Description of Tools
CHAPTER - IV

DESCRIPTION OF TOOLS

In the preceding chapters, the introduction to different variables under study, the review of the related literature, objectives and hypotheses were discussed. The present chapter deals with the description of tools.

For the present investigation following tools were used:

4.1 Cattell’s Culture Fair Test of Intelligence
4.2 Group Embedded Figure Test by Witkin.
4.3 Development of Summative Criterion Test (Developed by the investigator)
4.4 Development of Achievement Test (Developed by the Investigator)
4.5 Development of Formative Criterion Tests (Developed by the Investigator)
4.6 Development of instructional material (Developed according to Inquiry Training Model)
4.7 Development of instructional material (Developed according to Mastery Learning Model)
4.8 Self concept Test by Dr. G.P. Sherry, Dr. R.P. Verma and Dr. P.K. Goswami

4.1 CATTELL’S CULTURE FAIR TEST OF INTELLIGENCE

Cattell’s Culture Fair Test of mental abilities was used to measure intelligence of pupils. The Culture Fair Intelligence Test measures individual’s intelligence in a manner designed to reduce, as much as possible, the influence of verbal fluency, cultural climate, and educational level. The tests which may be administered individually or in a group are non verbal and require only that examinees be able to perceive relationships in shapes and figures. Each scale contains four subsets, involving different perceptual tasks, so that the composite intelligence measure avoids spurious reliance on a single skill.

Culture Fair Tests are valuable instruments for all uses to which an
intelligence test is rationally applied. The main aims in the design and construction of the tests were:

1) to create a psychometrically sound instrument upon a comprehensive theoretical foundation, with maximum possible validity and reliability.

2) to provide high administrative and scoring convenience and maximum economy of testing time.

3) discovering whether the achievement of each individual is what would be expected from his intelligence, thus facilitating identification of emotional or learning problems where they exist.

4) making a more reliable and informed decision whether a particular child may require special educational facilities or a programme of cultural enrichment.

5) selecting students for accelerated educational programmes within an age-group grade.

6) increasing the effectiveness of vocational guidance decisions, both for students and adults.

In all such applications, the special advantage of the Culture Fair Test is that they make a cleaner separation of natural ability from specific learning and thus permit better analysis and prediction of the individuals ultimate potentialities.

Contents:

To measure the intelligence of the students scale-2, Form A of culture Fair Test was used. Scale-2 consisted of four subsets.

In the first subset, the individual is presented with an incomplete progressive series. Students' task is to select from among the choices provided, the answer which best continues the series. Second subset is
related to classification. The individual is presented with five figures. He must select one which is different from the other four. In the third sub-test related to matrices, the individual is asked to correctly complete the design matrix presented at the left of each row. The fourth subset, conditions (or topology) requires the individual to select, from the five choices provided, the one which duplicates the conditions given in the left box. Before each subset, examples are given, so that the task requirements are clear to examine. A copy of the test is given in Appendix IX.

ITEMS AND TIME ALLOTTED TO EACH SUBSET IN SCALE- 2

<table>
<thead>
<tr>
<th>SCALE 2, FORM A</th>
<th>No. of Item</th>
<th>Time Allotted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1 Series</td>
<td>12</td>
<td>3 Mins</td>
</tr>
<tr>
<td>Test 2 Classification</td>
<td>14</td>
<td>4 Mins</td>
</tr>
<tr>
<td>Test 3 Matrices</td>
<td>12</td>
<td>3 Mins.</td>
</tr>
<tr>
<td>Test 4 Condition</td>
<td>8</td>
<td>2½ Mins.</td>
</tr>
<tr>
<td>Total</td>
<td>46 Items</td>
<td>12½ Mins.</td>
</tr>
</tbody>
</table>

Reliability

The reliability of the test (Scale-2, Form A) by Spearman Brown formula is 0.79 and K.R. formula 21 is 0.81. Direct concept validity of the test is 0.85.

4.2 GROUP EMBEDDED FIGURES TEST BY WITKIN.

(WITKIN, OLTMAN, RASKIN AND KARP, 1971)

Group Embedded Figures Test (GEFT), an adaptation of the original individually administered Embedded Figures Test (EFT) was
preferred as a suitable instrument to measure the cognitive style
dimension of field-independence/dependence of the sample primarily
because it makes group testing possible and scores for a large number
of individuals can be conveniently obtained in a single test session of 20
minutes. It has 18 complex figures, 17 of which are taken from EFT.
Moreover, the value of Pearson's product moment coefficient of
correlation between scores obtained on two administrations of GEFT (as
reported by Vasesi, 1985) on a representative sample of the school
students (Class IXth and Xth, N = 50) is 0.86. It indicates a high
reliability of the test.

