CHAPTER - V

DEVELOPMENT OF THE TOOLS

CHAPTER - V

DEVELOPMENT OF THE TOOLS

5.1 INTRODUCTION

The present investigation aimed at measuring instructional effects of Advance Organizer Model, namely, development of conceptual structures and meaningful assimilation of information andideas; and nurturant effects, namely, interest in inquiry and habits of precise thinking. Apart from measurement of these effects, the study also attempted to assess the students' retention of meaningful assimilation of information and ideas and their reactions towards Advance Organizer Model. All these required measurement devices which could assess students responses pertaining to the above aspects. But, ready-made tools to measure the above aspects except reactions toward AOM were not available. Because measurement of instructional effects and retention of meaningful learning is content specific. Measurement of nurturant effects has not been attempted by the earlier researches as it is evident from the review of literature presented in the Chapter III. Although, tools measuring reactions of students towards AOM have been developed by the researchers earlier, the investigator decided to develop a separate tool on this aspect. As a result, the following tools were developed by the researcher.

i) Tests on Conceptual Structures.

ii) Tests on Meaningful Assimilation of Information and Ideas.

iii) Test on Interest in Inquiry

iv) Test on Habits of Precise Thinking

v) Retention Test on Meaningful Assimilation of Information and Ideas

vi) Reaction Scale to Study Students' Reactions Towards Advance Organizer Model

The procedures of development of the above mentioned tools are presented one by one in the coming sections.

5.2 TESTS ON CONCEPTUAL STRUCTURES

i) Rationale

Development of conceptual structures has been pointed out by Joyce and Weil (1980) as one of the instructional effects of Advance Organizer Model. While describing the model, they have discussed organising information, the structure of the discipline and cognitive structure in the context of conceptual structures. An analysis of Ausubel's model reveals that there is a parallel between the way subject matter is organised and the way people organize knowledge in their mind, their cognitive structures. Ausubel (1963) also expresses the view that each of the discipline has a structure of concepts (and/or proposition) that are

organised hierarchically. That is, at the top of each discipline are a number of very broad, abstract concepts that include the more concrete concepts at lower stages of organisation. He also believes that the conceptual structure of the subject matter should be taught to the students to facilitate meaningful verbal learning.

Therefore, development of conceptual structures is a direct outcome of the model. While deciding for the evaluation of such development, one has to opt for a testing procedure. The testing procedure given by Novak et al.(1981) has been opted for the measurement of conceptual structures in the present study. Novak, alongwith others developed a tool called concept mapping in 1972 to represent student's knowledge structures before and after instruction (Novak, 1990). This technique is based on Ausubel's theory of meaningful verbal learning which requires a conscious effort on the part of the student to relate new knowledge to knowledge previously acquired. Concept mapping is a pedagogical technique to help students see explicitly how new concepts can be related to previously learned concepts (Novak et al.,1981). They have provided guidelines both for construction and scoring of a concept map/concept structure.

ii) Construction of a Concept Map

Ault (1985) has given the following steps for construction of a concept map:

1) Select an item for mapping. This could be an important text, passage, lecture notes, or laboratory background material.
2) Choose and underline key words or phrases; include objects and events in the list.
3) Rank the list of concepts from the abstract and inclusive to the most concrete and specific.
4) Cluster the concepts according to two criteria: (a) Concepts that function at a similar level of abstraction, and (b) concepts that interrelate closely.
5) Arrange the concepts as a two-dimensional array analogous to a road map.
6) Link related concepts with lines and label each line in propositional or prepositional form.

Novak et al. (1981) have also emphasised the above steps while constructing a concept map. However, they stressed on selection of shorter text or passage so that concept map does not become too large and contain too many concepts.

iii) Development of the Tests

Instructional effect on conceptual structures of the model required the students to develop conceptual structures or concept maps on the concepts assimilated by them during the instruction. It, therefore, necessitated to develop measurement device to evaluate students' concept maps on

the concepts assimilated by them. While doing so, the testing procedures namely, unit testing and comprehensive testing were adopted. It is because unit testing and comprehensive testing would provide good measures for the validation of the model in terms of the instructional effect i.e. Development of conceptual structures. As a result, one test for each of the four units and a comprehensive test were developed. A few broad and subordinate concepts from a unit were selected to constitute the passage which would be given to the students as the conceptual structure test. Based on this passage, students were expected to develop a concept map. In this way, four unit tests on the four units were developed.

