CHAPTER THREE

THE PROCEDURE

1. INTRODUCTION

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

2. CHOICE OF VARIABLES

The review of studies presented in the previous chapter highlights the idea that classroom climate is affected by a multitude of variables. Almost all the studies point to the fact that direct observation of the classroom is the best method of approaching the topic. They also indicate that the behaviour of the teacher is a key variable in developing the classroom climate. The sex factor has been a major variable in several
studies. A few studies have attempted to relate sociometric data to school and classroom climate. The attitude of the teacher is also found to be an important variable. A few studies, based on factor-analysis, point to the fact that variables related to pupil-teacher relationship are also significant in the study of classroom climate. Thus it is necessary to include several variables which are likely to affect the classroom climate.

A careful scrutiny of the variables used in the various studies referred to in the previous chapter was made and a list was prepared. There were some variables which could be easily combined because they referred to different aspects of the same variable. After a little bit of editing, the list included about 30 different variables. It would be highly ambitious to attempt to include all of them in any single investigation. Practical considerations and limitations of time and other factors suggest that a selection should be made and not more than 10 variables could be included in the study. In order to select these ten variables, it was decided to take the views of the consumers of this research - the heads of schools, experienced teachers and faculty of training colleges.

3. **VIEWS OF EDUCATIONISTS AND TEACHERS**

In order to decide the variables to be included in the study, a short questionnaire was prepared to get the reactions
of the consumers of this research. The questionnaire contained just four questions. In the first question, the respondents were asked to express what they understood by the term 'classroom climate'. The aim was to get at the raw thinking of the respondents. In the second question, they were asked to suggest three ways by which classroom climate could be measured. The aim was to get at the suggestions of the persons having intimate contact with classrooms. The third question contained a list of 30 variables, selected from the research studies reported in chapter two. The respondents were asked to indicate which of the variables, in their estimation, affected classroom climate. They indicated their response by putting a tick mark (✓) against variables of their choice. In order to generate free thinking, the last item of the check list was kept open where the respondent was given freedom to add any other variable not included in the checklist. In the last question, the respondents were asked to indicate their preference, in order of importance, regarding the variables they had selected. A copy of the questionnaire is presented in Appendix-A.

The printed copies of the questionnaire were sent to a number of headmasters and headmistresses of leading high schools in the locality, the members of the faculty in the Training Colleges in the State and educationists who were interested in the project. Since the aim was to get at the thinking of educationists, no attempts were made to follow strictly the procedures of sampling.
However a good representation of various types of schools was included. The questionnaires were sent to 200 individuals. 120 individuals were kind enough to send back the questionnaire with their responses. These responses were analysed. The following scoring system was used to convert the responses to a numerical index.

If an item was marked Rank 1 – The weight was 4
If an item was marked Rank 2 – The weight was 3
If an item was marked Rank 3 – The weight was 2
and If an item was just ticked – The weight was 1

The score for each item in the checklist was found by adding the weights and the items were ranked on the basis of the scores. Table 1 gives the results of the analysis.

The results indicated that the following ten variables topped the list, as most pertinent and related to classroom climate.

1. Teacher-pupil relationship.
2. Teaching procedure followed
3. Teacher attitude towards pupils
4. Teacher attitude towards profession
5. Pupil attitude towards learning
6. Teachers' capacity to control the class
7. Freedom to children, to ask questions etc.
8. Size of the class
9. Pupils' capacity to grasp what is taught
and 10. Pupils' attitude towards teacher.
The respondents, however, could not react in any specific way to the first two questions - viz. what they understood by the term 'classroom climate' and how it could be measured. An analysis of the responses to these questions brought out that classroom climate was mostly associated with teacher-pupil interactions and dealings which could either facilitate or debilitate the learning process intended to happen in the class. Most of them laid emphasis on the psycho-social factors and conditions obtaining in the classroom. More than 80 per cent of the respondents suggested observation of the classrooms, directly or indirectly, as the appropriate method to be followed. The observations should lead to an assessment of the teacher behaviour, the behaviour and response of the students, and an evaluation of the work going on in the class. A few respondents have also said that the physical environment should also be taken into consideration, though in the responses to the items in the check-list this aspect did not rank high.

In the light of these responses, the investigator decided to begin the investigation with the following variables. For convenience and organisation they were classified into four groups.

(a) The teacher: Sex, Professional qualifications, attitude towards profession, attitude towards pupils and perception of the class.
(b) **The Class:** The economic and educational background of the families from which the pupils come, the class composition in terms of socio-metric data.

(c) **The Interactions:** The teacher behaviour in the class and the pupil behaviour in the class during the lesson.

(d) **Other factors:** The physical conditions of the class and the environment.

These variables broadly cover all the ten variables suggested by the expert educationists. Having decided the tentative list of variables to be included in the study, the investigator next directed his attention towards the development of appropriate tools for collection of data.

4. **SYSTEMATIC OBSERVATION**

Scientific inquiry is often based on careful systematic observation. In order to be useful, the observation has to be made more systematic and reliable. Observational systems, developed for the purpose of identifying, classifying, quantifying and analysing classroom behaviours and interactions, are in use throughout the profession, by the teachers, researchers and teacher educators. Systematic observation is an accepted method of organising observed teaching acts with accuracy and reliability.
Several systems of observation have been evolved since the days of Anderson (1930) and a dozen such systems have already been discussed in chapter two. In India, the system evolved by Flanders seems to have attracted the attention of research workers in the field. It appears as though most (if not all) of the studies in this area rest upon observations made using Flanders Interaction Analysis Categories.

Flanders Interaction Analysis Categories is primarily concerned with the verbal behaviour of teachers. The system is based on the assumption that the verbal behaviour of the teacher in an adequate sample of his total behaviour. The chart presented below, gives a summary of the categories, with brief definitions for use by the observer.

SUMMARY OF CATEGORIES IN FIAC.

1. **Accepts Feelings**: accepts and clarifies the feeling tone of the students in a non-threatening manner. Feelings may be positive or negative. Predicting and recalling feelings are included.

2. **Praises or encourages**: praises or encourages student action or behaviour. Jokes that release tension, not at the expense of another individual, nodding head or saying "wh, huh?" or "go on" are included.

3. **Accepts or uses ideas of student**: clarifying, building or developing ideas or suggestions by a student. As teacher brings more of his own ideas into play, shift to category five.

4. **Ask questions**: asking a question about content or procedure with an intent that a student should answer.
5. **Lectures**: giving facts or opinions about content or procedure; expressing his own ideas; asking rhetorical questions.

6. **Gives directions**: directions, commands or orders with which a student is expected to comply.

7. **Critizes or justifies authority**: Statements, intended to change student behaviour from non-acceptable to acceptable pattern; bawling someone out; stating why the teacher is doing what he is doing; extreme self-reference.

8. **Student talk - response**: Talk by student in response to teacher. Teacher initiates the contact or solicits student statement.

9. **Student talk - initiation**: Talk by students, which they initiate. If "calling on" student is only to indicate who may talk next, observer must decide whether student wanted to talk. If he did, use this category.

