CHAPTER III
THE PROCEDURE

3.1 Introduction

"Like the tools in a carpenter's box, each research tool is appropriate in a given situation to accomplish a particular purpose". The selection of suitable instrument and their proper administration are of vital importance for useful research.

A detailed description of the procedure adopted by the investigator to study the classroom behaviour of teachers and creativity of students, is presented in this chapter. A discussion on the choice of variables is followed by a description of the various tools used for data collection. Towards the end of the chapter, the reliability and validity of the measures are discussed.

3.2 Choice of variables

The review of studies presented in the previous chapter points to the fact that direct observation of the classroom is the best method of approaching the topic. These studies also indicate that the teachers' behaviour in the classroom is a key variable in influencing creativity among pupils. The nature of creativity has been studied from a number of viewpoints - psychological, environmental, cultural, physical and intellectual. A study of the environment where creativity is valued involves complexes of such environments as those of family, school,

community and culture. The modern view has linked creativity to the study of teaching situation and classroom environment as the teacher can identify and liberate the creative stock in his classroom. Review of related literature also shows that creative thinking is likely to be influenced by several factors like age, sex, socio-economic status of pupils, personality factors, achievement, intelligence etc. Thus, it is necessary to include several variables which are likely to affect the creativity of pupils. The investigator in this study selects the following variables which are believed to influence creativity among pupils:

1) Teachers' classroom behaviour
2) Pupils' personality factors
3) Pupils' Intelligence
4) Pupils' Academic Achievement.

Since the topic, "A study of the Relationship of Teacher Behaviour with creativity of Higher Secondary Pupils", deals with pupils' creativity, creativity is also taken as one of the variables. Having decided the variables to be included in the study, the investigator next directs his attention towards the selection of appropriate tools for data collection.

3.3 Systematic Observation

Since the study deals with the behaviour of the teachers, observation technique is the best suited method of data collection regarding teacher behaviour. Systematic observation is an accepted method of organizing observed teaching acts and any observational system usually includes some type of carefully defined items or categories so that observers can become skilled in identifying and recording brief ideas.
3.3.1 The Technique of Classroom Interaction Analysis

Interaction Analysis technique developed by Flanders has been used in the present investigation. Interaction analysis is an observational technique designed to observe and code classroom verbal behaviour of the teacher for every three seconds. Using ten categories, the observer sits in the classroom and observes what goes on, the interaction between the teacher and the pupils; at the end of the three second periods, codes what has been going on and makes a tally in the relevant category. The observations are in a sequence. The verbal behaviour of teachers can be coded into ten categories (FIACS) of which seven deal with the teacher talk, two with pupil talk and one is for silence or confusion. The most important element in this system is that of the type of influence the teacher has in the classroom. The influence is of two types. One is "Direct Influence" and the other is "Indirect Influence". The first four deal with 'Indirect Influence Pattern', which is characterised by giving more freedom of action for the students through accepting their feelings, facilitating their participation by praising and encouraging them, accepting their ideas and stimulating them to talk by raising questions. Categories five, six and seven deal with 'Direct Influence Pattern' in which there is little or no freedom of action to the students. When the teacher lectures continuously or gives definite directions and justifies his authority or reprimands the students, he is using 'Direct Influence'.
3.3.2 Flanders Interaction Analysis Categories

Teacher Talk (INDIRECT INFLUENCE)

i. **Accepts feelings**: Accepts or clarifies the feelings of the students in non-threatening manner. Feelings may be positive or negative, predicting and recalling feelings are included.

ii. **Praises or Encourages**:

Praises or Encourages students' action or behaviour. Jokes that release tension not at the expense of another individual, nodding head or saying "UM", "hm" are included.

iii) **Accepts or uses ideas of pupils**:

Clarifying or building or developing ideas suggested by a student. If the teacher brings more of his own ideas into play shift to category five.

iv. **Asks Question**:

Asking question about content or procedure with the intent that a student may answer.

Teacher Talk (DIRECT INFLUENCE)

v. **Lecturing**:

Giving facts or opinions about content or procedure expressing his own ideas, asking rhetorical questions.
vi. Giving Directions:

Directions or orders to which a student is expected to comply.

vii. Criticising or Justifying Authority

Statements intended to change students' behaviour from non-acceptable to acceptable pattern, hauling some one out, stating why the teacher is doing what he is doing, extreme self-reference.

viii. Student Talk Response:

This category is used when the teacher has initiated the contact or has solicited student statements, when the student answers a question asked by the teacher.

ix. Student Talk - Initiation:

Talk by students which they initiate, unpredictable statements in response to teacher.

x. Silence or confusion:

Pauses, short periods of silence and periods of confusion in which communication cannot be understood by the observer.

