RESEARCH TOOLS

The tools used for collection of data for this study have been briefly discussed in the foregoing chapter. Attitude towards Science developed by Dr. Avinash Grewall to measure the attitude of secondary school students towards Science. The achievement test in Science constructed by the investigator to measure the achievement in Science of secondary school students. The present chapter is devoted to a discussion of the methods of their construction.

4.1 Science Attitude Scale

Thurston has defined attitude as the degree of positive or negative effect associated with some psychological object. A psychological object, according to him, may be a person, an institution, a religion, a community, an idea, a system, a political party, or a minority community. Under the new curriculum Science will be compulsory subject up to high school stage. One of the objectives of teaching Science is to inculcate scientific attitudes among the pupils.

The purpose of this scale would be to know whether or not the students have developed favourable attitudes towards Science as a discipline. The underlying assumption being that one of the outcomes of Science education is the development of positive attitude towards the subject.

4.1.1 Science Attitude

The Science attitude has operationally defined as a generalized attitude towards the universe of Science content and being measured in terms of its favourableness or unfavourableness estimated from the scores obtained by the subjects on an attitude scale toward Science comprising of the four categories from

the universe of content 'Science Attitude', (i) positive intellectual (ii) negative intellectual, (iii) positive emotional and (iv) negative emotional attitudes.

The Science Attitude Scale is a dependable tool for measuring student's attitude towards Science. It appears to be useful for teachers of Science, guidance workers and research scholars. It can be used by the curriculum specialists to measure the outcomes of teaching Science. The students of Psychology and Education can also use it to study the development of their attitude towards Science.

4.1.2 Try out

The investigator left 20 items after dropping the least discriminating items. The scale was then administered on 515 higher secondary students drawn from 6 schools of Bhopal. This was done with a view to determine the reliability, validity and norms of the scale.

4.1.3 Scoring

Each of the ten positive items (S. Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20) of the scale are assigned a weight ranging from four (Strongly Agree) to zero (Strongly Disagree). In the case of ten negative items (S. Nos. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19) the scale scoring is reversed ranging from zero (Strongly Agree) to four (Strongly Disagree). The attitude score of a subject is the sum total of scores on all the twenty items of the scale. For each student a total score on the scale can be obtained by summating his scores for the individual items. Thus a maximum of 80 scores can be obtained by a subject. However, the administration of the test reveals that the scores ranged from 25 to 70.

The present scale is a 5 point scale consisting of 10 positive statement item (at even serial no.) and 10 negative items (at odd serial no.). The following table shows the scoring.

Response Options	Weightage given to each item	
	Positive Statement	**Negative Statement**
Strongly agree (SA)	4	0
Agree (A)	3	1
Undecided (U)	2	2
Disagree (D)	1	3
Strongly Disagree (SD)	0	4

The sum total of all items given the quick measures of attitude towards Science at the respondents.

4.1.4 Reliability of SAS

The reliability of the Science Attitude Scale (SAS) was estimated by the split-half (0.86) and test-retest (0.75) methods which was found to be quite satisfactory. This compares favourably with reliability (0.765) found by Sood (1975) for his scale of attitude towards Science and scientists. Reliability of the scale was further checked by two methods of scoring by administering the scale to a small sample of 50 subjects with instructions to check the statements in accordance with the usual Thrustone's instructions and the Science subjects were then asked to check for each item on one of the five alternatives in accordance with the usual Likert instructions. The coefficient of correlation found between the scores on two scales was 0.94.

4.1.5 Validity

The SAS appears to have content validity and the method of selecting items supports his supposition. In addition, difference in mean scores were found among the selected groups of known preference for Science i.e. Arts (Mean = 46.41) and Science (Mean = 50.58) students which is highly significant (t=6.62) at 1 percent level.

4.1.6 Time for administration

The SAS is a self-reporting inventory consisting of 20-items designed to assess the attitude of individuals towards Science. There is no time limit but normally in takes about 5 minutes to explain the test and the subjects require about 15 minutes for giving responses to the items of the scale. The new scores can be converted into percentile norms which have been derived from the data collected on the norming population.

4.1.7 Usefulness

The SAS is a dependable tool for measuring student's attitude towards Science. It appears to be useful for teachers of Science, guidance workers and research scholars. It can be used by the curriculum specialists to measure the outcomes of teaching Science. The students of Psychology and Education can also use it to study the development of their attitude towards Science.