10. **Silence or confusion**: Pauses, short periods of silence, and periods of confusion in which communication cannot be understood by the observer.

The observer notes down the teacher/student behaviour using one of the codes. A record is made every three seconds. When not certain in which of two or more categories a statement belongs, the observer chooses one that is numerically farthest from category 5. The tone of the teacher is to be considered in deciding whether the teacher talk is direct or indirect. The observer must not be very much concerned with his own biases or with the teacher's intent. The question is simply, "what category
best describes this particular bit of interaction?" If more than one category occurs during the three second interval, then all the categories used in that interval are to be recorded; therefore, each change in category is recorded. If no change occurs within three seconds, the same category number is repeated. If silence prevails for a period of 3 seconds or more, it is recorded as 10. The sequence of events is of importance and therefore the observer notes down the category number either horizontally or vertically always preserving the sequence.

By convention all records start with 10 and end with another 10.

Using these general ground rules, Flanders and others have successfully coded many classroom verbal behaviour of teachers to yield fruitful results.

The chief advantages of this system are two. First the number of categories is only ten so that the observer can easily memorize these categories. Thus the system is compact. Second the categories are well defined and the observer, within a short period, can master the categories and code the teacher behaviour easily and accurately. These two advantages have perhaps attracted many Indian researchers to use the FIAC in their studies. The studies conducted at CASE, Baroda and elsewhere are all based on FIAC.

However, the system is not without a few limitations. Chiefly, the system is based on the verbal behaviour of teacher to the exclusion of other types of behaviour. Oral communication
to be effective, has to be supported with non-verbal behaviour. Hence it was felt necessary to modify the system by introducing certain non-verbal behaviours which so often happen in the classrooms. For example, teacher demonstrations, teacher using the black board, or teacher giving individual help to weak or needy student are important classroom procedures. Similarly demonstrations by students or student using the blackboard are also significant activities of the class. Another important behaviour is mass response by the class. If the teacher assigns a work and the students work silently - it is a feature to be counted. Thus the investigator felt that, for the purpose of this study, a few more categories of behaviour are to be added to the list of categories by Flanders. The following categories were therefore framed and added to the first nine categories of FIAC.

<table>
<thead>
<tr>
<th>CATEGORY No.</th>
<th>DESCRIPTION/BEHAVIOUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Teacher demonstrates - conducts an experiment in the class - uses maps and other aids.</td>
</tr>
<tr>
<td>11</td>
<td>Student demonstrates, conducts an experiment in the class - uses maps and other aids.</td>
</tr>
<tr>
<td>12</td>
<td>Teacher writes or draws a figure on the black-board.</td>
</tr>
<tr>
<td>13</td>
<td>Student uses the black-board.</td>
</tr>
<tr>
<td>14</td>
<td>Teacher gives individual guidance to a needy student; checks the work of an individual student.</td>
</tr>
<tr>
<td>15</td>
<td>Class works silently on a given assignment.</td>
</tr>
<tr>
<td>16</td>
<td>Mass response</td>
</tr>
<tr>
<td>17</td>
<td>Silence/confusion/apparently no work going on.</td>
</tr>
</tbody>
</table>
It can be seen that the above modification does not disturb the first nine categories of Flanders classification. Flanders category number 10 is now taken last and is marked category 17. The new categories 10 and 11, and categories 12 and 13 are reciprocative in nature, in the sense that they describe the same behaviour (a) by the teacher and (b) by the student. The RCS system, described in chapter two, has the whole system based on this idea of teacher and pupil action. Thus the modified system has seventeen categories. The ground rules for observing and recording were the same as in the case of FIAC, in spirit. The complete list of the categories is presented in Appendix B. This modified form was tried during the pilot study. It was expected that this observation would yield data regarding (i) Teacher - pupil relation (ii) Teaching procedure, (iii) Teacher's control over class (iv) Pupil behaviour in the class.

5. **ATTITUDE MEASUREMENT**

In order to study the attitude of teachers towards the profession, it was decided to construct and standardize attitude scales.

Attitude expresses the individual's potential direction of reaction towards certain object or value. It is an indication of the future reaction that can be expected in an individual.

Allport defines attitude as a 'mental and neural state of readiness,
organised through experience, exerting a directive or dynamic influence upon the individual's responses to all objects and situations with which it is related.\(^1\) Newcomb similarly speaks of an individual's attitude towards something as a state of readiness for motive arousal; it is his predisposition to perform, perceive, think, and feel in relation to the thing.\(^2\) Attitude determines a consistent characteristic mode of reaction in relevant situations, which may involve people, an issue, an institution, an object or value.\(^3\) These definitions clearly bring out the fact that attitudes refer to the mental readiness to act in a particular direction with respect to the object under consideration.

Attitude scales have been extensively used in the measurement of attitudes. Several procedures for the construction of scales to measure attitudes have been suggested. The three commonly used procedures are: the method of equal — appearing intervals, the method of summated ratings and scalogram analysis. The first method has been made use of by Thurstone and Chave. The method of summated ratings was popularised by Rensis Likert. Guttman popularised the scalogram technique.

2. Newcomb, T.M. in Otto Klineberg, op.cit. 482
The method of equal appearing intervals is fully described by Thurstone and Chave in "Measuring attitude towards Church."¹

The first step is to collect a number of statements or propositions which express some kind of opinion about the object under study. The statements should be selected under two major criteria:

(a) the statements should express opinion ranging from the most favourable shade to the most unfavourable shade and (b) they are simple and un-ambiguous. These statements are then given to a number of judges, who individually classify them on an eleven point continuum. The interval between these points are assumed to be equal and hence the name 'equal appearing intervals'. The next step involves an analysis of the placements of each statement by the several judges. Two statistics are calculated for each statement: the median 'M' and the quartile deviation 'Q'. It is customary to use graphical methods for this purpose. The median 'M' is taken as the scale value of the statement, while the quartile deviation 'Q' serves as a measure of ambiguity. "A small quartile deviation indicates relative agreement of the judges about the scale value of the statement, while a large quartile deviation indicates relative disagreement among the judges."²

The final scale is built up by choosing a number of items which have low Q-values and whose scale values spread over the

1. Thurstone and Chave (1929). *Measuring attitude towards Church.*

entire range as evenly as possible. The subjects are asked to check the items which closely express their own attitudes. The mean or median scale value of the items checked by the subject is taken as the attitude score of the subject.

The chief criticism to this method stem out of the use of judges in determining the scale value of the statements. Fransworth P.R. has shown that the intervals between the eleven categories were not interpreted by the judges as equal. Nor was the midpoint of the scale consistently regarded as representing a neutral attitude.\(^1\) Further the effects of the attitude of the judges cannot be completely eliminated.

