with Asthana's inventory — — .80
2. Aligarh Adjustment Inventory, A Jamil Quadri and Guidance unit, with ratings based on clinical interview — — .73
3. Kishor Abhiyojan Prasnavali, Jai Prakash, with Saxena's inventory — — .60
4. Adjustment Inventory, S. Bhattacharya, M.M. Shah and J.C. Parikh, with expert opinion ratings .63
(5) Adjustment Inventory for college students,
A.K.P. Sinha and R.P. Singh, with hostel
superintendent ratings --- --- .58
(6) Personality Questionnaire, J.N. Lal,
with teachers ratings --- --- .68
(7) Personality Inventory, Bureau of Educational
and Psychological research, David Hare College,
Calcutta, with teachers rating --- --- .29 to .70
(8) Indirect Academic Adjustment Inventory,
S. Narayan Rao, with teachers rating --- .67
with Borov's College Inventory of Academic
Adjustment --- .41
SUMMARY

Chapter 9

In the first section of this chapter the meaning of validity is discussed along with the difficulties of establishing validity in a personality test. The second section deals with the different methods of establishing validity and in the third section, the validating criteria used for establishing the validity coefficients of the present test are stated.