GEFT required the subjects to locate a simple visual figure embedded
with in a more complex one. Besides the seven simple forms (A, B, C, D, E,
F, G) that have to be located, the test has three sections, first section
comprising of a seven-item practice set which served the purpose of
providing practice to the subjects and is not to be scored, second and third
sections are comprised of nine difficult figures which are arranged in
ascending order of difficulty within each section. Thus, the test has a scored
set of 18 items administered in two equal parts and for which, subjects are
allowed a time limit of 5 minutes each, where as for the practice set only 2
minutes are to be allowed (Manual of GEFT, 1971). The total number of
simple form correctly traced in second and third sections combined is the
individual's scores, since the items in practice set are not scored but merely
scanned to ensure that the instructions have been understood properly by
the subjects. Omitted items are scored as incorrect. In order to receive
credit for an item, all lines of the sample forms must be traced. All incorrect
lines must be crossed. No extra lines are to be added.

Since GEFT is speed test, an appropriate method of estimating
reliability is the correlation coefficient between parallel forms with identical
time limit. Correlation between the second and third sections were computed and calculated by Sperman Brown prophecy formula, producing a reliability of 0.82 for both males (N=80) and females (N=97) from an eastern liberal arts college (as reported in GEFT Manual, 1971). The test along with scoring key, is appended in the Appendix No X.

4.3 DEVELOPMENT OF THE CRITERION TEST

The Criterion reference test is a sample of items yielding information that is interpretable directly with respect to both a well defined domain of tasks and to specified performance standards (Glaser and Nitko, 1971). This definition reflects three characteristics of criterion referenced measurement.

(i) Definition of a well specified content domain (Backel, 1974; Hambleton and Novick 1973; Millan 1972)
(ii) Delineation of valid performance criteria. (Hambleton, 1980) and
(iii) Development of procedures for generating appropriate samples of test (Goldstein, 1982; Popham 1980)

Popham (1975) maintained that the most important aspect of criterion referenced measurement is that a well-explicated domain of behaviour be delineated and an individual's performance in relation to his behaviour domain be ascertained. All these components stress the educatiometric and psychometric of criterion referenced tests.

Following steps were followed in the construction of criterion test (Popham, 1975)

Step I : Domain definitions
Step II : Generating Items
Step III : Improving Items
Step IV : Reliability and Validity
Step I : Domain Definitions

This is the most difficult and important step in the construction of a criterion test. Here the limits of behaviour that the test items would measure and to which all individual performance be referenced were determined. The domain definitions were kept brief and at the same time sufficiently circumscribing the class of behaviors under consideration so that they, in fact measure the behaviour described in that domain. Next aspect regarding domain definition was selection among competing domain alternatives, which was affected by transferability within domain alternatives and transferability outside the domain.

Regarding generation of domain definitions, two guiding factors were given due consideration viz. (1) economy of description (2) ambiguity reduction for developing stimulus elements. While working on a domain, a middle position was taken between the two extremes of:

(1) Sufficient detail for complete stimulus homogeneity of resulting test items and

(2) Economy of resource investment

Although a domain did not limit all possible test items, it markedly reduced the ambiguity associated with the class of learner behaviours under consideration. After identification and definition of learning outcomes, from the selected chapters of mathematics text book of class V, the content was divided into five units.
Table 4.1
CONTENT SPECIFICATIONS

<table>
<thead>
<tr>
<th>Units</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit-I</td>
<td>Factors and multiples H.C.F &amp; L.C.M, Multiplication and division of fractions</td>
</tr>
<tr>
<td>Unit-II</td>
<td>Decimals, Multiplication and Division of decimals.</td>
</tr>
<tr>
<td>Unit-III</td>
<td>Temperature, Averages, Ratio, Speed, Distance, and Time</td>
</tr>
<tr>
<td>Unit-IV</td>
<td>Lines, Angles, and Circles, Triangles</td>
</tr>
<tr>
<td>Unit-V</td>
<td>Quadrilateral, Areas and Volumes</td>
</tr>
</tbody>
</table>

The instructional objectives developed on this content have been placed in Appendix-I. In order to make sure that the criterion test provides a representative sample of the desired behaviour the specifications were developed. Thus for testing 90 instructional objectives, 120 items of multiple choice were planned.

**Step II : Generating Items**

A number of items may be constructed for any given objective, even a highly specific objective could have a potential item pool of well over several thousand items (Hively 1970, 1973; Bormulh 1970). In terms of feasibility, a survey of current measures revealed that the usual practice is to use about 3-5 items per objective. This practice, however does not have any sound foundation in psychometric theory or technology (Klein and Kosecoff, 1976).