4.2 Science Achievement Test

Some mechanism has always existed for assessing the worth of an individual in accordance with the needs of society. Even in ancient Greece in Sparta, tests were devised to measure the physical competence of youthful pupils (Schwartz, et al. 1962). In 225 B.C, Chinese were selecting civil servants through examinations (Schwartz, et al. 1962). The modern testing methods can, however, be traced to the begging of 20^{th} century. In contemporary times, annual examination continues to be the most widely used system of evaluation.

Generally speaking, achievement implies the net result of an individual's effort over a period of time. In the case of academic achievement, it is a combination of knowledge and skills which a child acquires on going through a process of formal

instructions. Academic achievement is related to the objectives of syllabus content to be covered during a full term. (Deale 1975).

Achievement tests generally measure the present proficiency, mastery and the understanding of general and specific areas of knowledge. Largely, they are measures of effectiveness of instruction and learning. The achievement tests can be broadly classified into two categories.

i. Standardized Achievement Test.

ii. Specially Constructed Achievement Tests.

Standardized tests published group of tests that are based on general educational content common to a large number of educational systems, while as especially constructed tests are teacher-made tests devised by teachers to measure limited and specific achievements. Such tests are also constructed by educational researchers for measuring limited areas of achievement or proficiency like the present Science Achievement Test.

4.2.1.1 Purpose of Science achievement test

Achievement tests serve a number of purposes. They are used to evaluate teachers' effectiveness, the effectiveness of different teaching methods, maintaining the school standards and in making surveys of pupil's performance. These tests also provide information for classification and placement of individuals in relatively homogenous groups for the purpose of differentiated instruction. These tests enable teachers and counsellors diagnose each pupil's strength and weaknesses. Achievement tests thus enable us to know the progress of pupils in a specific area of work or the whole performance in learning process of subjects for a particular period of time. In other words as put forth by Garrett (1959) "The purpose of educational

achievement tests-like that of ordinary school examination is to discover how much a pupil knows about the subject he has studied or is studying."

Different types of tests can be put to use to serve the purpose like Essay type tests, objective type tests or performance tests. But the investigator is more concerned with the objective type tests. An objective type test is a systematic procedure in which the individual tested is presented with a set of constructed stimuli to which he/she responds. These responses enable the examiner to assign a numeral or a set of numerals from which inferences can be made about examinees, possession of whatever the test is supposed to measure. There are a number of objective tests which can be put forth in the following classification.

4.2.1.2 Objective type tests

For many generations' pupils, progress was measured through the utilization of teacher made tests and often subjectively marked examination of its essay type. But the psychologists of Nineteenth century made commendable strides in the development of techniques for measuring the results of sensory simulations. These efforts reached their peak in the establishment of laboratory for the experimental study of psychology in the University of Leipzig by Wilhelm Wundt. However, little was done until the end of Nineteenth Century in formulating or attempting to formulate the standardized and objective techniques of evaluating the outcomes of teaching. The contribution of J.M. Rice in the construction of objective type tests proved a milestone "Although there was a much criticism on Rice's work among psychologists and Schoolmen. He persisted in his studies of testing techniques in various school subjects. His point of view and his attempts at test construction become a starting point in the development of objective testing technique as an aid to

improved teaching methods." (Crow & Crow, 1984). Objective type tests are composed of items or questions on learners progress in,

1. Subject matter mastery.
2. In the power to understand the significance of and to apply the material learned.
3. In the growth of attitude and social competence and
4. Above all in learning objectives put forth earlier.

The investigator has selected only one subject Science syllabus for test construction. There are many stages in the process of test development. To construct an achievement test the following steps are followed. These steps are presented in the form of a flow chart explained subsequently.

4.2.1.3 Planning of the Science achievement test, for Class-IX

Science Achievement Test for class-IX was designed to measure knowledge, comprehension and application skills in Science.

Keeping in view the importance of analysis of subject in the construction of Science achievement test, the investigator first reviewed the text books of Science for classes VIII prescribed by NCERT and U.P. board. The subject matter included in these text books was analysed thoroughly and an outline of the content was prepared. It was presumed that students entering into any class have the concepts and mastery of the content learned in the previous class. So students of class IX are expected to have sufficient knowledge and understanding of the Science they have studied in class VIII.