The method of summated ratings owes its origin to Rensis Likert who made use of item – analysis procedures for selecting attitude statements.\(^2\) The steps involved in this procedure may be summarised as follows. As in the method of equal appearing intervals, the first step is to collect a number of attitude eliciting statements, according to certain criteria, which express different shades of opinion. The carefully edited list of statements is then presented to a jury and classified into two categories as those which express a favourable attitude and those which express an unfavourable attitude. Next, the statements are tried on a sample of subjects who are asked to respond to


\[^{2}\] Remmers. H.H. *op.cit*. 94
each statement on a five point scale, ranging from strong agreement to strong disagreement. The responses are then quantified using a system weights where-in strong agreement with the favourable statements and strong disagreement with unfavourable statements receive high weights. A subject's score is the sum total of the weights assigned to the responses he has made. The items are then analysed for their discriminating power. Several statistical procedures have been developed for this purpose. Most of them are based on comparing two extreme groups showing highly favourable and highly unfavourable attitudes. Statements which show high discriminating power are selected for the final attitude scale.

The chief advantage of this method over Thurstone's method is the avoidance of the elaborate classification of statements using a number of judges at the initial stage. The main criticism against Likert's method rests on the fact that the same attitude score can be obtained through different patterns of responses to the various statements. However in practical applications, this is not found to have any adverse influence.

Scalogram Analysis, as popularised by Guttman and others, differs considerably from the methods described above. If a set of statements with a common content is to constitute a Guttman scale, then an individual with a higher rank or score than another individual on the same set of statements must also
rank just as high or higher on every statement in the set as the other individual. This means that a person with a more favourable attitude score than another person must also be just as favourable or more favourable in his response to every statement in the set, than the other person. When this criterion is satisfied, the set of statements is said to constitute a 'unidimensional scale.'

The Cornell technique assumes that the subjects have made their responses in either of the categories agree or disagree, so that the weights are 0 or 1. The technique describes the procedure for analysing a set of statements and select from them so that 'the co-efficient of reproducibility' is as high as possible. Several modifications and improvements have been suggested by Goodenough, Edwards, Kilpatrick and others to this procedure. This method has been found to work well when a relatively small number of well conceived attitude statements are initially available.

It was decided, in the light of the relative merits and limitations of the three major procedures, that two attitude scales be constructed using the Likert technique to estimate the teacher attitude towards (a) pupils and (b) the profession.

In this context, it was felt that it would be economic and time saving to build further on an attitude scale already constructed by the investigator as part of his M.Ed. thesis on

"The attitude of Graduate Teacher Trainees towards the Teaching Profession". This scale consists of 50 statements selected after appropriate item analysis of an initial list of 200 statements. However it was decided to do an item analysis a second time, based on the reactions of teachers (instead of teacher trainees). To study the attitude towards pupils a new set of 50 statements was prepared. These statements, after necessary editing, were presented to a jury consisting of five professors from a training college and classified into favourable/unfavourable categories. Thus in the try-out form, each attitude scale consisted of 50 statements. This form was used to estimate the attitude of eleven teachers whose classes were observed during the pilot study.

Simultaneous with the pilot study, the attitude scales were administered to 400 teachers drawn from 36 schools. A copy of this form of the attitude scales is presented in Appendix C. The teachers were requested to read each statement in the attitude scale and express their degree of agreement or disagreement on a five point scale. The five points were: SA - Strongly Agree, A - Agree, N - Neutral, D - Disagree and SD - Strongly Disagree. To avoid confusion and to facilitate tabulation all the statements were numbered serially from 1 to 100. However during analysis they

2. This sample is described in Table 2.
were separated; the first fifty deal with attitude towards pupils and the rest with attitude towards profession.

The responses were then scored according to the following scheme of weights:

<table>
<thead>
<tr>
<th></th>
<th>SA</th>
<th>A</th>
<th>N</th>
<th>D</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favourable Statements</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Unfavourable Statements</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

An attitude score was struck for each teacher by adding the weights assigned to the 50 statements. The maximum possible score is, thus, equal to 250 and the minimum is 50.

The data were then punched on IBM cards for item analysis. Two cards were punched for each teacher: the first card contained data regarding the scores on attitude towards pupils and the second card, attitude towards teaching profession. The following was the scheme of field arrangement for punching.

<table>
<thead>
<tr>
<th>Col.No.</th>
<th>Data - code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Punched 1 attitude towards pupil or 2 attitude towards profession</td>
</tr>
<tr>
<td>2 - 4</td>
<td>Three numbers - Serial number identifying the teacher-001 to 400.</td>
</tr>
<tr>
<td>5 - 6</td>
<td>Two number code - school identification number -01 to 36</td>
</tr>
<tr>
<td>7 - 10</td>
<td>Blank Space</td>
</tr>
<tr>
<td>11 - 60</td>
<td>The score on each of the 50 statements (item-score)</td>
</tr>
<tr>
<td>61 - 64</td>
<td>Blank Space</td>
</tr>
<tr>
<td>65 - 67</td>
<td>Total Score on the scale (3 digits)</td>
</tr>
<tr>
<td>68 - 80</td>
<td>Blank Space</td>
</tr>
</tbody>
</table>
The item analysis was done separately for each scale. The following procedure was adopted. The cards were arranged in order of the total score. One hundred cards from the high score end were selected to form the high group and one hundred cards from the low score end were selected to form the low group. The mean score and standard deviation for each statement were then computed. The difference between the means of the high and the low group was tested for significance using the t-test. This analysis was done on IBM 1620. The computer programme, in FORTRAN II, evolved by the investigator is given in Appendix D.

The results of the analysis are presented in Tables 3 and 4. Items found to have significant difference between means were tentatively selected.

To further strengthen the scale, the item - total correlations were also worked out. For this purpose a systematic sample of 100 cards, choosing every 4th card from the entire sample was selected. The computer programme developed for this purpose is given in Appendix E. It is written in FCS Language. The programme was run on TDC - 321. The results of the calculations are given in Tables 5 and 6. Items which showed a statistically significant correlation of 0.3 and above were considered to be good.

Combining both the criteria of significant mean difference and significant item - total correlation it was found feasible to select 25 statements for the final form of each attitude scale. This final form has been made use of at the final study stage.
6. **TEACHER AND PUPIL PERCEPTION OF THE CLASS**

Besides these two major variables, it was planned to include a study of the teacher's perception of the pupil and the pupil perception of the class. This would be helpful in understanding how each one feels about the class. For this purpose it was thought appropriate to use an evaluation procedure using bipolar descriptions. Lindzey and Byrne have reported that "A version of the semantic differential has been employed in several socio-metric studies. The sum of the ratings on the evaluative (for example, good - bad) scales are summed to yield a subject's attitude towards any object, including another individual. (Dentch and Solomon, 1959; Fishbein, 1963, 1965)..."

A number of investigations have employed an adjective check list consisting of positive and negative descriptive terms of adjective pairs of polar opposites or ratings on a series of traits. Attraction toward another individual is indicated in terms of positive and negative descriptions (for example, Davis and Warnath, 1957; Jones and Daugherty, 1959; Jones et.al. 1959; Lerner, 1965; Newcomb, 1961)."

Hence it was decided to select and use appropriate bi-polar descriptions to study the teacher perception of the pupil and the pupil perception of the class. Osgood et.al. in

their book "Measurement of meaning", have presented a long list of bi-polar descriptive, adjectives and their research shows that these can be grouped into three categories: (1) those which have high loading on "Evaluative" dimension, (2) those which have high loading on "Potency" dimension and (3) those which have high loading on "Activity" dimension. Since the objective of the present study is more evaluative in nature, it was decided to choose those bi-polar pairs which have high loading on the evaluative dimension.