As opined by Gronlund (1977) each item was started as a multiple choice items and was converted to other types of items when the learning outcomes of the subject matter made it desirable to do so. In writing the items proper care was taken regarding the coverage of the content structure of items type and language. The item of each content points were exhaustively written so as to test every aspect thoroughly. After generating items, each type of items were grouped together and appropriate directions...
for the students to attempt these items were written. The scoring key was also developed. Thus the first draft of the criterion test was developed.

**TABLE 4.2**
BLUE PRINT OF FIRST DRAFT OF CRITERION TEST

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Content</th>
<th>Knowledge Category (O₁)</th>
<th>Comprehension Category (O₂)</th>
<th>Total No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unit - I</td>
<td>19</td>
<td>20</td>
<td>39</td>
</tr>
<tr>
<td>2</td>
<td>Unit - II</td>
<td>10</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>Unit - III</td>
<td>12</td>
<td>18</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>Unit - IV</td>
<td>33</td>
<td>9</td>
<td>42</td>
</tr>
<tr>
<td>5</td>
<td>Unit-V</td>
<td>14</td>
<td>9</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>88</td>
<td>62</td>
<td>150</td>
</tr>
</tbody>
</table>

**Step III : Improving Items**

In order to ensure that the entire class of learner behaviours had been circumscribed, a try-out was conducted.

The preliminary draft was administered to a sample of 350 students of class V who had already studied the content. No time limit was imposed. The answer sheets were scored with the help of scoring key prepared for the test.

For each correct response, 1 mark was given and for wrong response no mark was given. The total scores thus obtained by the students were arranged in descending order which were utilized for item analysis of the test items.

On the basis of total scores obtained, the students were divided into 3 groups, the higher group, the average group and lower group. The higher and lower group comprised 27% each of total number of students. 27% was taken because it was optimum. Turman kelley (1939) showed that when
extreme groups each consisting of 27% of the groups are used, the ratio of
the difference in average ability of the group to the standard error of their
difference is maximum. The number of correct responses for each item in
the higher and lower groups were used for finding out the difficulty value of
each item (D.V.). The difficulty value of each item was calculated using the
formula.

\[ D.V. = \frac{R_u + R_l}{N} \]

Where  
\( R_u = \) Number of right responses in upper group
\( R_l = \) Number of right responses in lower group
\( N = \) Total No. of students in both the groups

<table>
<thead>
<tr>
<th>Difficulty Range</th>
<th>Frequency</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>91 - 100</td>
<td>23</td>
<td>Very Easy</td>
</tr>
<tr>
<td>81 - 90</td>
<td>8</td>
<td>Easy Items</td>
</tr>
<tr>
<td>71 - 80</td>
<td>8</td>
<td>Easy Items</td>
</tr>
<tr>
<td>61 - 70</td>
<td>24</td>
<td>Easy Items</td>
</tr>
<tr>
<td>51 - 60</td>
<td>12</td>
<td>Moderate</td>
</tr>
<tr>
<td>41 - 50</td>
<td>23</td>
<td>Moderate</td>
</tr>
<tr>
<td>31 - 40</td>
<td>34</td>
<td>Difficult Item</td>
</tr>
<tr>
<td>21 - 30</td>
<td>11</td>
<td>Difficult Item</td>
</tr>
<tr>
<td>11 - 20</td>
<td>4</td>
<td>Difficult Item</td>
</tr>
<tr>
<td>1 - 10</td>
<td>3</td>
<td>Very Difficult</td>
</tr>
</tbody>
</table>

TABLE 4.3
DISTRIBUTION OF ITEMS ON THE BASIS OF
ITEM DIFFICULTY INDEX.

Some items were rejected on the basis of difficulty values with
respect to the types of learning and categories of objectives, the final draft
of the test containing 120 item was prepared. The difficulty values of these
120 items have been placed in the item analysis chart in Appendix II.
TABLE 4.4
BLUE PRINT OF FINAL DRAFT OF CRITERION TEST

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Content</th>
<th>Knowledge Category (O₁)</th>
<th>Comprehension Category (O₂)</th>
<th>Total No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Multiple Choice</td>
<td>Multiple Choice</td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>Unit - I</td>
<td>14</td>
<td>16</td>
<td>30</td>
</tr>
<tr>
<td>2.</td>
<td>Unit - II</td>
<td>8</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>3.</td>
<td>Unit - III</td>
<td>10</td>
<td>16</td>
<td>26</td>
</tr>
<tr>
<td>4.</td>
<td>Unit - IV</td>
<td>27</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>5.</td>
<td>Unit - V</td>
<td>11</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70</td>
<td>50</td>
<td>120</td>
</tr>
</tbody>
</table>