Flanders' Interaction Analysis Categories can be written as shown below:

1. Accepts feelings
2. Praises or encourages
3. Accepts or uses ideas of pupils
4. Asks questions
5. Lecturing.
6. Giving directions
7. Criticising or justifying authority
8. Student talk response
9. Student talk initiation
10. Silence/confusion.

3.3.3 Some Modifications

The Flanders' System as described concentrates only on verbal interaction between teachers and pupils. So only verbal behaviour can be analysed. It is fact that non-verbal behaviour is also equally important, so that more data may be collected in the classroom and thus more insight can be gained into the teaching behaviour. So additional categories were added to Flanders categories.

The following additional categories have been added to the FIACS.

1. Teacher reads aloud the text.
2. Pupil reads aloud the text.
3. Teacher uses the blackboard.
4. Pupil uses the blackboard.
5. Teacher gives guidance individually.
6. Pupils work silently - Teacher apparently doing nothing.
7. Mass answering by the pupils - class getting out of control.
3.3.4 Modified Categories of classroom Behaviour

1. Accepts feelings
2. Praises or encourages.
3. Accepts or uses ideas of pupils.
4. Asks questions.
5. Lecturing.
7. Criticising or justifying authority.
8. Student talk - response
9. Student talk - initiation
10. Teacher reads aloud the text
11. Pupil reads aloud the text.
12. Teacher uses the blackboard.
13. Pupil uses the blackboard.
15. Pupil Works silently.
16. Mass answering by the pupils.
17. Silence / Confusion.

3.3.5 Ground Rules

a. When not to which of two or more categories a statement belongs, choose the category that is numerically farthest from category five with the exception of category 10.

b. If more than one category occur during three second interval, then all the categories used in that interval are recorded. If no change occurs within three seconds, repeat the category number.
c. Directions are statements and result in observable behaviour on the part of pupils.

d. If there is discernible period of silence, record one 17 for every five seconds of silence.

e. When the teacher repeats a student's answer and the answer is a correct answer, this is recorded as a '2'. This reveals that the student has the right answer and therefore functions as praise.

f. When the teacher repeats a student's idea and communicates only that idea which is considered or accepted as something to be discussed, 5 is used.

g. If a student begins talking after another student a '17' is inserted between the 9's or 8's to indicate the change of student.

h. Statements such as 'ah', 'huh', 'yes', 'all right', 'okay', which occur between 9's are recorded as '2'.

3.3.6 Methods of Observation

Before starting the work, the investigator has to memorise all the categories, so that she may equip herself with the method of handling the tool. The investigator sits on the last bench of the classroom and observes the teacher when he is teaching. At an interval of every three seconds she writes down the category number which best represents the communication event just completed. For instance, when the teacher is lecturing, the investigator puts '5', when he asks questions, he puts
'4', when the pupils work silently, he puts '15'. This procedure of recording the events goes on at the rate of 20 to 25 observations per minute. In the end, a long series of numbers is obtained which is entered in a 17 x 17 matrix. When the matrix is prepared, two numbers are taken at a time. The first number stands for row and the second number for column. Thus each number in the series is used once as the row number and once as the column number.

As an example, the following sequence numbers can be classified as illustrated:

```
II     IV
4 ........ 4 ........ 12 ........ 5 ........ 5 ........ 5 ........
I III V
```

Tabulations are made in the matrix. They represent pair of numbers. The first pair is (4,4). The second pair is (4,12). The third pair is (12,5). The fourth pair is (5,5) and so on. Each pair overlaps the next. The total number of observations 'N' will always be tabulated by (N - 1) tallies in the matrix. Observation may continue for the whole period for 40 or 45 minutes. Within the period, the investigator may write 800 to 900 categories. From the tabulated matrix, a number of ratios representing different communication patterns may be calculated. The teachers may be asked to fill in the proforma, a copy of which is given in Appendix 1.

3.4 Assessment of Pupils' Personality Factors

Personality is 'The dynamic organisation within the individual of those psychological systems which determine his unique adjustment
To assess the personality factor of the pupils, the investigator used the Sixteen Personality Factor Questionnaire (16 P.F.), Form D, developed by Cattell, R.B. A copy of the 16 personality Factor Questionnaire, Form D, is given in Appendix 2. The 16 Personality Factor Questionnaire is an objectively scorable test devised by basic research in psychology to give the most complete coverage of personality possible in a brief time. The test was designed for use with individuals aged sixteen and above. Form D of the Questionnaire is most appropriate for literate individuals whose educational level is roughly equivalent to that of the normal high school student. The test can be scored by hand or by machine and various types of answer sheets are available for this purpose.