The rating scale used in the pilot study consists of the following pairs of descriptions:

(a) **pupil perception of the teacher**

The pupils in the class were asked to express their feelings about the teacher who had just taught them, by placing a tick mark (✓) inside the appropriate bracket. There were seven positions between the two poles as follows:

Good ( ) ... ( ) ... ( ) ... ( ) ... ( ) ... ( ) Bad.

The other bi-polar pairs were: Intelligent - dull; Interested - Uninterested; Agile - slow; Strict - easy going; Honest - dishonest; Young - old; Calm - emotional; Praiseworthy - condemnatory; Brave - coward; Rich - poor; Has an aim - aimless; Capable - not capable; Clear - confusing.

(b) **Teacher perception of the class:**

In this case, the teacher was asked to evaluate the class, as a whole, using the same procedure as in the previous case. Twenty-four pairs of bi-polar descriptions were included in this case. They are: Good - bad; Compact - unwieldy; Homogeneous-heterogeneous; Sharp - dull; Interested - uninterested; Willing to learn - not willing to learn; Obedient - troublesome; Very active-passive; Happy - sad; Hard working - easy going; Pleasant - unpleasant; Honest - dishonest; young - overaged; Rich - poor; Brave - cowardly; Overt - covert; Social - reserved; Quick - slow; Original - imitative; Exhibiting - hiding; Permitting - preventing; anticipate - unexpected; white - black; Leading - following.

The usual scoring procedure involves assigning a weight of 7 for a tick in the most favourable position and a weight of 1 for a tick in the most unfavourable position. More detailed discussion on the scoring aspects is presented in a separate section below. (Reb. p. 3.30)

7. **SOBIO-METRIC ANALYSIS**

The size of the class has been suggested by the headmasters and teacher educators as an important variable affecting classroom climate. (Supra. P.4). Further discussion with colleagues and teacher educators brought out the fact that it is not the mere size of the class that matters, rather the socio-metric status or composition of the class that really
matters. Accepting this view, the investigator has attempted to study the structure of the class.

"In simple terms, a socio-metric measure is a means of assessing the attractions, or attractions and repulsions, within a group. It usually involves each member of the group privately specifying a number of other persons in the group with whom he would like to engage in some particular activity and further, a number of persons with whom he would not like to participate in the activity."¹

Accordingly the students in the class were asked to mention the names three other members of the class with whom the subject would like to work during "Group Study". This theme was particularly considered significant in the context of academic work of the class.

Suitable socio-gram was prepared for each class, using the first choice only. It was possible to identify and count the number of subgroups and their sizes. Based on these results, it was planned to develop a class homogeneity index using the formula

\[
\text{Class homogeneity index} = \frac{100 \times N}{n \times N}
\]

Where \( N \) is the size of the class.
\( n \) is the number of subgroups.
\( M \) is the size of the largest subgroup.

¹ Lindzey, G. and Byrne. op.cit. 455
The rationale behind the formula may be briefly summarised as follows. The index, as defined here, decreases as the number of sub-group increases. Thus a group which subdivides into two subgroups will show a higher numerical index than a group which splits into 3 or more subgroups. Secondly, for any fixed value of \( n \) (Number of subgroups), the index will increase with the size of the largest sub-groups. Thus the index takes into account both the number of subgroups and their size. It also takes into account the size of the original group. If a group is completely homogeneous, there will be no subgroups. In other words 'n' will be equal to 1 and the value of \( K \) will be the same as \( N \). So the index will be 100. This index is simple by definition and also easy to compute. For fairly large groups, the index tends to the value \( 100/n \) (i.e.) 50 in case the group splits into 2 subgroups, 33.3 in the case of 3 subgroups and so on.¹

8. OTHER VARIABLES

An assessment of the physical conditions of the class was made by observing the conditions of cleanliness, blackboard, chalk and duster, seating, facilities for teacher, disturbances and distractions and lighting and ventilation. A few others questions to get at the bio-data of the teachers and the content taught were also added, to give a finish to the work.¹


(In the article under reference, the index is termed Class Cohesiveness Index; the terminology, class homogeneity index, as used in this thesis has been suggested as the latter is more meaningful)
9. **ORGANISATION OF DATA COLLECTION**

Before launching the pilot study, it was necessary to arrange the various variables in proper form and order. This kind of organisation is necessary for efficient and smooth collection of data.

There are three kinds of data to collected: (1) data that are to be collected through systematic observation by the investigator (2) data that are to be collected from teachers teaching the classes and (3) data that are to be collected from the pupils in the class. Therefore two tools were got ready in the printed form to collect data from teachers and from pupils. For the sake of convenient reference, they will be referred to as questionnaires.

The questionnaire to the teacher consists of the following sections. Section A: Personal data - getting information about age, sex, experience, qualification etc. of the teacher; Section B: About the lesson; Section C: Teacher's perception of the class; Section D: Attitude towards pupils; and Section E: Attitude towards the profession..

The questionnaire to the pupils includes questions regarding name, age, socio-economic family group etc; a question intended to study socio-metric structure of the class; a question regarding what he has learnt during the lesson and lastly a rating scale to study the pupil perception of the teacher.
Copies of these two tools are presented in Appendices 3 and 6.

10. PILOT STUDY

(a) Observation of Classes:

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 (a) to find out the feasibility and the adequacy of the tools of data collection (b) to refine and if necessary, modify them and (c) to get better insight of the problems of data collection, planning the analysis and getting at results. Adequate copies of the questionnaires were got ready and the investigator observed eleven classes. This sample includes (a) teachers of both the sexes (b) Teachers of varying experience (c) teachers teaching different subjects and (d) teaching different standards IX to XI. This sample was so drawn as to include two schools managed by private bodies and two schools run by the Government. Details of the sample are given in Table 7.

With the permission of the head of the institution and the teacher, the investigator sat in the last row of seats in the class. While the teacher began his lesson and warmed up, the investigator noted his observations regarding the physical conditions prevailing in the class. This was followed by the systematic observation of the teacher at work, using the modified version of Flanders interaction Analysis Categories. One coded record was made every three seconds. This process was continued till the end of the period.
Soon after the period, the teacher was requested to answer the questionnaire and the attitude scale items. Immediately after that the pupils in the class were suitably instructed and the student - questionnaire was administered to the whole class. While both the students and the teacher were answering, the investigator prepared the seating chart. All the doubts raised by the pupils or the teacher regarding the questionnaire were cleared then and there.

Thus two periods, each of 45 minutes duration, were made use of: one for the observation of the teacher at work and the other for the administration of the questionnaire to the teacher and the students.

The teachers were not informed of the visit of the investigator beforehand so as to prevent their making any special preparation or arrangement. The objective was to observe the class in as natural as setting as possible.