The investigator decided to construct multiple-choice objective type items because these types of items have several additional merits as compared to other

types of objective questions. Multiple-choice type of items can be used to measure more complex outcomes like comprehension and application along with knowledge objective. Murry (1938) regarded multiple choice type as one of the best measure of test judgment that is available, Hawkes & Lindquist (1936) too considered it definitely superior to other types of items for measuring educational objectives. Cronbach & Murwin (1960) observed that multiple choice items have deserved popularity as an aid in assessing achievement, ability and personality.

4.2.1.4 Content of the Tests

Keeping in view the importance of analysis of subject in the construction of Science achievement test, the investigator first reviewed the text books of Science for classes VIII prescribed by Utter Pradesh State Board of School Education and Central Board of Secondary Education. The subject matter included in these text books were analyzed thoroughly and a blue print of the content was prepared. It was presumed that students entering into any class have the concepts and mastery of the content learned in the previous class. So students of class IX are expected to have sufficient knowledge and understanding of the Science they have studied in class VIII.

4.2.1.5 Flow Chart of a Blue Print for SAT

There were many stages in the process of test development. To constructs an achievement test the following steps were followed. These steps were presented in the form of a flow chart explained subsequently.

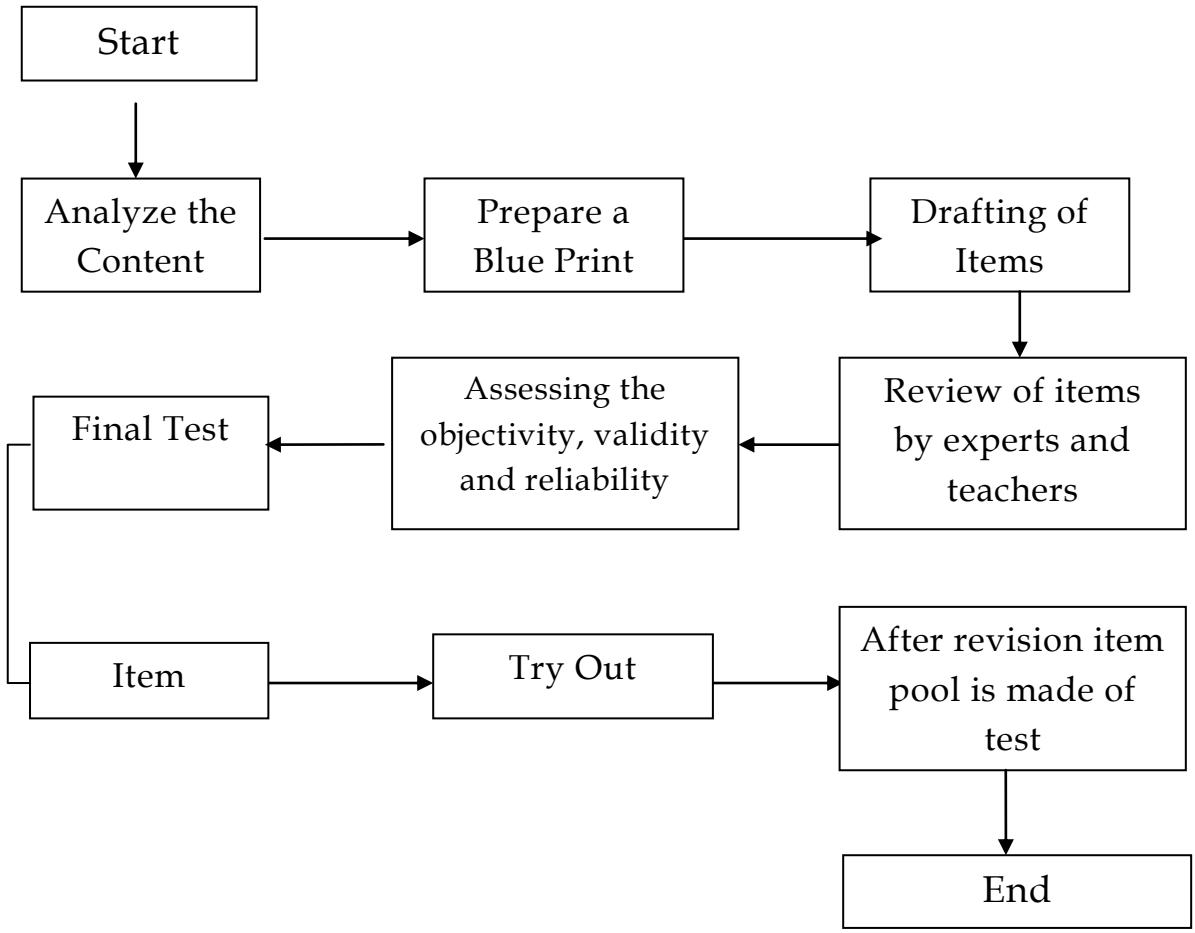

Fig 4.1 Flow Chart of a Blue Print of SAT

The whole course content of the subject of Science was the main focus in the test construction. The SAT is based on nearly 90% course content of Science already mentioned on the ground that the test administration was scheduled at the beginning of the session in order to ensure that all respondents are in a position to attempt it. The detail of content and its learning objectives for the subject was placed in the blue print prepared.

4.2.1.6 Objectives of the Science achievement test

The purpose of testing is to determine the extent to which the pre-determined objectives are fulfilled. Thus, the first step in measuring achievement is to establish a clear statement of objectives. Hence, the researcher was concerned with the educational objectives, both which are related to total process of education and those

which are related to the subject. Each item of the test was related to one or other of the educational objectives. The test items were constructed keeping in view the weight-age as regards to objectives for the levels of learning knowledge, comprehension and application. Test on the subject was based on the syllabus prescribed by Utter Pradesh State Board School Education and Central Board of Secondary Education.