The personality factors measured by the 16 P.F. are not just unique to the test but instead rest within the content on general theory of personality. These 16 dimensions or scales are essentially independent. Any item in the test contributes to the score on one and only factor so that no dependencies were introduced at the level of scale construction.

In addition to the 16 primary factors, the test can be used as a measure of eight secondary dimensions which are broader traits scorable from the component primary factors.

3.4.1 Arrangement of Questions

In Form D, there are eight items for the Factor B scale and seven items for the motivational distortion scale.
The questions are arranged in a roughly cyclic order determined by a plan to give maximum convenience in scoring by stencil and to ensure variety and interest for the examinee. Three alternative answers are provided for each of the questions and the respondent should choose one among them.

3.4.2 Instruction of administration

3.4.2.1 General

Simple and clear instructions are printed for the examinee on the cover page of the test booklet. Although the test can be virtually self-administering, it is always important to establish good 'rapport' with examinees, whether tested individually or in groups.

3.4.2.2 Detailed Instruction

Answers are always made on a separate answer sheet never on the reusable test booklet. The investigator has to inform the respondents to use a separate answer sheet which is to be provided. He has to ask the respondent to enter his name, etc., at the top of the answer sheet and then to read the instruction on the cover of the test booklet to himself and then to work the four examples. The test is untimed; but it is good to remind examinees that they should not delay but should give immediate answers and move along. Care must be taken to ensure that one, and only one answer is given for every question on the test.

3.4.3 Principles and Mechanics of Scoring

Each answer scores 0, 1 or 2 points, except the Factor B
(intelligence) answers which score 0 (incorrect) or 1 (Correct). The score of each item contributes to only one factor total. Tests can be either hand scored with a stencil key, or machine scored. Hand scoring is accomplished by key, easily, rapidly and in a standard manner.

3.4.4 Converting Raw Scores to Stens

The meaning of raw scores from any form or combination of forms of the 16 P.F. depends upon the particulars used. Consequently, before these raw scores can be evaluated and interpreted, they must be converted into a system which places the examinee's score in relation to scores obtained by other people in some defined population.

The standardisation tables convert raw scores to stens, a practice consistent with best modern usage aiming at a good, but not unrealistically refined degree of accuracy in expression of results.

Sten scores are distributed over ten equal interval standard score points from 1 to 10, with the population average fixed at sten 5.5. The available selection of norm tables permits the conversion of any given raw score for any of the 16 personality factors to stens. The tables cover the general adult population and various sub-samples, with various tables for individual forms and for combinations of forms.

3.4.5 Calculation of Second-order Factor Scores

The 16 P.F. can be scored for broad secondary factors as well as for the sixteen primaries. Eight second-order factors have been identified and replicated at the present time. The first four, which will generally of most interest to the practitioners are:
Q I   Introversion vs. Extraversion
Q II  Low Anxiety vs. High Anxiety
Q III  Tenderminded Emotionality vs. Tough Poise.
Q IV  Subdueness vs. Independence.

Second order scores are more easily derived from the sten scores on the primaries than from raw scores. If the primary sten scores are combined in the manner shown in the manual, the resulting second-order scores will also be in sten form.

3.5. Assessment of Pupils' Intelligence

"Intelligence is the inherent capacity of profiting by experience, adoption to environment and ability to learn". To assess the intelligence of the pupils the investigator used the Culture Fair Intelligence Test Scale - 3, Form A, developed by Cattell. A copy of the CFIT Scale-3, Form A is given in Appendix 3. The Culture Fair Intelligence Tests measure individual 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. The test scale contains four sub-tests involving different perceptual tasks, so that the composite intelligence measures avoids spurious reliance on a single skill. The special advantage of the Culture Fair Intelligence Tests, is that they make a cleaner separation of natural ability from specific learning and thus permit better analysis and prediction of the individual's ultimate potentialities. The tests can be reasonably administered in a single session.
3.5.1 Description of the sub-tests

In the first subtest the individual is presented with an incomplete, progressive series. His task is to select from among the choices provided, the answer which best continues the series. The second subtest is classification. Here the individual is presented with five figures. The examinee must correctly identify two figures which are in some way different from the others. In the Matrices subtest the task is to correctly complete the design or matrix presented at the left of each row. The final subtest, conditions or topology requires the individual to select from the five choices provided, the one which duplicates the conditions given in the far left box. Before each subtest, examples are given so that the task requirements are clear to the examinee.