(b) **Finalising the Attitude Scales**:

Simultaneous with this pilot study, copies of the attitude scales were mailed to a number of teachers. The aim was to get adequate data to conduct item - analysis and finalise the attitude scales. About 500 copies of the printed attitude scales were mailed and 405 returned. The item - analysis was finally based on 400 returns from 36 schools. An outline of the procedure adopted for item-analysis has already been presented. Table 7 gives details regarding the schools and the teachers who responded to this part of the study.
11. **ANALYSIS OF THE DATA - PILOT STUDY**

The data were then analysed to yield several indices. The following paragraphs give details regarding the procedure adopted in the analysis.

(a) **Analysis of verbal behaviour:**

The coded records of observation made during the lesson were counted to yield frequency counts. Based on the frequency counts two indices were calculated as follows: The D/I ratio (Direct vs Indirect influence) for each teacher (or class) was calculated by dividing total of the frequencies in categories 5, 6, 7, 12, and 14 by the total of frequencies in categories 1, 2, 3, 4 and 10. This is based on the assumption that the effect of categories 5, 6, 7, 12 and 14 is more direct in nature and the effect of the categories 1, 2, 3, 4 and 10 is indirect.

Flanders et. al. have suggested the reciprocal of this ratio as I/D because theoretically a teacher who is able to bring about a lot of indirect influence to play in the class is considered superior. However several studies made in India, have shown that a typical teacher is more direct in his approach than indirect. Uday Pareek and Venkateswara Rao\(^1\) have given the following distribution for the I/D ratio, based upon 147 observations.

---

I/D ratio: 0-.50, .51-1.00, 1.01-1.50, 1.51-2.00, 2.01-18.50

Percentage frequency: 67.34, 21.77, 6.12, 2.72, 2.04

Clearly the distribution shows a high clustering in the range 0 to .50. It was, therefore, felt that by taking the D/I ratio the clustering at the lower end could be resolved to a certain extent.

Similarly a S/T ratio was calculated by dividing the total frequencies in categories 8, 9, 11, 13 and 15 denoting student behaviour by the total frequencies in categories 1, 2, 3, 4, 5, 6, 7, 10, 12 and 14 denoting teacher behaviour.

A third index, called class control Index, was also calculated by expressing the total frequencies in categories 16 and 17 as a percentage of the total number of observations.

Expressed symbolically,

\[
\frac{\text{D/I}}{\text{S/T}} = \frac{\text{Sum of frequencies in categories 5, 6, 7, 12 & 14}}{\text{Sum of frequencies in categories 1, 2, 3, 4 & 10}}
\]

\[
\text{Class control Index} = \frac{\text{Sum of frequencies in categories 16 & 17}}{\text{Total frequencies}} \times 100
\]

The interpretation of these indices is quite simple. If the D/I ratio equals one, then the class can be considered to be well balanced with respect to direct/indirect influence of the teacher. If D/I is greater than one, the teacher is more direct
in his approach. If the ratio is less than one, the teacher is more indirect. Usually the lesser the value of this ratio, the better.

A similar interpretation can be made of the S/T ratio. If it equals one, then the class time is perhaps shared equally by the teacher and the student. If this ratio is greater than one, then the student behaviour is more prominent than the teacher behaviour. If this ratio is less than one, it may be interpreted that the teacher behaviour dominates over the student behaviour.

Category 16 corresponds to mass response and category 17 is silence and confusion. The more the number of tallies in these two categories, the less the control the teacher has over the class.

(b) **Teacher perception of the class:**

This index was based on the responses made by the teacher to the rating using 24 bi-polar discriptions. There were seven points between the two poles. The favourable end was weighted 1 and the unfavourable end 7. The teacher had marked one position on each of the 24 traits or descriptions. These were converted into weights, added and divided by 24 to yield the mean point. This was taken as an index of the perception the teacher has.

The interpretation of this index is as follows. If a teacher has a highly favourable perception, then he would be consistently rating high on the descriptions and consequently
the mean point will be low. If, on the other hand, the teacher had an unfavourable perception, the mean point or index will be high—near to 7.*

(c) Teacher attitude towards the pupil and the profession:

The responses made by the teachers to the attitude statements were scored according to the scheme already explained. (Supra I.49). Two totals were struck separately to yield (i) a score indicative of the teacher's attitude towards the profession and (ii) his attitude towards the pupils.

(d) Pupil—background:

In order to understand the background of the pupils, it was decided to study (i) the economic conditions of the family and (ii) the educational background of the family. Though the expert opinion obtained through the questionnaire did not include these variables, the investigator felt that it may throw light on certain aspects of the problem under investigation. This hunch was based upon the experience and comments often made by some teachers that they are not able to reach higher standards of attainment or do better in the class because the pupils come from such poor families.

Each student provided information regarding the monthly income of the family on a six point scale. The responses were weighted as follows:

* It would have been better if the weights had been taken in the reverse order—7 for the favourable end and 1 for the unfavourable end. This error was realised after the scoring. Hence suitable adjustment was made, when the raw scores were converted into Z—scores. (See Page 37).
<table>
<thead>
<tr>
<th>Family Income</th>
<th>Less than Rs.100</th>
<th>Rs.100 to Rs.249</th>
<th>Rs.250 to Rs.499</th>
<th>Rs.500 to Rs.749</th>
<th>Rs.750 to Rs.999</th>
<th>Above Rs.999</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

The index for the class was determined by finding the mean weight for the entire class.

Similarly, the information provided by the student regarding the highest educational qualification amongst the members of the family was weighted as follows and the index for the class was the mean weight of the class.

<table>
<thead>
<tr>
<th>Qualification</th>
<th>Less than Std.VIII</th>
<th>Std.VIII to Std.X</th>
<th>Std.X to XI</th>
<th>XI to Dip.</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

(e) **Homogeneity Index**:

This index was based upon the sociogram of the class as explained earlier (Supra 1.23).

(f) **Grasping capacity**:

Each student was asked to write out what he had learnt during the lesson observed. This was compared with what the teacher, in his questionnaire, had mentioned as taught. The responses were scored on a three point scale as follows:
2 points: if the response is quite clear and detailed.

1 point: if the response indicates that the pupil had been attentive, but lacks clarity and details.

0 point: if the response indicates inattentiveness and lacks clarity and detail.

One score was struck for each student and the mean was calculated for the class.

(e) Pupil perception of the teacher:

This index was based upon the ratings made by the pupils using the 14 pairs of bipolar adjectives/descriptions. Weights were assigned to the seven possible rating positions as indicated under (b). The class index was obtained by adding the weights over the entire class and all the 14 descriptions and finding the mean. The procedure was similar to that followed in the case of teacher perception of the class.

(h) Physical conditions of the class:

The physical conditions of the class was assessed on the seven aspects: (i) cleanliness, (ii) blackboard, (iii) chalk and duster, (iv) seating, (v) teacher facilities, (vi) disturbance and distractions and (vii) lighting and ventilation. The conditions obtained in the class were rated on a three point scale and the total point was taken as an index describing the physical conditions of the class.
To summarise, the responses made by the teachers and the pupils to the various items in the respective questionnaires and the observation records were quantified to yield 12 measurements or indices for each class, as described below:

i. Family Economic Background (of pupils)
ii. Family Education Background (of pupils)
iii. Homogeneity Index
iv. Grasping capacity
v. Pupil perception of the teacher.
vi. Teacher attitude toward profession
vii. Teacher attitude toward pupils
viii. D/I ratio
ix. S/T ratio
x. Class control
xi. Teacher perception of the class.