While developing the passages, the steps for construction of concept map given by Ault (1985) were taken into consideration. This was done to ensure that all the components required to construct a good concept map are contained in the passages. However, passages were not made complex in terms of concepts and propositions as it was thought that students might face difficulty to develop maps if passages were complex. After the passages were developed, they were given to five experts to examine if the concepts and propositions were presented in terms of hierarchy and other features of a concept map. Their comments and suggestions were taken into account to modify the passages. In this way, content validation of the tests was ensured.

While developing the final test on conceptual structures, two passages, namely A and B were developed from two units out of four units on a random basis. This was done as students might face difficulty if they are asked to develop four concept maps on four units in the comprehensive test. However, while developing the two passages for the comprehensive test, they were made longer and complex in comparison to the passages used for individual units. The procedure followed to develop passages for the four units was adhered to while developing the two passages for the comprehensive test. Both the passages for the four unit tests and the passages for the comprehensive test to measure conceptual structures are given in Appendices A & B respectively.

iv) Scoring Procedure

The scoring procedure followed by Novak et al. (1981) to score a concept map or conceptual structure was adopted to score the concept maps developed by the students. The scheme of scoring procedure is presented in Table 5.1.

TABLE 5.1 : Scheme of Scoring a Concept Map

Sl.No.	Elements	Scoring Procedures
1	Relationships	One point is given for each relationship between two concepts provided the relationship is content correct and explicitly stated. No additional credit is awarded for duplication of the same relationship on the concept map.
2	Hierarchy	Points are awarded depending on the degree of hierarchy in the concept map. The number of points given for hierarchy depends upon the number of levels that are identified in in the constructed map. Use the map to the right for illustration. One point is given for at least one correct relationship per level up until two levels beyond the branching if the map remains linear.
3	Branching	The branching of the concept map refers to the level of degree of differentiation among the concepts that are illustrated

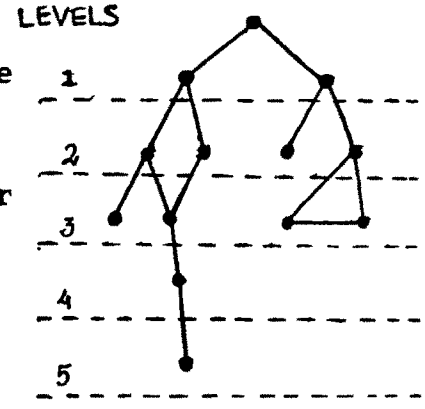

Table 5.1 contd.

Sl.No.	Elements	Scoring Procedures
		in the hierarchy. That is, it attempts to rate the degree that specific concepts are connected to more general or inclusive concepts. That rating is as follows: One point for the first branching where two or more concepts are connected to the concept above. Three points for any subsequent branching where there is an example of two or more concepts connected to a concept above. The illustrated map above would receive a score of seven: 1 point for level 1, and three points each for levels 2 and 3. Note that since no branching occurs in levels 4 and 5, no further points are awarded for this criterion.
4	General to specific	The concept map receives an additional rating for illustrating a general to specific pattern. Whether one concept is more general than the one below it depends upon the line which connects the two concepts. If no general to specific relationships exist, or less than 10% of the relationships are general to specific, the map receives a score of zero.

Table 5.1 contd.

Sl.No.	Elements	Scoring Procedures
		if 10-20 % are correct = 1 point 30-40 % are correct = 2 points 50-60 % are correct = 3 points 70-80% are correct = 4 points 90-100% are correct = 5 points
5	Cross links	:Inter relatedness in a student's concept map indicates an integration of concepts, and is depicted as cross links on the concept map. Cross links show a relationship between concepts on one branch of the hierarchy with concepts on another branch. Notice the two examples of cross links in the illustration above. A rating of one point is given for each cross link showing the integration among concepts. Two additional points are awarded for duplication of the same cross link, that is, showing the same integration of concepts.

Source: Novak et al.(1981), page V-9.

While scoring students' concept maps the first four elements of the concept map, namely, relationship, hierarchy, branching, and general to specific were taken into account. The fifth element, namely, cross links, was not considered as it was thought that students of class VIII might find difficulty in establishing cross links on one branch of the hierarchy with concepts on another branch.