3.5.2 Items and Time allotted to each sub-test in Scale 3

<table>
<thead>
<tr>
<th>Test</th>
<th>Items</th>
<th>Time Allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>Series</td>
<td>12</td>
</tr>
<tr>
<td>Test 2</td>
<td>Classification</td>
<td>14</td>
</tr>
<tr>
<td>Test 3</td>
<td>Matrices</td>
<td>13</td>
</tr>
<tr>
<td>Test 4</td>
<td>Conditions</td>
<td>10</td>
</tr>
</tbody>
</table>

3.5.3 Directions for Administration

CFIT can be administered to groups or individually. Answers can be marked in the test booklet or on separate answer sheets thus permitting reuse of the test booklets.

The examiner should read the directions to himself several
times beforehand to familiarize himself with the test sequences, the timing and the instructions. He should have and be able to use efficiently a stop watch or a precise interval timer. It is desirable to begin with some appropriate introductory remarks to put the examinees at ease and to motivate them to do as well as possible. The instructions should be read exactly as they are given so that the examinees are prepared to do their best. Major departures from the instructions may seriously affect the results and render the norms inappropriate. Time limits must also be strictly adhered to and under no circumstances altered to allow fewer or more persons to complete a given subtest.

The room should be well lighted and the desks widely spaced. The examinees may be asked to fill in the information at the top of the answer sheet or cover page.

3.5.4 Scoring and use of the Norm Tables

Once the tests have been administered, there remains the task of obtaining raw scores from the answer sheets and converting these raw scores to an interpretable standard score. Scoring key is available for scale 3. A copy of the scoring key is given in Appendix 4. It greatly reduces the personnel time required to obtain the raw scores. Before placing the scoring key over the answer sheet, the examiner should check first for patterned responses that would indicate an invalid protocol. When he has remedied the problem, the test should be readministered. He should also check that inappropriate multiple responses have not been made. When inappropriate multiple responses are encountered, the item should be treated as if it were left blank.
The score obtained from the answer sheet is a raw score and needs to be transformed before it can be interpreted. Using the appropriate norm tables, the total raw score can be converted into a normalised standard score I.Q., thus permitting direct comparisons with many other available intelligence measures.

3.6 Assessment of Pupils' Academic Achievement

Achievement test was not selected as a tool for the present study, since one particular test in any one subject alone cannot do full justice in determining a student's academic achievement. Moreover, administering tests in all subjects to all students involves heavy consumption of time. Hence, it was resolved to assess the performance of the sample pupils through the total marks obtained by each of them in all subjects in the Half-yearly Examination. The total marks of each pupil was converted into percentage in order to bring them into common base.

3.7 Assessment of Pupils' Creativity

To identify the creative thinking of pupils, Baqer Mehdi's verbal and non-verbal tests of creative thinking are used, which will give scores for fluency, flexibility and originality in verbal, and elaboration and originality in non-verbal creative thinking. The total of these verbal and non-verbal tests gives a comprehensive score of creative thinking.

The verbal and non-verbal tests of creative thinking were administered to the sample and scored as per the instruction given in the manual. The total scores on fluency, flexibility and originality in verbal test, and elaboration and originality in non-verbal test, were tabulated and totalled up to get
the comprehensive creative thinking score.

Copies of the tools for measuring verbal and non-verbal creative thinking of pupils, are given in Appendix 5 and 6 respectively.

3.7.1 Instruction for Administration

The preliminary instructions to be given to pupils are recorded on the test booklet. After the instructions for motivating the children have been given, the test administrator should distribute the test booklets and ask them to fill in the columns for name, class, etc. After they have filled in the required columns, the test administrator should ask them to put down their pens and pencils. He should then ask them to look at the general instructions which the test administrator should read aloud letting the children to go through with him silently.

3.7.2 Verbal Test of Creativity

The verbal test of creativity includes four sub-tests, namely, Consequences Test, Unusual uses Test, Similarity Test and Product Improvement Test.

3.7.2.1 Description of the sub-tests

i. Consequences Test

The consequences Test consists of three hypothetical situations
(a) What would happen if man could fly like birds? (b) What would happen if our schools have wheels? (c) What would happen if man does not have any need for food?.

The subject is required to think as many consequences of these
situations as he can, and write them under each situation in the space provided. The test encourages free play of imagination and originality. The time allowed for the three problems is 4 minutes each.

ii. Unusual Uses Test

This test presents the subject with the names of three common objects - a piece of stone, a wooden stick, and water - and requires him to write as many novel, interesting and unusual uses of these objects as he may think of. The time allowed for the three tasks is 5 minutes each.

iii. New Relationship Test

This test presents the subject with three pairs of words apparently different - tree and house, chair and ladder, air and water, and requires him to think and write as many novel relationships as possible between the two objects of each pair in the space provided. The test provides an opportunity for the free play of imagination and originality. The time allowed for each pair of words is 5 minutes.

iv. Product Improvement Test

In this test, the subject is asked to think of a simple wooden toy of a horse and suggest addition of new things to it to make it more interesting for the children to play. The time allowed is 6 minutes.