If a class could get high indices on all the twelve traits, it was assumed that the conditions reflected acceptable or good classroom climate. Under this assumption, it was, therefore the next step, to compare the indices.

Comparisons based on raw scores are misleading because all the 12 measurements mentioned above have not been made on the same scale. The arbitrary weights would have acted in directions unknown and hence it was felt that some kind of conversion to a common scale was necessary before making comparisons.
Garrett. H. E.¹ has suggested that it is more meaningful to compare standard scores rather than raw score. "Deviations from the mean expressed in \( \sigma \) - terms are called \( \sigma \) - scores or Z - scores or Standard Scores .... \( \sigma \) - scores are often small decimal fractions and to avoid this difficulty they are usually converted to another distribution with mean \( M' \) and \( \sigma - 1 \)."...

The shift from raw to standard score requires a linear transformation of the form

\[
\frac{X' - M'}{\sigma'} = \frac{X - M}{\sigma}
\]

and such a linear transformation does not change the shape of the distribution in any way.

Using this idea, the raw indices were transformed to a standard score with mean = 4 and \( \sigma = 1 \) by using the formula

\[
Z - 4 = \frac{X - M}{\sigma}
\]

(i.e) \( Z = 4 + \frac{X - M}{\sigma} \)

The mean and standard deviations were calculated for each of the twelve variables separately, using the 11 classes. The complete details of the various measurements are presented in Tables 8 and 9. In the case of pupil perception of the teacher and the teacher perception of the class, the formula was

\[
Z = 4 - \frac{X - M}{\sigma}
\]

since lower values indicate more favourable disposition.

Attempts were made to read through the values of the 'Z' score and interpret the classroom climate. No one class was consistently scoring high on all the 12 variables but still it was not difficult to locate classes with deficiencies. For example, class where the teacher domination was extra-ordinary (Class No. 9), class which showed poor grasping power (Class No. 6) and so on could be easily located.

A thought occurred at this point, that some sort of profile may be developed to express the conditions of the class graphically. Two kinds were tried one using parallel lines and other using 12 equally spaced radii of a circle; when the graphs were presented to a few teachers for comments, the latter attracted their attention. Samples of both the graphs are presented in Appendix H.

Studying further the graphical representation it became clear that if a class had scored high on all the 12 variables, then their plots will be away from the centre of the circle. And so the polygonal area will be large. If, on the other hand, the class has a poor climate with low values on the twelve variables, then the corresponding points will be nearer to the centre and the polygonal area will be low. This appeared to be a good and easy hint to interpret the results.

Another hint also occurred for the quantifying the concept. The idea was: why not use the area of the polygon as an index of classroom climate?
It is proved in elementary trigonometry that the area of a triangle with two sides \( b, c \) and the angle between them \( A \) is given by the formula

\[
\text{Area} = \frac{1}{2} bc \sin A.
\]

In the graph, we have twelve triangles, all with a common vertex \( O \) and with same vertex angle. Therefore, it is easily deduced that the area of the polygonal region equals

\[
\frac{1}{2} z_1 z_2 \sin A + \frac{1}{2} z_2 z_3 \sin A + \ldots + \frac{1}{2} z_{11} z_{12} \sin A + \frac{1}{2} z_{12} z_1 \sin A
\]

\[
= \frac{1}{2} \sin A \left( z_1 z_2 + z_2 z_3 + \ldots + z_{11} z_{12} + z_{12} z_1 \right)
\]

Since \( \frac{1}{2} \sin A \) is a constant, we can drop it conveniently.

The sum of the cross products (i.e.)

\[
z_1 z_2 + z_2 z_3 + \ldots + z_{11} z_{12} + z_{12} z_1
\]

can be taken as the classroom climate Index.

The last line in Table 9 gives the classroom climate Index for the eleven classes observed. It is now clearly seen that Class No. 11 has the best climate. It is a small class, handled by a capable teacher, with balanced teacher-pupil interaction, fairly homogeneous and very good grasping capacity, compared with other classes.

Thus the results of the pilot study could indicate that the tools were fairly good for the purpose of the study and that it is possible to carry out further analysis. The pilot study has also led to evolving of a single numerical index indicative of the classroom climate.
12. VIEWS AND COMMENTS OF EXPERTS

A brief report outlining the research, procedure and results was prepared and copies of the report were sent to a handful of other experts in the field. The aim was to get at their comments about the procedures adopted. The comments received were highly useful in several respects. Most of the experts agreed with the use of systematic observation technique to study classroom climate. Two contradictory suggestions were also received - one suggesting the use of multivariate analysis and factor analysis and the other advising the investigator to keep away from it. One expert suggested to reduce the number of variables. The use of socio-metric data was commended and a suggestion was made to use all the three choices instead of the first choice alone. The circular graphs were appreciated and the experts pointed out that a more detailed explanation of the rationale behind the various formula could have been included in the report.

The one aspect which did not meet with the approval of the experts was the technique used for the study of Teacher perception of the class and pupil perception of the teacher. The use of the particular set of bipolar adjectives and the process of summation over the class were not accepted.

Also, the inclusion of Economic background, Education background, Physical conditions as variables classroom climate was
not well received on the ground that they do not form any 'Scale'. Further, they were based on the reports made by the pupils which may not be objective and reliable.

Based on these comments, it was decided to revise the collection of data suitably and to reduce the number of variables.*

13. **REVISIONS MADE IN COLLECTION OF DATA**

It was a well accepted fact that teacher-pupil relationship and mutual perception was an important aspect to be included in the study. If the use of bipolar adjectives was not well conceived, then what other method could be appropriate? This question could not be answered immediately in a satisfactory way. Hence the investigator had to do some more review of studies already made seeking a solution to the problem.

In an exploratory study at the Midwest Psychological Field Station,¹ instead of requiring students to talk about their teachers, the investigators invited students to tell about their classes. These students were asked to identify as 'plus', the several classes "which were especially good, satisfying, worth while". They were then asked to explain for each: "What I did, what went on, what made class good, 'plus', what I enjoyed or got out of the classes". Similar questions were used for 'minus' classes. This open method invited the subject to reveal the salient aspects of the classroom situations as experienced by him.


* The Investigator gratefully acknowledges the help rendered by the experts in this context.
The student responses fall into five categories: the teacher (his qualities and behaviours), the subject matter, the classroom activity, the classmates (their qualities and behaviours) and the working conditions. The responses were distributed approximately as follow: Teacher - 25%, Subject matter - 47%, Activity - 23%, Classmates - 3.5% and working conditions - 1.5%.