TABLE 4.5
DISTRIBUTION OF ITEMS ON THE BASIS OF ITEM DIFFICULTY INDEX (FINAL DRAFT)

<table>
<thead>
<tr>
<th>Difficulty Range</th>
<th>Frequency</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>91 - 100</td>
<td>19</td>
<td>Very Easy</td>
</tr>
<tr>
<td>81 - 90</td>
<td>4</td>
<td>Easy Items</td>
</tr>
<tr>
<td>71 - 80</td>
<td>7</td>
<td>Easy Items</td>
</tr>
<tr>
<td>61 - 70</td>
<td>19</td>
<td>Easy Items</td>
</tr>
<tr>
<td>51 - 60</td>
<td>8</td>
<td>Moderate</td>
</tr>
<tr>
<td>41 - 50</td>
<td>22</td>
<td>Moderate</td>
</tr>
<tr>
<td>31 - 40</td>
<td>28</td>
<td>Difficult Item</td>
</tr>
<tr>
<td>21 - 30</td>
<td>10</td>
<td>Difficult Item</td>
</tr>
<tr>
<td>11 - 20</td>
<td>1</td>
<td>Difficult Item</td>
</tr>
<tr>
<td>1 - 10</td>
<td>2</td>
<td>Very Difficult</td>
</tr>
</tbody>
</table>

Preparation of Directions of the Test

Simple and concise directions, including purpose of the test and procedure of recording the answer was developed by the investigator. The test comprised of multiple choice items. Therefore general directions were
provided for the entire test. Scoring key was developed. A copy of the test has been enclosed in the Appendix III.

Reliability and Validity

The idea of reliability is as crucial for criterion test as for all other tests yet for criterion test, there are some important cautions that need to be taken care of and a somewhat different way of conceptualizing certain reliability procedures.

Popham (1975) suggested that when marked range restrictions are present it may be necessary to employ less sophisticated but more meaningful, reliability estimates. Less fancy methods of calculating consistency might be used, for instance, rather than using correlational method, simple computation of the percentage of student scores, that were on two different testing occasions, different by 0-5 percent, 6-10 percent etc. The criterion test was administered to a group of 100 students on two different occasions. The percentage of students corresponding to different scores have been given in Table 4.6.

<table>
<thead>
<tr>
<th>Percentage of Scores</th>
<th>Difference in Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 5</td>
<td>66</td>
</tr>
<tr>
<td>6 - 10</td>
<td>34</td>
</tr>
<tr>
<td>11 - 15</td>
<td>-</td>
</tr>
<tr>
<td>16 - 20</td>
<td>-</td>
</tr>
<tr>
<td>21 - 25</td>
<td>-</td>
</tr>
<tr>
<td>26 - 30</td>
<td>-</td>
</tr>
<tr>
<td>31 - 35</td>
<td>-</td>
</tr>
<tr>
<td>36 and Above</td>
<td>100</td>
</tr>
</tbody>
</table>
The table 4.6 shows that the difference between the scores of all students fall in first two categories. It depicts that the students show a marked level of consistency across the score levels. So this test may be considered reliable for measuring performance of students.

This test was validated against the criterion and content validity. It was the only kind of validation which was feasible under the circumstances.

Gronlund (1977) also opined that the type of validity of greatest importance for criterion references test is content validity. Hence the content validity of the test was established by relating the task to instructional objectives.

The table 4.4 indicates that there are 120 items in the criterion test and these measure 90 instructional objectives. The final draft of the criterion test has been given in the Appendix-III.

4.4 DEVELOPMENT OF THE ACHIEVEMENT TEST

"Achievement Testing" refers to the assessment of the outcomes of formal instruction in cognitive domain (Dwyer, 1982). It can also be thought of as a sample of indicator of a student's knowledge taken at a particular point of time (Ebel, 1979) or achievement test may mean a sample of behavior that provides opportunity for comparison with performance standard, as in criterion-referenced testing (Gronlund, 1973; Hambleton et al, 1978, Popham, 1978). It aids both the teachers and students in assessing learning readiness, monitoring learning progress, diagnosing learning difficulties and evaluating learning outcomes (Gronlund, 1977). Strictly speaking, achievement and criterion test differ only with respect to interpretation of results, thus both types of interpretations could be applied to same test (Gronlund, 1977). An achievement test is distinguishable from a criterion test on the basis of purpose of the test besides its rigorous
In the present study, achievement test was developed from those criterion items that exhibited sufficient response variance to measure the performance of the students with two different instructional treatments.