The total time required for administering the test is 48 minutes in addition to the time necessary for giving instructions, passing out test booklets to children and collecting them back.
3.7.2.2 Procedure for Scoring

The following points have to be kept in mind while scoring. Each item is to be scored for fluency, flexibility and originality. The scores may be directly entered on the answer sheet by closely following the scoring guide. If the scorer comes across responses which are not mentioned in the scoring guide, he should briefly mention them on the backside of the answer sheet and score them for originality after all the tests scripts have been scored.

i. Scoring for fluency:

In scoring for fluency, the scorer would go through the responses to the item in question carefully and strike off those which are irrelevant and/or have been repeated. He should then count the remaining number of responses and enter this number as the fluency score for the items.

ii. Scoring for flexibility:

In scoring for flexibility, the scorer should first acquaint himself with the categories of responses given for each item in the scoring guide. For convenience sake, he should note in bracket against each response, the alphabet serial of the category to which it belongs. If he comes across a response which has not been mentioned in the scoring guide, he should himself determine the category to which it would seem to belong. After he has gone through all the responses to a given item, he should see how many different categories have been used by the testee. This can be easily determined on the basis of the number of different
alphapet serials used. The flexibility score will be the total number of different alphabet serials used.

iii. Scoring for Originality:

Originality scoring is done on the basis of statistical uncommonness of responses. The more uncommon the response, the higher the originality weight. If the response has been given by 0.1% to 0.99% of the testees, then the response will get an originality weight of 5; if a response has been given by 1% to 1.99% of the testees, then the response will get an originality weight of 4; if a response has been given by 2% to 2.99% of the testee, then the response will get an originality weight of 3; if a response has been given by 3% to 3.99% of the testees, then the response will get an originality weight of 2; and if a response has been given by 4% of 4.99% of the testees, then the response will get an originality weight of 1. Responses given by 5% or more of the testees, will get an originality weight of zero.

As given by the author of the tool, the originality weight for original responses were calculated by the researcher with respect to the sample under consideration, following the same guidelines.

3.7.3 Non-verbal Test of creativity

The non-verbal test of creative thinking is intended to measure the individual's ability to deal with figural content in a creative manner. Three types of activities are used for this purpose, namely, picture construction, picture completion and triangles and ellipses. The total time required for administering this test is 35 minutes, in addition to
the time necessary for giving instructions, passing out booklets and collecting them back. A brief description of these tests is given below:

3.7.3.1 Description of the Sub-Tests

i. Picture Construction Activity

The activity presents the subject with two simple geometrical figures, a semi-circle and a rhomb, and requires him to construct an elaborate picture using each figure as an integral part. The subject is allowed to turn the page to use the figure in any way he likes for making the picture. Emphasis is put on originality and elaboration. Originality is emphasised by the instruction that the subject should try to make as novel a picture as possible, such that no one else will be able to produce. Elaboration is emphasised by the instruction that the subject may add as many details as he thinks necessary in order to make the picture tell us complete and as interesting a story as possible. Ten minutes are allowed for the two tasks.

The pictures are scored for elaboration and originality. The subject is also asked to give an interesting and unusual title to each picture. The titles may also be scored for verbal elaboration and originality and the scores added to the verbal creativity score obtained on the verbal creativity test. The scoring of titles, however is optional.

ii. Incomplete Figures Activity

This activity consists of 10 line drawings which could be made
into meaningful pictures of different objects. The subject is asked to make a picture which is novel and asked to give an interesting and suitable title to each picture he makes. The subject is given 15 minutes for the 10 items. Each item is scored for elaboration and originality. Titles may also be scored for verbal elaboration and originality and the scores added to the verbal creativity score obtained on the verbal creativity test. The scoring of titles is optional and the researcher has included it in the computation of creativity scores.

iii. Triangles and Ellipses Activity

In this activity, the subject is provided with 7 triangles and 7 ellipses and he is required to construct different meaningful pictures based on the given stimuli. Each item is scored for elaboration and originality. A total time of 10 minutes is allowed for this activity.

The subject is also asked to give an interesting and suitable title to each picture which should also be scored for verbal elaboration and originality and the scores added to the verbal creativity score obtained on the verbal creativity test. The scoring of the titles, however, is optional.

The researcher, in the present study, while computing the scores, has also scored the title for all the three afore-said activities.

The three activities taken together provide ample opportunity to the subject to use his imagination with different types of figural tasks and come out with some novel ideas.
3.7.3.2 Procedure for Scoring

The following points have to be kept in mind while scoring the tests. Only the items in Activity III may be scored for flexibility. Scoring, however, is optional. The definitions of these terms are given below:

i. Elaboration

Elaboration is represented by a person's ability to add pertinent details (more ideas) to the minimum and primary response to the stimulus figure. The minimum and the primary response to the stimulus figure is that response which gives essential meaning to the picture.