5.3 TESTS ON MEANINGFUL ASSIMILATION OF INFORMATION AND IDEAS

i) Rationale

Meaningful assimilation of information and ideas has been mentioned as an instructional effect of the Advance Organizer Model of teaching by Joyce and Weil (1972, 1980). While evaluating such an effect, one is required to judge, in the learners, the meaningful assimilation of information and ideas transacted during the instructional process. The concept of meaningful assimilation of information and ideas has been clarified in Chapter-II. However, to put it precisely, it refers to acquisition and internalisation of information and ideas in a meaningful manner. Such a process of acquisition and internalisation of information and ideas require certain cognitive abilities on the part of the learners, namely, knowing and understanding the information and ideas which correspond to knowledge and comprehension categories under the cognitive domain of the Taxonomy of Educational Objectives given by Bloom (1956). This Taxonomy comprises six categories of cognitive abilities, namely, knowledge, comprehension, application, analysis, synthesis and application. The first two categories constitute lower levels of cognitive abilities. Most of the advance organizer studies have taken the first two levels of the cognitive domain to assess the meaningful learning (Healy, 1989). As a result, two levels of cognitive

domain, namely, knowledge and comprehension were taken to constitute the components of the test on meaningful assimilation of information and ideas.

ii) Criterion Referenced Testing

While evaluating the above instructional effect, the procedures of criterion-referenced testing and comprehensive testing were followed instead of following only achievement testing at the end of the experimentation. Such an approach to testing was followed due to certain reasons. Firstly, the above instructional effect required the measurement of all the information and ideas assimilated by the students during the instructional treatment. Criterion-referenced testing in the form of unit testing could achieve this objective. Secondly, criterion referenced testing as well as comprehensive testing provide good measures for the validation of the model in terms of the above instructional effect. Such a type of evaluation procedure has been followed by many researchers to validate the instructional strategies developed by them (Biswal, 1980; Menon, 1984; and [illegible] Vardhini, 1983). With the above justifications, criterion - referenced tests in the form of unit tests and the comprehensive test had to be developed. As per the requirement of the study, four unit tests and one comprehensive test had to be developed. Development of four unit tests and one comprehensive test required content analysis of the units and formulation of terminal behaviours.

iii) Content Analysis and Formulation of Terminal Behaviours

Before formulation of terminal behaviours, contents of the four units had to be analysed. This was done to identify the main information and ideas which were to be taught to the students. Based on these ideas and information, terminal behaviours pertaining to knowledge and comprehension categories were formulated. Definite terms have been suggested by Bloom (1956) to indicate these levels of behaviour. These terms were followed by the investigator to state the terminal behaviours under knowledge and comprehension categories. The list of terminal behaviours, in respect of each unit, thus arrived at, are presented in Appendix C.

iv) Development of Unit Tests

Unit tests had to constitute evaluation items related to each of the terminal behaviours representing information and ideas to be taught to the students. Because such items aimed at evaluating whether information and ideas had been assimilated by the students. Therefore, suitable evaluation items were developed to cover all the terminal behaviours under the four units. All these items covered knowledge and comprehension categories. The items were of mainly objective types, namely, multiple-choice, fill-in-blanks, true-false and one-word/phrase answer. The items framed in this way for all the unit tests were edited by the researcher from the point of view of language suitability and ambiguity. The content

validity of the tests was obtained by subjecting them to five experts for scrutiny who studied the testsin terms of their content coverage and terminal behaviours comprising knowledge and comprehension categories. On the basis of their suggestions, modifications were done to the items in the tests. Arrangement of items was done to place similar types of items together. Each set of items was preceded by specific guidelines as to how to respond to the item. The tests were used as power tests. So, students were allowed to take their own time to complete the tests. However, all the students were able to complete the items of the tests within the school period of 40 minutes. While scoring the items, one mark was given to each correct answer and zero to wrong answer. A hand made scoring key for each of the tests was prepared to score the items. Unit tests along with their respective scoring keys are given in Appendix D . The details regarding item types and objectives in respect of each unit test are presented in Table 5.2.

v) Development of Comprehensive Test

In addition to the unit tests, a comprehensive test on meaningful assimilation of information and ideas had to be developed.