First Try Out

The test was administered to 350 students of class V. The discriminating power was computed for each item after forming top 27 percent and bottom 27 percent groups from the total subjects as suggested by Kelly (1939).

TABLE 4.7
BLUE PRINT OF FIRST DRAFT OF ACHIEVEMENT TEST.

<table>
<thead>
<tr>
<th>Type content</th>
<th>Knowledge category ((O_1)) Multiple choice</th>
<th>Comprehension category ((O_2)) Multiple choice</th>
<th>Total No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit-I</td>
<td>14</td>
<td>16</td>
<td>30</td>
</tr>
<tr>
<td>Unit-II</td>
<td>8</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Unit-III</td>
<td>10</td>
<td>16</td>
<td>26</td>
</tr>
<tr>
<td>Unit-IV</td>
<td>27</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>Unit-V</td>
<td>11</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>50</td>
<td>120</td>
</tr>
</tbody>
</table>

TABLE 4.8
DISTRIBUTION OF DISCRIMINATING POWER (D.P) OF ITEMS OF FIRST DRAFT OF ACHIEVEMENT TEST

<table>
<thead>
<tr>
<th>Discriminating Range</th>
<th>Frequency</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>.40 and above</td>
<td>86</td>
<td>Very good Items</td>
</tr>
<tr>
<td>Between .30 and .39</td>
<td>8</td>
<td>Reasonably good</td>
</tr>
<tr>
<td>Between .20 and .29</td>
<td>4</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>Below .19</td>
<td>22</td>
<td>Poor Items</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td></td>
</tr>
</tbody>
</table>
Second Try Out

The revised version of the achievement test was administered on another group of students. Again discriminating power for 96 items was computed. The distribution of the discriminating power of items can be seen in table 4.9

**TABLE 4.9**

**DISTRIBUTION OF DISCRIMINATING POWER (D.P) OF ITEMS OF FINAL DRAFT OF ACHIEVEMENT TEST**

<table>
<thead>
<tr>
<th>Discriminating Power</th>
<th>Frequency</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>.40 and above</td>
<td>86</td>
<td>Very good Items</td>
</tr>
<tr>
<td>Between .30 and .39</td>
<td>8</td>
<td>Reasonably good</td>
</tr>
<tr>
<td>Between .20 and .29</td>
<td>2</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>Below .19</td>
<td>-</td>
<td>Poor items</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td></td>
</tr>
</tbody>
</table>

In the light of above results, out of 120 items, 24 items below the discriminating power of .2 were dropped and 2 items were improved and 96 items were retained. This led to the preparation of final draft of the Achievement Test. This draft of achievement test comprised of 96 items. This table of specification for final draft of the Achievement test has been presented in the table 4.10

**TABLE 4.10**

**BLUE PRINT OF FINAL DRAFT OF ACHIEVEMENT TEST**

<table>
<thead>
<tr>
<th>Type content</th>
<th>Knowledge category (O1)Multiple choice</th>
<th>Comprehension category(O2) Multiple choice</th>
<th>Total No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit-I</td>
<td>6</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>Unit-II</td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Unit-III</td>
<td>10</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>Unit-IV</td>
<td>17</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>Unit-V</td>
<td>9</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>49</td>
<td>47</td>
<td>96</td>
</tr>
</tbody>
</table>

154
Reliability and validity of Achievement Test

Reliability concerns the extent to which measurements are repeatable, i.e. when different persons make the measurements on different occasions with supposedly alternative instruments for measuring the same things (Nunnally, 1982). In other words, measurements are intended to be stable over a variety of conditions where essentially the same results should be obtained. The Achievement test was administered to 100 students of different occasion with time interval of 30 days. Product moment correlation was computed between the two sets of scores & reliability of the test was found to be .98.

Validity

"Content Validity" is determined by showing that the behaviors demonstrated in testing constitutes a representative sample of behaviours to be exhibited in a desired performance domain. The domain usually involves learned knowledge and skills. Thus content validity is commonly use in Achievement Tests (Wolf, 1982). The content validity is determined by comparing the items in a test with the content and objectives of a particular domain. A copy of test has been placed in Appendix IV.

4.5 DEVELOPMENT OF THE FORMATIVE CRITERION TESTS

The formative criterion tests were developed to measure the performance of the students after teaching each of the five units. The tests were developed through the following steps:

Step 1 Planning the test,
Step 2 Preparing the first draft of the tests,
Step 3 Try Out
Step 4 Preparing the final draft of the unit tests.
Step 5 Validity.
Step 1 : Planning the Tests

Planning is very important in the process of development of any test. It covers selection of the topic, developing of instructional objectives and determination of items used. The investigator selected five units from syllabus of Algebra of class V. The following units were selected.