It is recommended, as a general rule, that the criterion for determining the primary and minimum response is what is most essential for identifying the response; in other words, only these parts will be considered most essential without which the figure cannot be identified what it is meant to be. Thus in a human head, eyes and an indication of nose and mouth will be enough to identify it as head and so all other parts like hair, ear, neck, etc., should be considered as elaboration.

It is important for the scorer to see that the primary and minimum response is meaningful and relevant to the stimulus before it is scored. If the figure is not relevant and meaningful, it should be ignored. The total elaboration score will consist of a score of one for the primary and minimum response plus one score each for all the additional new ideas. An idea once scored in a picture should not be scored again in the same picture.
ii. Scoring of the Title

It has to be kept in mind that titles too are to be scored for elaboration and originality. These scores will be considered as verbal rather than non-verbal and will have to be calculated separately. In scoring the title for elaboration, again care has to be taken to identify the primary and minimum response and all additions to be taken as elaborations over it. For example, if a picture is of an aeroplane and the title reads 'Pilot returning after bombarding the enemy', the elaboration score for this title will be the number of ideas included in the title - one for the pilot, one for returning, one for bombardment, one for enemy (total score = 4). As the word aeroplane is not verbally mentioned no scores for aeroplane will be given. The primary response here, however, is aeroplane. As the title is conveying an interesting and novel idea which is relevant and meaningful, it is to be scored.

iii. Originality

Similar to 'originality' under verbal test.

As given by the author of the tool, the originality weights for original responses were calculated by the researcher with respect to the sample under consideration following the same guidelines.

3.8 The Pilot Study

In order to ascertain the feasibility of the study and to refine the tools, a pilot study was considered necessary. More specifically, the objectives of the pilot study were (i) To identify the defects in
the tools of data collection for remedying before final study, (ii) To identify in advance the problems in the methodology of research so that care could be bestowed to rectify them before starting the final study.

One tenth of the teacher and pupil samples was selected for the pilot study. The investigator observed 6 Post-Graduate teachers teaching English at Standard XI in three Higher Secondary Schools in Coimbatore for one period each. Modified version of FIACS was used for observation of the classes. About 150 pupils of the same teachers were also given the four tests, viz., 16 P.F. Questionnaire, CFIT, verbal and Non-Verbal Tests of creative thinking. 17 x 17 matrices were prepared for all the six teachers. All the three test papers of the pupils were scored and analysed.

The pilot study was helpful in identifying the problems in the methodology of research which were rectified before starting the final study.

3.9 Final Study

3.9.1 Sample of the Study

There are about 95 Higher Secondary Schools in Coimbatore Revenue District managed by different agencies such as Government, Corporation, Municipalities and other Private Bodies. The schools are of different types such as Boys' Schools, Girls' Schools or Mixed Schools. The schools are located at Rural, Urban and Semi-Urban areas. Pupils study either in the English Medium classes or in Tamil medium classes in these schools. The teachers also vary with respect to sex, experience,
educational and professional qualifications. Stratified Random Sampling Technique was adopted for the selection of schools for the present study. The distribution of the teacher sampling is given in the Appendix 7. The distribution of the pupil sampling is also given in the Appendix 8.

The Teacher sample comprises 26 female post-graduate English Teachers and 34 male post-graduate English Teachers (60 Teachers) handling standard XI. The pupil sample comprises 775 female pupils and 725 male pupils (1500 pupils).

3.9.2 Collection of Data

The collection of data for the final study was done personally by the investigator who visited all the 60 classes in 30 Higher Secondary Schools in Coimbatore Educational District.

The classrooms of 60 Post-Graduate teachers teaching English at Standard XI, in 30 Higher Secondary Schools in Coimbatore District were observed for one period each. The investigator used the modified version of FIACS for the observation of these teachers' classroom behaviour. The bio-data of the observed teachers were obtained through the proforma as given in the Appendix 1. 17 x 17 matrices were prepared for all the 60 teachers and were analysed separately.

The pupils of the teachers whose classroom behaviour had already been observed, were given the four tests, viz., 16 P.F. Questionnaire Form D, CFIT, verbal and non-verbal tests of creativity, one after another. On an average, 3 hours were taken by the pupils to complete all the four tests. While administering the CFIT, time limits were strictly adhered to, and under no circumstances altered, to allow fewer or more persons to complete a given subtest. In the case of verbal and non-verbal
tests of creative thinking also, time limits were strictly followed. But there is no such strict time limitation for completing the 16 P.F. Questionnaire. The pupils were asked to fill in the information at the top of the answer sheets. Proper care was taken to see that the pupils clearly understood the instructions how to respond to the tests. The answer papers were scored, tabulated and analysed.