Taking a clue from this study, the investigator observed a class and after the class, the students were asked to discuss and react to the lesson. Of course, the teacher who taught the lesson was not present and that made the students free to express both 'plus' and 'minus' aspects. From this, it was possible to formulate a list of questions to which the students can respond and show their reactions to the class and the teacher. A list of 50 such questions was formulated and they covered the same aspects - teacher, subject matter, activity, classmates and working conditions almost in the same proportion as found in Midwest study. It was also thought appropriate that the same questions (after a little editing and revision) could be given to the teacher to react. If the reactions of the teacher and the pupils could show considerable agreement it is a sign of closer ties between the teacher and the pupils and it is a healthy sign. If there is dis-agreement then it can be interpreted to mean that there is divergence between the perceptions of the teacher and the pupils.

In the light of this experience, it was decided that the ratings using the bipolar adjectives be deleted and in its place,
the 50 questions evolved from classroom discussion of students should be used instead. All the questions were of the 'Yes' or 'No' type.

The criticism that the family economic status, the education status and the physical conditions obtaining in the class were not objectively measured was considered valid and hence it was decided not to use them as an aspect of classroom climate. With these modifications, the tools for collection of data were finalised.

14. THE FINAL STUDY

The experiences gained during the pilot study and the experts' reactions on the report of the pilot study have helped in deciding the procedure for the final study. As Paul V. Gump remarks: the inter dependency of multiple parts (Classroom elements) implies that utilization of system variables, as opposed to controlled single variables, may provide a feasible and fruitful approach.\(^1\) It is also worth noting that the variables selected for the present investigation are in close agreement with some of the aspects of the seven variable model for teacher - effectiveness suggested by Bruce J. Biddle.\(^2\) He has, in the seven - variable model suggested the following variables: (i) Formative experiences (ii) Teacher properties (iii) Teacher behaviour (iv) Immediate effects (v) Long-term consequences (vi) Classroom conditions (vii) School and Community contexts.

\(^1\) Paul. V. Gump. (1964). op.cit. 183

\(^2\) Bruce. J. Biddle (1964). The Integration of Teacher Effectiveness Research, in Biddle and Ellena (Eds.) op.cit. 6-18
The present investigation (in the final study) has included the following:

1. **General Information**: Class, school and subject - providing general description of the sample.

2. **Teacher variables**: Sex, age, experience and qualification - these are what Biddle classifies as formative experiences - information obtained through the self-report of the teachers.

   Attitude towards pupils, towards profession and Teacher-pupil agreement. These are aspects which Biddle classifies as 'Teacher properties'.

3. **Class variables**: Strength of the class, number of sub-groups, size of the largest sub-group, grasping level, liking for the subject, liking for the teacher, homogeneity index - These may be termed 'Class conditions'.

4. **Teaching Variables**: D/I ratio and T/S ratio - These are outcomes of the 'teacher behaviour'.

In order to gather data on these variables, three major tools are made use of: (1) Systematic observation of the classroom using modified form of Flanders Interaction Analysis Categories; (2) A questionnaire to the teacher which consists of six sections:

   (a) Introductory letter (b) Personal data (c) About the lesson, (d) Reactions to the lesson (e) Attitude towards pupils and (f) Attitude towards the profession; and (3) A questionnaire to students which consists of four sections: (a) personal data, liking for subject matter, liking for teacher (b) A section where the pupil records what he has learnt in the lesson, (c) A question for studying the socio-metric structure of the class and (d) fifty questions eliciting reactions to the lesson.
The copies of the tools used in the final study are given in Appendices I and J.

15. **The Sample**

In many research studies, it is customary to take a sample consisting of a small collections of the objects under study and study those units. The aggregate of objects about which the study is made is called the Population. Every individual object is called the Sampling unit. In order to arrive at valid conclusions, the researcher has to follow certain well laid procedures and select the 'sample' for his study. As Snedecor points there are two problems involved in this process: "to design and conduct his sampling so that it shall be representative of the population; then, having studied the sample, to make correct inferences about the sampled population."¹

The present investigation is concerned with classrooms and thus the sampling unit is the class. But in order to have a manageable, economic, representative sample, it is necessary to consider the different strata which are of significance. 'Classrooms' are part of 'Schools' and in Tamil Nadu we have different types of schools. Obviously to include all teachers and all classes in a single school will also be not desirable. In such cases, a sampling in two stages will be appropriate.

In Coimbatore area, there are different types of schools which can be classified as follows: (a) Government High Schools (b) Private Schools and (c) Municipal Schools. There are also separate schools for boys and girls. Many schools have co-educational classes. In some big schools they admit both boys and girls but run parallel sections. Another factor to be considered was the sex of the teacher. While girls' schools employ only lady teachers, it is not uncommon to see lady teachers in other schools teaching boys' class. Thus in order to have a 'Representative' sample it was necessary to take into consideration all the above facts.

Yet it was not the end of the matter. There was another problem - psychological in nature. The study involves observation of live classes. And in order to observe the teacher and the class in as natural a setting as possible, it was decided not to give advance information to the teacher. Therein comes the willingness of the teacher to permit the investigator. Only such classes where the teacher was willing to co-operate could be observed. This psychological consideration is of primary importance for the success of getting data, as otherwise the teacher may adjust his lesson to please the investigator.

Taking all these facts into account and also the factors of time and cost, it was decided to observe at least a hundred classes. At the first stage of sampling, the schools were selected
and necessary permission was obtained from the head of the institution. At the second stage of sampling, the selection was made regarding classes to be observed. This was dependent on the time-table to be followed in the school in the day of visit, the presence (or absence) of teachers and also their willingness to co-operate. Teachers who were hesitant were left out.

It was also thought feasible to observe equal number of men and women teachers. English, Tamil, Mathematics, Science and History/Geography are the five core subjects studied by all the students in the high schools. It was decided to observe equal number of classes in each subject discipline. In the reorganised pattern of education in this state, the elective subjects will be abolished in the next year (1977) and so these classes were dropped.

Thus, it was possible to observe a sample of 106 classes. On grounds of reliability 6 classes were rejected and the data collected from a hundred classes were utilised for statistical analysis. Table 10 presents a description of the sample according to the various strata.

16. **COLLECTION OF DATA**

The collection of data for the final study was done personally by the investigator who visited all the 106 classes. In order to establish reliability, a second observer also participated in the data collection. The helper was trained
in observation by practice observations in a local school along with the investigator. After adequate training, he accompanied the investigator in actual data collection. The reliability of observation was estimated by the comparison of the coded observation. This is fully discussed under a separate heading in the next section.

As it happened in the pilot study, it took two periods to observe and gather data for any particular class. The first period was entirely devoted to the observation of the class while the teacher was teaching. During the next period, the teacher and the pupils answered the questionnaires meant for them. Each period was of 40-45 minutes duration approximately.

Naturally data collection was spread over a period of time, because of (i) a fairly long time required for observation and administration of the questionnaires and (ii) intervening holidays and examinations. The collection of data started some time in October 1975 and ended in March 1976.