TABLE 4.11
CONTENT SPECIFICATION FOR FORMATIVE CRITERION TESTS

<table>
<thead>
<tr>
<th>Units</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit-I</td>
<td>Factors and multiples, H.C.F &amp; L.C.M, Multiplication and division of fractions</td>
</tr>
<tr>
<td>Unit-II</td>
<td>Decimals, Multiplication and Division of decimals.</td>
</tr>
<tr>
<td>Unit-III</td>
<td>Temperature, Averages, Ratio, Speed, Distance, and Time</td>
</tr>
<tr>
<td>Unit-IV</td>
<td>Lines, Angles, Circles, Triangles</td>
</tr>
<tr>
<td>Unit-V</td>
<td>Quadrilateral, Areas and Volumes</td>
</tr>
</tbody>
</table>

Step-II : Preparing the first draft of the tests

The first draft of the tests was developed keeping in mind the instructional objectives. After consulting the mathematics teacher, the unit tests were developed. Different types of items viz., multiple choice, matching, fill in the blanks and short answer type were developed.

Step-III : Try Out

The five unit test were tried out on 50 students of class VI. There was no time limit. After the test, the students were asked to point out any items, which were not meaningful, or comprehensible.

Step-IV : Preparing the final Draft of the Unit Tests

Comments and suggestions of students were incorporated and final draft of the tests was developed. Copies of the unit tests have been placed in Appendix-V.
Step-V : Validity

"Content validity" was determined by comparing the items in a test with the content and objectives of a particular domain to see how well they match, as it is essential for a valid test to reflect the content of a particular domain. The five unit tests represented a fairly well-defined universe of content, so they were considered valid.

4.6 DEVELOPMENT OF THE INSTRUCTIONAL MATERIAL (ACCORDING TO INQUIRY TRAINING MODEL)

The following steps were followed in developing the instructional material based on Inquiry Training Model.

Step-I Selection of the content

The content was selected from the fifth class syllabus of mathematics of C.B.S.E., as given in table 4.1 of content specification.

Step-II Division of the content

Content was divided into 45 sub-units, so that each unit may be taught in 35 minutes duration. Care was taken to place each sub-unit in a logical and psychological sequence.

Step-III Writing of Instructional objectives

Instructional objectives were written for each sub-unit. They have been placed in Appendix-I.

Step-IV Developing the lesson plans according to inquiry Training model

The instructional material was developed to be administered as instructional treatment to V class students. During planning, capabilities, interests and needs of the students were kept in mind.

This enabled the assessment of student's competency on instructional
objectives prior to the instructional programme implementation. The lesson plans were developed with the help of guidelines from the book, "Multiple Models of Teaching for Educators." By L.C Singh, discussion with experts in this field and review of some sample lesson plans in previous research work. According to L.C Singh (1995) the teaching models have direct relevance to the designing of instructional materials, as each model aims at specific instructional objectives, the planning of the subject matter accordingly, methods of presentation and evaluation, content and instructional objectives are as important as the learner in the educational process because they interact powerfully with the teaching method. Instructional materials have a powerful effect on the learning environment. Teaching method transforms content and material through emphasis and its process, and both vary according to the model used. Models of teaching are thus an important guide for the developers of instructional material.

The emphasis of the Inquiry Training Model is upon making the learners aware of the inquiry process through structuring and re-structuring of questions and sequencing them appropriately. It is more of a process oriented model, with relatively less emphasis on content. The lessons were developed keeping the significance of the various phases of Inquiry Training Model.
<table>
<thead>
<tr>
<th>Phase</th>
<th>Major Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>First - Presentation of the problem</td>
<td>The teacher presents the pre-planned discrepant event and explains the inquiry procedure, objectives and the procedure of yes/no questions.</td>
</tr>
</tbody>
</table>

**Inquiry session**

<table>
<thead>
<tr>
<th>Sub-phase one</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Second - Data gathering and hypothesizing</strong></td>
<td>The students inquire about the nature and identity of the objects, events properties and conditions related to the problem (through the type of questions) here the teacher guides the students to frame such questions which can be answered in yes/no whenever the students drift from this rule.</td>
</tr>
<tr>
<td><strong>Sub-phase two</strong></td>
<td>The students ask questions in order to verify the various hypotheses they formulate. These hypotheses are relationships among the variables identified in the first sub-phase.</td>
</tr>
</tbody>
</table>

| Third - Theorizing, i.e. organizing data and formulating explanation | The teacher asks the students to organize the data that they have gathered in the second phase and to give the most appropriate explanation which fits the data. |

| Fourth - Analysis of Inquiry Process | The teacher asks the students to analyze their pattern of inquiry. They may discuss as to which questions were most effective and which were not, or the type of information which they could have asked for. Here the teacher can guide the students by asking them, if they could ask more direct and relevant questions so as to arrive at the solution of the problem. Similarly, which questions were unnecessary and could have been avoided. |