3.10 Reliability and Validity of the Tools

The reliability for a set of scores from a group of examinees is the co-efficient of correlation between that set of scores and another set of scores on an equivalent test obtained independently from members of the same group. It helps to know the representativeness and the reliability of the measurements made through the test.

A test is valid if it measures what it purports to measure. The validity of the measures and thereby of the tool is lowered if the measures are contaminated by assessment of some other feature or trait besides the target one. There are different types of validity depending both on what the scores measure and what for they do it, i.e., on the matter and the purpose of measurement.

3.10.1 The Problem of Observer Training and Reliability

Individuals differ in their ability to become reliable observers. Both accuracy of judgement in classifying verbal interaction and consistency in judging the behaviours correctly are necessary qualities of a good observer. Unless recorded behaviours are actual observed behaviours, a system's usefulness is limited. The greater the disparity
between observed and recorded behaviours, the less useful the system. To provide an estimate of an observer's validity or reliability, the collected data is compared to that collected from the same setting by an expert. Sets of data should be collected under similar conditions, if not identical, for the exact accepted time sample.

Consistent observation by a team requires group training, discussion of common ground rules, each observer's understanding of his own unique biases and regular meetings after training to discuss unusual categorization problems. The ideal observer team is a group of like-minded individuals who will respond consistently with the same category number when presented with the same communication events.

During the pilot study, a trained observer also observed the same classes which the investigator was observing. Inter-observer reliability was calculated for each class and the training in observation continued till the reliability reached a sufficiently high level.

Generally, a coefficient of 0.6 is frequently established, as an acceptable level to illustrate mastering of the system's collection skill. The investigator applied the Scott's formula for finding out the reliability correlation of her tool. According to Scott's formula, reliability is equal to the total agreement between the two observers minus the agreement that occurs by chance divided by the greatest possible agreement that occurs by chance. This reliability is a comparison of the agreement between the observers not resulting by chance and the greatest agreement possible that does not result from pure chance.
Reliability = \frac{\text{Total agreement between two observations} - \text{chance agreement}}{	ext{Greatest possible agreement} - \text{chance agreement}}

The calculation of the reliability correlation is given in Table 3.1

<table>
<thead>
<tr>
<th>Category No.</th>
<th>Investigator percentage</th>
<th>Observer percentage</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.67</td>
<td>0.69</td>
<td>--</td>
</tr>
<tr>
<td>2.</td>
<td>8.08</td>
<td>9.02</td>
<td>--</td>
</tr>
<tr>
<td>3.</td>
<td>13.48</td>
<td>13.98</td>
<td>--</td>
</tr>
<tr>
<td>4.</td>
<td>7.64</td>
<td>7.88</td>
<td>--</td>
</tr>
<tr>
<td>5.</td>
<td>13.25</td>
<td>12.75</td>
<td>0.50</td>
</tr>
<tr>
<td>6.</td>
<td>4.49</td>
<td>4.52</td>
<td>--</td>
</tr>
<tr>
<td>7.</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>8.</td>
<td>14.15</td>
<td>13.03</td>
<td>1.12</td>
</tr>
<tr>
<td>9.</td>
<td>8.98</td>
<td>9.02</td>
<td>--</td>
</tr>
<tr>
<td>10.</td>
<td>3.37</td>
<td>3.39</td>
<td>--</td>
</tr>
<tr>
<td>11.</td>
<td>3.82</td>
<td>3.71</td>
<td>0.11</td>
</tr>
<tr>
<td>12.</td>
<td>7.64</td>
<td>7.52</td>
<td>0.12</td>
</tr>
<tr>
<td>13.</td>
<td>4.94</td>
<td>5.02</td>
<td>--</td>
</tr>
<tr>
<td>14.</td>
<td>4.04</td>
<td>4.13</td>
<td>--</td>
</tr>
<tr>
<td>15.</td>
<td>1.34</td>
<td>1.36</td>
<td>--</td>
</tr>
<tr>
<td>16.</td>
<td>3.59</td>
<td>3.52</td>
<td>0.07</td>
</tr>
<tr>
<td>17.</td>
<td>0.44</td>
<td>0.46</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.92</td>
<td>2.00</td>
</tr>
</tbody>
</table>
According to Scott's formula,

\[
 r = \frac{P_0 - Pe}{1 - Pe}
\]

Where \( P_0 \) = Greatest possible agreement - disagreement

\[
P_0 = 1 - \text{disagreement}
\]

\[
= 1 - 0.0392 \text{ (Percentage value is converted into decimal values)}.
\]

\[
= 0.96
\]

Agreement due to chance \( Pe \)

\[
= (0.135)^2 + (0.142)^2 = 0.04
\]

Therefore \( r = \frac{0.96 - 0.04}{1 - 0.04} \)

\[
= \frac{0.92}{0.96} = 0.96
\]

This shows that the reliability of the agreement between the two observers is quite high and significant.