The objective of the achievement test in Science is to test the knowledge, comprehension and application level of the respondents. Before the construction of items, an outline of the whole test items in the form of a chart called 'Blue Print' was prepared for Class-IX. A well prepared blue print (Singh, 1996) was consulted while framing the items. A 'Blue Print' of Science achievement test was prepared on the basis of nearly 90% course content in Science for Class IX. This blue print helped the investigator to construct the achievement test taking into consideration the major educational objectives like knowledge, comprehension and application. The items in test were constructed in the light of these objectives.

4.2.1.7 Form of Items/Questions

The SAT consists of multiple choice items only. Every item was in form of a statement either affirmative or interrogative or a simple statement called the stem of the item. Every item was followed by four alternatives or choices. Among the four given choices, only one was correct and the rest three choices were distracters.

4.2.1.8 No of items

The achievement test in Science for class IX students that is used in the present study was constructed by the investigator. This is very comprehensive test based 16 chapters of class VIII Science text book U.P. Board and CBSE Board. The test, at the initial stage consists of 140 items of multiple choice types, representing achievement at

various units of Science, such as 60 items in physics, 40 items in chemistry and 40 items in biology. The test was based on the latest syllabus prescribed by the NCERT and U.P. board.

From the item pool, initially 140 items for science achievement test for class IX were drafted and discussed with subject teachers. These were also shown to two experts in Science. Based on their constructive criticism improvement was made in the items. The language was also checked by three experts in language. Similarly, the objectives of the SAT were to test the knowledge, comprehension and application of learning material. A blue print of SAT for class IX was prepared. In the light of the blue print, items were constructed for SAT for class IX.

The draft form thus prepared was released for expert's opinions that were requested to judge the worth of each question against the following criteria.

- Appropriateness of content
- Accuracy of the scoring key
- Consistency of the text items
- Avoiding undesirable over lapping
- Accuracy of language

As a result of experts comments some of the questions were modified and some omitted. The revised version of the initial Science achievement test contained 100 items.

4.2.1.9 Duration of the test

The time allotted for attempting 100 items was 120 minutes.

4.2.1.10 Codification of answer-booklets

All the answer booklets were coded from 001 to 300 avoiding any confusion.

4.2.2 Try out

The initial form of Science Achievement Test (SAT) was administered on 300 class 9^{th} students (150 male 150 female) randomly sampled from five-secondary schools, out of which two were girls, two were boys and one was coeducational schools located in Barabanki district. These schools belong to different categories of management and range from good to poor in regard to standards of performance of their students. Thus, the sample selected for tryout of the SAT constituted cross-sections of secondary school students. After permission was obtained from the principals of the schools, the sheets were distributed to the students of these schools. From the 300 students in the five schools 300 completed sheets were received responding a 100% response rate.