The class room environment in this model is of a cooperative and participatory nature. It may be initiated in a structured manner with a greater control by the teacher. This is the beginning stage where the teacher presents the problem and guides the nature of questions. Later, as the teacher encourages the students to initiate inquiry and the students learn the principles of inquiry, the learning system becomes more open.
Step-V Try out of lesson

The lessons were tried out on a group of 25 students to obtain their response regarding the effectiveness of the lessons.

Step-VI Validation of the lesson

Responses of the students to the lessons and scores on post test indicated that they were instructionally sound. Changes were made when needed with respect to sequence, content, presentation and clarity in language. The suggestions of the experts of Punjab University, M.D. University and Kurukshetra University were incorporated and the lessons were again reviewed. A sample lesson developed according to Inquiry Training Model has been placed in Appendix -VI.

4.7 DEVELOPMENT OF THE INSTRUCTIONAL MATERIAL (ACCORDING TO MASTERY LEARNING MODEL)

Bloom’s mastery learning strategy is a group based and teacher-paced mastery learning strategy. So role of the teacher in this strategy is, sequencing the learning units, designing the instructional material and other alternate material, and presenting it in such a form that the participation of students is maximum (Block, 1974). Learning for mastery is probably a novel experience to the students. Hence, time should be reserved for orientation towards mastery learning strategy before its onset.

The orientation period focuses on:

i) What the students are expected to learn;
ii) the level of learning they are expected to learn;
iii) how the students are expected to learn;
iv) on what basis the evaluation will be done.

35 lessons were developed by the investigator, that consisted of a variety of instructional events according to the nature of desired learning on
the context specified in table. The student teacher participation was given
due emphasis while organizing events of lessons. An appropriate order of
events in which they were employed and the function served by the events
were as follows:

♦ Gaining attention
♦ Informing the objectives to the learners
♦ Stimulating recall of prerequisite learned capabilities
♦ Presenting the stimulus material
♦ Providing 'Learning guidance'.
♦ Attaining the performance
♦ Providing feedback
♦ Assessing the performance

Gaining Attention

In teaching lessons, various events were employed to gain the
learners' attention. The use of stimulus change and appeal to the learners
interest were the two frequently used methods of gaining attention and one
or more options of communications were designed to gain attention for a
single lesson.

Informing the objectives to the learners

Instructional objectives which may not be initially obvious to students
were informed in a simpler language to the students.

Stimulating recall of prerequisite learned capabilities

Component ideas constituting new learning must be mastered before
hand to make new learning a success. In organizing instructional events, for
a new learning, previously acquired capabilities were stimulated by recalling
them just before the new learning took place. The recall of previously
learned capabilities were stimulated by asking questions or a recognition of diagram, examples etc.

**Presenting the stimulus material**

The stimuli to be communicated to learner were those involved in the performance which reflected learning. In these lessons, the stimuli were presented orally and with the help of diagrams and examples on the chalkboard. Explanation, definitions were usually given orally whereas all the diagrams were presented on the blackboard.

Gagne and Briggs (1974) advocated that stimuli presentation for the learning of concepts and rules required the use of a variety of examples. In learning both concepts and rules, one may proceed either inductively or deductively. Keeping this in view, a variety of examples, diagrams were used in presenting stimuli for learning concepts, rules and problem solving situations of the selected content of the lessons.

**Providing learning Guidance**

To facilitate learning guidance was provided to the learners. According to Wittrock (1966), the amount of learning guidance is the number of questions and the degree to which they provide direct or indirect prompts, will obviously vary with the kind of capability being learned. Learning guidance was provided to stimulate learners' thought in the direction of terminal behaviour. This guidance was provided in the form of questions which in turn provided direct or indirect prompts leading to the desired knowledge combining subordinate knowledge. Learning guidance was presented in parallel and alternate forms of the actual learning material. The amount of questioning or hinting varied from lesson to lesson and it also varied from learner to learner.
**Attaining the performance**

After providing sufficient amount of learning guidance, the learners were carried to the point where the actual internal event of learning had taken place. The teacher constructs brief, ungraded, student scored, diagnostic-progress tests called 'formative' evaluation instruments for all learning units (Bloom, Hastings and Madaus, 1971). Bloom et al., (1971) found the term 'formative evaluation' helpful for instruction and learning outcomes as it helps in pacing student learning and seeks to identify learning weaknesses prior to the completion of instruction in a course segment. Formative evaluation is diagnostic in nature and therefore does not grade students, it refers to achievement or non-achievement of mastery level performance. It also pin points what the students has not learnt and what else he needs to learn in the light of specified objectives. The following suggestions given by Bloom et.al., (1971) and Gagne (1965) were duly considered in developing the formative tests and its parallel forms.

i) Analyzing the instructional material.

ii) Mapping the hierarchical structure of the learning tasks and

iii) Administrating achievement test after completing study material.