TABLE 5.2 : Distribution of Items in Terms of Types and Objectives for the Unit Tests on Meaningful Assimilation of Information and Ideas

Units	Types of Items					Objectives		
	Multiple choice	Fill in Blanks	True False	One word or phrase	Total	Knowledge	Comp-rehension	Total
1. National Integration	10	5	4	4	23	10	13	23
2. India's Defence	17	5	6	6	34	21	13	34
3. Foreign policy of India	6	3	3	3	15	8	7	15
4. The United Nations	21	6	9	2	38	27	11	38
Total	54	19	22	15	110	66	44	110

This test, being a comprehensive test, each and every terminal behaviour could not be included in it. However, a sample of terminal behaviours from each unit were represented in this test. This form of sampling became a necessity to ensure the usability of the test. A table of specifications was prepared to represent the sample of terminal behaviours pertaining to knowledge and comprehension categories as well as the item types. While developing items for the sample of terminal behaviours chosen for the comprehensive test, a few items had to be retained from the unit tests and the rest of the items had to be constructed afresh. This was done with a view to not

repeating all the items from the unit tests which might have placed the students in a comfortable position while answering the items. The items developed for the test were of multiple-choice, fill-in-blanks, true-false and one word or phrase answer types. The items were scrutinised and edited by the researcher. Items, thus developed, were given to five experts for judging their content coverage and representation to two categories of terminal behaviours. On the basis of the expert's comments and suggestions, items were modified and re-written. In this way, the content validity of the test was obtained. Arrangement of the items was done to place similar types of items together. Each set of items was preceded by specific instructions as to how to respond to the items. The test was used as a power test. Hence, students were allowed to take their own time to complete the test. However, most of the students were able to complete the test within one period of 40 minutes. Of course a few students took a few more minutes to complete the test. While scoring the test papers, one mark was given to each correct answer and zero to wrong answer. Scoring was done using a hand made scoring key. The details regarding type of items and the objectives are presented in Table 5.3. The comprehensive test and the scording key is given in Appendix E.

TABLE 5.3 : Distribution of Items in terms of Types and Objectives for the Comprehensive Test on Meaningful Assimilation of Information and Ideas

Units	Types of items					Objectives		
	Multiple choice	Fill in blank	True False	One word or phrase	Total	Knowledge	Comprehension	Total
1. National Integration	6	1	2	1	10	3	7	10
2. India's Defence	8	-	2	1	11	7	4	11
3. Foreign policy of India	4	1	1	-	6	2	4	6
4. The United Nations	10	1	2	-	13	9	4	13
Total	28	3	7	2	40	21	19	40

5.4 TEST ON INTEREST IN INQUIRY

i) Rationale

Development of interest in inquiry has been mentioned as a nurturant effect of Advance Organizer Model of teaching (Joyce and Weil, 1980). This nurturant effect has two aspects - one is the act of inquiry and the other is the interest in the inquiry act. The act of inquiry has two connotations. One is related to someone's inquisitiveness or curiosity to know things and the other is related to

someone's ability to solve problems through a systematic process involving sensing the problem, formulating hypothesis and verifying hypothesis after collecting and analysing evidences. While both connotations are emphasised in teaching situations, the latter is more used as a teaching - learning strategy. An interest in inquiry implies someone's likings or dislikings in the inquiry act.

An analysis of the Advance Organizer Model of teaching shows that it promotes students' inquiry ability by way of encouraging them to initiate their own questions, to recognise the difference between hypothesis and fact, to find out meaning for their own questions, etc. which constitute inquiry as a problem solving act. Ausubel et al.(1978) also emphasize , "That existing cognitive structure plays a key role in problem solving is evident from the fact that the solution of any given problem involves a reorganization of the residue of past experiences so as to fit the particular requirements of the current problem situation". Hence, while operationalising this nurturant effect, inquiry as a problem solving act was taken into consideration. As a result, three major steps of the inquiry act, namely, sensing the problem, formulating hypothesis and verifying the hypothesis were taken for the test construction. Students' likings and dislikings in these three steps of inquiry formed the test on interest in inquiry. The three steps of inquiry constituted the three components of the test towards which students' interest had to be studied.

ii) Format and Nature of Items

Since the subject of teaching for the present study was Civics, students' interest in inquiry into social problems had to be studied. The best way of doing this is to study the behaviour representing the construct in students' social life situations. As this involves a number of constraints on the part of the researcher e.g. time to be devoted, number of students involved, the researcher decided to present these social situations in the form of items constituting a test to which students would respond in a written manner. While choosing social situations for item construction, students' own environment was taken into consideration as they are more sensitive to and take interest in social situations which concern them. Therefore students' own home, neighbourhood, school, classroom, peer circle formed the social situations based on which items were constructed.

Each test item consisted of a stem which described a social situation in which students had to exhibit their likings or dislikings. The situations selected were of three types. In first type of situations, students had to show their interest in sensing the problem. In second type of situations, they had to show their interest in formulating the hypothesis about the problem, and in third type of situations they had to evince their interest in verifying the problem.