4.2.3 Scoring of the test paper

The investigator scored the answer-booklets according to the scoring key prepared in advance for question booklet. Since, all the items in the test are multiple choice items the scoring process seemed quite easy and done comfortably and objectively. All the items on the test were assigned equal mark i.e. 'One' mark for each correct response. In case of a wrong response 'Zero' mark was given. Negative marking procedure had not been adopted in the marking system. When all the three hundred answer sheets were scored, they were arranged in descending order of scores for item analysis. The Science achievement score was the sum total of scores on all the 100 items of the test. Theoretically the range of scores on this test extended from zero to 100.

4.2.4 Item analysis

The objective of the item analysis was to obtain objective information concerning the items written for the test. "This information is valuable for several reasons. It provides an opportunity to check upon the test writers subjective judgments in selecting the items to compose the test." (Guilford, 1987). The test

maker gained new insights into the test construction and learned where and how item need to be rewritten. The most common use of item analysis was to enable the writer to modify the test in the direction he wanted. It provides a method of eliminating certain items on the basis of discrimination Index (DI) and difficulty value (D.V.) or facility Index (FI). "It is powerful tool for the improvement and accumulating a bank of high quality test items." (Ebel & Frisbile, 1991).

4.2.4.1 Item analysis procedure

For the present Science Achievement Test item analysis procedure was taken from " Ebel & Frisbile (1991)" Item analysis begins after the test had been scored. The scored test papers were arranged from high achievers to low achievers. The top 27% and the lowest 27% scored papers were taken for item analysis from the groups upon which test were administered. For each individual item the number of examinee's and their responses were counted separately for higher and lower groups, respectively. In order to estimate the index of item difficulty the following formula was adopted as given by Singh (1996).

Difficulty value $(D_1) = \frac{H+L}{2N} \times 100$

Where, H = No. of correct responses in high scoring groups.

L = No. of correct responses in the low scoring groups.

N = No. of examinees in the two groups.

Similarly, in the estimation of Index of Discrimination, the following formula was adopted.

$(D_1) = \frac{H-L}{N} \times 100$

Where, H = No. of correct responses in higher scoring group.

L = No. of correct responses in the low scoring groups.

N = No. of examinees in higher or lower groups.

4.2.4.2 Difficulty value or facility index

Difficulty value is the index which describes the percentage of students who attempted a particular test item correctly. Higher the difficulty value or 'facility Index', easier the item is considered. A test item to be valid from difficulty point of view, it should be in the range of 25% to 85%. An item possessing the difficulty value below 25% or above 85% normally gets rejected. (Singh, 1996).

4.2.4.3 Discrimination index

The quality of an item in distinguishing (discriminating) between good and poor students in technically called 'Discrimination'. The statistics showing discrimination value is called 'discrimination index'. The indices range from -1.0 to +1.0. If the criterion or the group changes the discrimination indices for the some question may change. (Natarajan, 1977).

The method of calculating Discrimination Index followed for the final selection of the test formats has been shown above in the formula. It is simpler to compute and to explain to others than indices of discrimination as point bi-serial correlation, Flangan's coefficient (Flangan, 1939) and Davis's co-efficient (Davis, 1996). It had very useful property, which most of other correlation indices lack of being biased in favour of items of middle difficulty. As we have already seen, "it is precisely these items that provide the largest amounts of information about differences in the levels of achievement and thus contribute most to score reliability". (Ebel & Frisbile, 1991).

The efforts were made to select only best items i.e. the item having high discriminating value in the final format. The items possessing discrimination index below 0.35 were out rightly rejected.

4.2.4.4 Item analysis chart

The two dimensional charts were prepared in order to analyze each item. In these charts the numbers of questions are written horizontally and the names or the codes of examinees are written vertically. In this chart, it could be seen that a student had done a particular test item correct or not. We put a mark one (1) for a correct answer, a zero (0) for wrong answer and a dash (-) for omitted question in the relevant box. Then the correct responses were added both horizontally and vertically. This way it was ascertained as to how many examinees attempt the item on the test correctly and secondly the responses of each individual item were counted for further analysis. In this manner, we got two different totals of responses from higher group and lower group. Lastly, the formula of difficulty value and discrimination index was applied on the scores reflected by these item analysis charts.