At this stage the students were asked to produce these answers according to the material presented earlier.

**Providing feedback**

The feedback concerning the corrections and degree of corrections of the learners performance was provided. Formative evaluation instrument provides detailed information about student' grasp of each unit objective. The teacher prepares a set of alternative learning materials or instructional correctives keyed to these formative tests. These correctives teach the
unit's objectives but in ways different from the teachers group based presentation (Block, 1974).

**Assessing the performance**

Teacher certifies the progress of students that is whether they are achieving as desired and requires those students who are not, to use the instructional correctives to complete their learning of unmastered objectives. The student is then given responsibility to use these correctives as necessary to complete his learning of objectives of the first unit before the teacher begins the group based instruction on the second unit (Kersh, 1971, Kim, 1971).

In organizing instructional events for teaching according to Bloom's mastery learning model, the above mentioned events were incorporated in all the lesson plans. These events did not individually occur exactly in the same order but with a little bit of variation here and there.

**Description of lessons**

Format of lessons have been written under same heads and follow a similar sequence Topic, objectives of the lesson, content sequence and Media, and instructional program. The main theme of content unit was taken as the topic of the lesson. Instructional objectives pertaining to a unit were specified and made part of each lesson. Objectives were orally presented to the students to give them guidelines or directions.

Instructional events were organized by linking content, events of the instruction and evaluation. Media used for teaching was chalkboard. Formative tests or unit for criterion test were prepared for each unit.

**Tryout of the lesson**

The lessons were tried out on a group of 25 students to obtain their responses regarding the effectiveness of the lessons.
Validation of the lesson

Responses of students to the lessons and scores on post test indicated that they were instructionally sound. Changes were made when needed with respect to sequence, content, presentation and clarity in language. The suggestions of the experts of M.D. University, Panjab University and Kurukshetra University were incorporated and lessons were again reviewed. A sample lesson developed according to Mastery Learning Model has been placed in Appendix -VII.

4.8 SBP- A TEST OF SELF-CONCEPT

The test of self-concept developed by Sherry, Verma and Goswami was used to measure the self-concept of pupils before and after the experimental treatment. The main objective of the test is to measure the self-concept of the students. Self-concept is the most important single attribute and key to understanding the behaviour of an individual. The importance of role of self-concept as a determinant of human behaviour and its acceptance as a concise measure of personality is increasingly being realized. Mostly, present day Educators and Psychologists consider that an individual’s self-concept is a critical facet of his personality. Adjustment, academic achievement and general behaviour are among the developmental features of an individual, which are subject to the influence of his self-concept. Consequently, many researchers are concerned with determining means for improving and channelizing the self-concept of the individual in desirable directions. For all this measurement of self-concept is the starting point.

Self-concept means – "those perceptions, beliefs, attitudes and feelings which the individual views as parts or characteristics of himself." It is his own conception of his health and physique, intellectual abilities,
academic status, behaviour, temperamental qualities, mental health, emotional tendencies and socio-economic status.

Swatvabodh Parikashan, is a forty-eight item test, yielding scores in eight different dimensions of the self-concept. Thus, the present test provides eight separate measures of self-concept. These dimensions are:

- Health and Physique
- Temperamental Qualities
- Academic status
- Intellectual abilities
- Habits and Behaviour
- Emotional Tendencies
- Mental Health
- Socio-Economic Status

The statements of the test are simple and declarative about self-seeking responses in 'Yes' or 'No'. Responses were obtained on an answer sheet. A high score on this test indicates a bright self-concept. While a low score shows a poor self-concept.

Booklets containing test items and answer-sheets were distributed to the sample students. An attempt was made to follow all the instructions laid down by Sherry, Verma and Goswami for conducting the test. Although no time limit was set, the pupils were advised to complete the test as quickly as possible. The pupils generally took 20 minutes on the average to complete the test. The scoring was done with the help of scoring stencil.

The test-retest reliability of the test is 0.733 and rational equivalent reliability using K.R. formula-21 0.761. This test is reported to have reasonable content validity. The test is appended in Appendix No XI.