17. RELIABILITY

Reliability refers to the consistency of measurements. In respect of tests, reliability has two somewhat different connotations. First, it refers to the extent to which the test is internally consistent. In other words, the question is how accurately does the test measure at a particular time. Second, reliability refers to the extent to which a measuring device
yields consistent results upon testing and retesting. Both these are important and several methods of estimating the reliability of tests have been suggested, especially when the final criterion is available as a numerical score. In the present investigation, systematic observation has been used to collect some basic data, using well defined categories of observation. The best available method of establishing the reliability of observations in such cases is to use Scott's co-efficient. It is customary that this co-efficient is calculated by comparing the observations recorded independently by two observers. Most of the classes included in this study were observed along with another trained observer and the reliability could therefore be obtained in such cases.

The formula for computing the co-efficient is as follows:

\[ \text{Scott's co-efficient} \quad \Pi = \frac{P_o - P_e}{100 - P_e} \]

Where \( P_o \) = percentage of agreement between the observers. 
\( P_e \) = Percentage of chance agreement.

An illustrative example, giving numerical details of calculation is presented in Appendix K. Table 11 gives the reliability co-efficient for the various classes. A scrutiny of the table will

reveal that the reliability co-efficients are quite high (above 0.7) in most cases) and therefore the observations are reliable. Since the observations are reliable, it is fair to assume that the ratios T/S and D/I based on these observations are also reliable.

Two attitude scales have been developed for the purpose of this study. Each scale consists of 25 items selected after item analysis. The reliability of the attitude measure was established by using the split-half method. All items with odd serial numbers formed one part and the rest formed the second part. The scores for each individual teacher were obtained for each part separately and the Pearson's Product Moment correlation was calculated. The obtained correlation was used in the Spearman Brown formula to obtain a more accurate estimate of the reliability. The formula is

\[ r = \frac{2r'}{1 - r'} \]

where \( r' \) is the first uncorrected reliability and \( r \) is the corrected reliability.

The reliability co-efficients were 0.83 for both the scales. Judged by the range of values usually obtained in such scales, these values are quite high enough to accept the scales as reliable.

Tables 12 and 13 present the two scattergrams and show the numerical details of the calculation of reliabilities.

18. **VALIDITY**

Validity of a measuring device refers to the fact how far it measures what it is designed to measure. Hence validity is concerned with a more fundamental idea.

Face validity is a term used to characterize materials that appear to measure what the author desires to measure. Evidently this is a weak form of validation.

Content validity is estimated on the basis of the content of the measuring device. This may be achieved by a built-in process as has been done in the present investigation. The variables included in the study were selected on the basis of expressed comments from experts. Also item-total correlations used in the construction of attitude scales, enhance the validity of the measures used.

Concurrent validation refers to the process of validating obtained results with the results already obtained in other similar studies or well established facts. Such agreements will be pointed out during discussion of results.

As Freeman points out: "The ultimate decision on the criteria of validity in any area of testing rests upon the analytical judgements and agreement of qualified specialists, who evaluate the test's objectives and the groups for whom it is intended."\(^1\) Thus it is difficult to think any measuring instrument as completely valid.

\(^{1}\) Freeman (1968). op.cit. 117
19. PLAN OF ANALYSIS

In order to have clarity and control over the whole process, a careful planning is necessary. Statistical methods are useful in summarising numerical data. In this investigation, several statistical techniques are made use of. They range from simple classification and counting (as done with the categories of observation) to correlations, testing of hypothesis and other similar techniques.

There are several variables involved in the study. It is planned to organise them as follows. Six major variables, as detailed below, are considered as primary variables or aspects of classroom climate.

Variable 1 \((V_1)\) Teacher Attitude towards the Pupil-Attitude scale.

Variable 2 \((V_2)\) Teacher Attitude towards the Profession - Attitude scale.

Variable 3 \((V_3)\) Class Homogeneity Index - Sociometric Status study.

Variable 4 \((V_4)\) T/S ratio - derived from Systematic observation.

Variable 5 \((V_5)\) D/I ratio - derived from Systematic observation.

Variable 6 \((V_6)\) Teacher-pupil Agreement - derived from the responses to 50 questions.

An attempt is made to study classroom climate in terms of these six primary variables.
There are other secondary variables; these include such facts as sex, age, experience, qualification of the teacher, size, structure and composition of the class, type of school, the subject matter of the lesson and so on. Some believe in controlling these variables and some use them for making differential and correlational studies. The latter approach is used in the present investigation.

An attempt is made in this investigation to develop two processes or models of calculating a climate index and compare them. For convenience they will be referred to as the Gross Product Model and the Additive Model. Complete details regarding these models are presented in the appropriate chapters.

20. SUMMARY

The aim of this chapter has been to present a detailed account of the procedure followed in the investigation. The several important stages have been the following.

1. Based on the review of literature and studies conducted in India, a list of pertinent variables was prepared.

2. The list was circulated to a group of teachers, headmasters and faculty members of training colleges. Based on the reactions of 120 respondents, the following ten variables were selected for study. (a) Teacher-pupil relationship (b) Teaching procedure followed (c) Teacher attitude towards pupils (d) Teacher
attitude towards profession  
(f) Teacher's capacity to control the class  
(g) Freedom to children  
(h) Size of the class  
(i) Pupils' capacity to grasp what is taught and  
(j) Pupils' attitude towards teacher.

3. For the sake of better organisation and convenience of data collection these, along with a few other aspects, were classified to fall into four groups  
(a) Variables pertaining to the Teacher: sex, age, professional qualifications, attitudes, perception of class etc.  
(b) Variables pertaining to the class: the economic and educational background of the families from which the pupils come, the class composition in terms of socio-metric data  
(c) Variables pertaining to interactions: including the teacher behaviour and the pupil behaviour in the class and  
(d) other variables such as physical conditions of the class and so on.

4. The following tools for collection of data were also finalised.  
(a) Systematic observation of live classes for studying the interaction variables  
(b) Two attitude scales to study the attitude of teachers towards (i) the pupils and (ii) the profession  
(c) Rating scales using bipolar descriptions to get at the teacher perception and the student perception of the class  
(d) A questionnaire to teachers and  
(e) A questionnaire to students.
5. In order to refine the tools and to ascertain the feasibility of the study, a pilot study was conducted by observing eleven classes. The data gathered in the pilot study were analysed to develop a classroom climate Index. Simultaneous with this observation, the attitude scales were administered to a large sample of 400 teachers in 36 schools. On the basis of the responses made by these teachers, suitable item-analysis was made to evolve the final attitude scales.

6. The report of the pilot study based on the observation of eleven class was circulated to a handful of expert research workers for critical comments. Based on the comments received from them, suitable changes were made in the selection of variables and their measurements. Aspects like economic conditions, educational family background, physical conditions of the classroom were not given importance, as aspects of classroom climate. The rating scales using bipolar adjectives was completely thrown out and instead of a fifty item question list eliciting the reactions of the teacher and the pupils was finalised.

7. The final study was conducted by observing a stratified random sample of 106 classes. The live classes were observed using the modified form of Flanders Interaction Analysis categories. Immediately after observation, the teacher and the
pupils answered a questionnaire which included questions regarding personal data, attitude measurement, reaction to the class and study of the socio-metric status of the class.

8. The reliability and validity of the various measures were also established. The inter-observer reliability for most of the classes was over 0.70. The attitude scales show a reliability of 0.83 by the split-half method.

The results of the study are presented in the next chapter.

###