For the present analysis only 27% respondents acquiring the high marks called higher group and 27% low achievers called lower group were considered separately for class IXth. The rest of respondents called middle group is not considered as they belong to average category.

4.2.4.5 Rejection of item

Initially the test book-let contained 100 items of Physics, Chemistry and Biology in Science for class VIIIth. After the item analysis was conducted upon the test, it was found that some items owing to difficulty value need to be eliminated. The problem was discussed with the supervisor who unanimously gave the instruction to eliminate these items which do not come under the required limits of discrimination index and difficulty value. Thus 25 items from the Science achievement test got eliminated after the item analysis was completed. Ultimately the final test format got reduced to 75 items.

Table 4.1: Distribution of items included in the final test according to the educational objectives

S.N.	Topic	Knowledge	Comprehension	Application	Total
1.	Combustion and Flame	2	3	1	6
2.	Force and Pressure	0	0	7	7
3.	Friction	0	3	1	4
4.	Sound	1	0	3	4
5.	Chemical and Magnetic Effects of Electric Current	3	0	0	3
6.	Light	2	4	3	9
7.	Stars and Solar Systems	2	0	0	2
8.	Metals and Non Metals	4	5	1	10
9.	Coal and Petroleum	1	2	3	6
10.	Synthetic Fibres and Plastics	1	2	1	4
11.	Crop Production and Management	1	3	0	4
12.	Microorganism : Friend and Feo	1	1	1	3
13.	Cell Structure and Functions	1	2	0	3
14.	Reproduction in Animals	1	1	1	3
15	Reaching the age of Adolescence	2	2	0	4
16.	Blood Structure: Blood Group and Blood Transfusion.	2	0	1	3
	Total	**24**	**28**	**23**	**75**

4.2.5 Reliability

Reliability refers to the consistency of scores obtained by the same person when re-examined with same test on different occasions or with different sets of equivalent items or under other variable examining conditions (Anastasi, 1982) "A test is said to be reliable to the degree that it measures accurately and consistently, yielding comparable results when administered a number of items", (Best & Kahn, 1992). Usually, contrary to person's, level of achievement, the physical quantities measured (like, length mass & time etc.) are quite stable and accounts all leading to measurement of high reliability. The same may not be true with achievement tests because measuring devices or tools are not so precise and accurate. However, the best and most convenient method of computing reliability of a teacher made test is the Split half method. Since, repeated exposure to a test is not advisable and developing parallel forms is a cumbersome. The researcher resorted to this method to compute reliability of Science achievement test for class IX^{th}.

Reliability is the most fundamental quality which any measuring instrument should possess. In order to find as to what extent the SAT developed by the investigator possesses this quality, split half method was employed. A split-half reliability coefficient was found by correlating scores of the subjects on odd items of the form with their scores on even items. There are several ways of splitting the scores into two, each with some merits and some demerits, but the most common method used by test constructors is odd-even method, specifically when the items are assumed to be homogeneous, which is the case in the present case. The reliability was calculated by using the scores of 300 pupils on the 75 items of the final Science Achievement Test. The formula used for estimating reliability was a simplified of general Kuder-Rechardsan formula. The estimate of reliability by this formula is 0.87.

4.2.6 Validity

The validity of a test concerns what the test measures and how well it does so (Anastasi, 1982). “Content validity is a matter of determining whether the sample is representative of larger universe it is supposed to represent”. (Gronlund & Norman,1968). There are three types of validity: content, concurrent and construct validity. In an achievement test content validity is the most important criterion. Content validity show how adequately the test samples the universe of knowledge, attitudes and skills a student's is expected to master. Content validity a built to a test from the outset through the choice appropriate items (Anastasi, 1982). In general, a test is valid if it measures what it claims to measure. It is that quality of data gathering instruments that enables it to measure what it is supposed to measure. (Best & Kahn, 1992). Besides, an achievement test has content validity if it represents faithfully the objectives set forth in the blue print. (Lindeman, 1967). Content validity of achievement test in Science was based on careful examination of course text books, syllabus and objectives and on the basis of judgment of three subject matter specialists.

The major types of validity are content validity and construct validity. Content validity is based on a careful comparison of the items to the definition of the domain being measured (Allen & Reyan, 1979). Content validity of the items was ensured through rational logical analysis of the Science teachers and experts in test construction. A copy of the test is given in Appendix-II.