**Validity of the Tool**

The investigator has observed the classes with a standard category system of observation. So, it is sure that the tool has a good and high content validity also.

### 3.10.2 Reliability and Validity of the 16 P.F. Questionnaire

The consistencies of the 16 P.F. scales, i.e., the agreement
of the factor score with itself, under some changes of conditions are given in all relevant ways in the manual. The first type of consistency to consider is reliability or the agreement of the factor score over time. Reliability may be further subdivided into (a) dependability, i.e., short-term test - retest correlations and (b) Stability, i.e., retest after a long interval.

The investigator readministered the same test to the 5% of the same pupil sample, 3 months after the first administration of the test. A correlation co-efficient was carried out between the two sets of scores and the results are given in Table 3.2.

From the Table 3.2, it is found that the tool is highly reliable.

Validity

The items in these final forms are the survivors from several thousands of items originally tried and constitute only those which continue to have significant validity against the factors after ten successful factor analysis (Cattell, 1973) on different samples. These analyses have both verified the existence and natural structure of the sixteen factors, and cross-validated the test items in their correlation with the factors on different adult population samples.

3.10.3 Reliability and Validity Estimations of the CFIT Scale 3 Form A.

The investigator adopted Test - Retest method for estimating the reliability of the test.
A group of students is given a test once and the scores are noted. After sometime, the same test is given to the same group of students and scores are noted. If the correlation between the results first obtained and those obtained for the second time, is sufficiently high, the test is said to be reliable. A limitation of this method is that the pupils might have remembered more or grown matured or have learnt something during the time that elapsed between the first trial and second trial of the test.

The researcher readministered the same test to 5% of the total sample of 1500 pupils after 3 months from the date of the first administration of the test. An attempt was made to calculate the correlation co-efficient between the two sets of scores by scattergram method. The results are given in Table 3.3.

From the results it is clear that the CFIT Scale 3, Form A, is highly reliable.

Validity of the CFIT Scale 3 Form A

The investigator attempted to find out the correlation co-efficient between the pupils' mean achievement score obtained from this scale and (a) the pupils' mean achievement scores on the Tests of creative Thinking and (b) the pupils' mean score on the Academic Achievement score. The correlation co-efficients were found to be (i) 0.496 and (ii) 0.720 respectively. These correlation co-efficients are significant at 0.01 level. Hence it is concluded that the CFIT scale 3, Form A, is highly valid.
3.10.4 Reliability and validity Estimations of the verbal and Non-verbal Tests of creative Thinking

To establish the reliability of the test, the investigator adopted the Test - Retest method. She readministered the same test to 5% of the total sample of 1500 pupils after three months from the date of the first administration of the test. An attempt was made to find out the corelation co-efficient between these two sets of creativity Test Scores by Scattergram method. The results are given in Tables 3.4 and 3.5. The calculated \( r \) value for the verbal Test of Creativity is found to be 0.985 and significant at 0.01 level ; and for the non-verbal test of creativity is found to be 0.944 and significant at 0.01 level. Hence it is concluded that the Test of verbal and Non-verbal creative Thinking by Bāqer Mehdi are highly reliable.

Validity of the Verbal and Non-verbal Tests of Creative Thinking

To establish the validity of the verbal Tests of Creative thinking, the investigator attempted to find out the correlation co-efficient between the pupils' mean scores of the verbal creativity scores and (i) the pupils mean scores on the Non-verbal Creativity Test, (ii) the pupils' mean scores on I.Q., and (iii) the pupils' means scores on Academic Achievement. The correlation co-efficients were found to be 0.828, 0.439 and 0.379 respectively and significant at 0.01 level.

To establish the validity of the Non-verbal Test of creative thinking, the investigator attempted to find out the correlation co-efficient between the pupils' mean scores of the non-verbal creativity scores and
(i) the pupils' mean scores on the verbal creativity Test, (ii) the pupils' mean scores on I.Q., and (iii) the pupils' mean scores on Academic Achievement. The correlation co-efficients were found to be 0.828, 0.500 and 0.483 respectively and significant at 0.01 level.

Hence, it is concluded that the verbal and Non-verbal Tests of Creative Thinking by Baqer Mehdi are highly valid.

3.11 Conclusion

The data collected from 60 teachers and 1500 pupils, have been statistically analysed and details are dealt with in the next chapter.