The items were provided with three possible alternative responses from which the student was required to select one response. These three responses represented three degrees of students' interest. One response depicted that students were 'most interested', the second, 'moderately interested', and the third, 'least interested'. Each stem containing the situtation was carefully constructed so that students did not face any ambiguity in understanding the stem and got biased to a particular response. The test was patterned on the Likert's (1932) method of summated rating as the attribute was composed of various components. However, unlike the Likert method which requires all the items pertain only to a single reference point which is the psychological object, the test under consideration described the reference point as being composed of certain components. Therefore, while judging, a student's interest in inquiry, his/her likings or dislikings in all the components of inquiry had to be taken into account. Unlike the Likert's method where directional statements are given, to which the extent of agreement or disagreement is required to be known, deviation was made to present definite responses in terms of possible behaviours in particular situations. This was preferred because the students of Class VIII might find it difficult to discriminate between various degrees of agreement or disagreement.

iii) Response Pattern

While responding to the test items, the respondents had to read the stem properly, place him or her in that particular situation and think of the solution to the situation in terms of the alternative solutions to the situation given at the end of the stem. Then, he/she had to encircle the number against the alternative to which he/she thought as the best. This involved a mental process which included understanding the situation presented, distinguishing between the alternatives, placing oneself in a particular situation and choosing the best alternative. An example of an item from the component - sensing the problem, is given below. The same pattern was followed for other two components.

iv) Illustrative Item on Sensing the Problem

Your father comes back from his office regularly at 5.00 p.m. but one day he does not reach home on time. Would you then like to:

a) wait till your father comes back from his office.
b) think that something might have happened to him but you cannot do anything about it.
c) think that something might have happened to your father and that you must do something to find out your father.

v) Development of Items

An initial pool of 30 test items were constructed. These items were examined by the researcher himself. Then, the test comprising the pool of items were given to six experts for examination in terms of their validity, ambiguity of the words used, length of the stem, the appropriateness of the situations taken and the alternative solutions to each situation formulated. Four experts returned the test with their comments and suggestions to modify some of the items. However, nobody suggested deletion of any item. Based on their comments and suggestions, all those items requiring modifications were rewritten by the researcher. Thus, a total of 30 items were retained in the test for try-out. The test containing the items was given to an English editor for editing. It was, then, given to two eighth grade students of English medium school, who were average learners, for their comprehensibility. It was found that the test items were comprehensible for them. The distribution of items with respect to the components is shown in Table 5.4.

TABLE 5.4: Distribution of Items with respect to the Components on the Tryout Form of the Test on Interest in Inquiry

	Components	No. of items
1.	Sensing the problem	10
2.	Formulating hypothesis	10
3.	Verifying hypothesis	10
	Total	30

In arranging the items constituting the test, care was taken to see that the items belonging to a particular component were distributed across the test. This step was taken to eliminate the patterned responses which result if all the items belonging to a particular component were concentrated at one place of the test. The test also included the relevant information from the students and the directions to respond to the test items. The format of the test used for the tryout study is given in Appendix F .

vi) Tryout of the Tool

To make a selection from the pool of 30 items, tryout was done. This needed VIIIth standard students of English medium schools in the urban setting other than the school where the experimentation was to be carried out. Therefore, Laxman Public School, New Delhi, where consent could be obtained, was selected for the tryout purpose. A school in New Delhi was selected due to certain reasons. The researcher serves as a teacher in New Delhi. It was, therefore, thought by the researcher that the time to be spent on tryout from the limited time that he could avail of from the institution where he serves would be utilised for the experimentation. However, while doing this, the researcher did not hamper the sanctity of tool construction. The researcher ensured that tryout sample and experimental

sample were equal in relation to medium of instruction, age, settings of the school, etc. Both the schools representing tyrout sample and experimental sample follow English as well as Hindi as the medium of instruction, are situated in urban settings, are affiliated to the C.B.S.E. Students in both the schools belong to different parts of the country, thereby representing the same national culture. The age level of the students in both the situations was also the same.

The sample for the tryout consisted of a total of 72 VIII standard students of Laxman Public School. During the tryout time, students were told about the purpose of the test, namely, to measure their interest in inquiry. Students were placed in a comfortable situation during the tryout. The test was used as a power test. The students were allowed to take their own time to respond to the items. However, on an average, the time taken by students was about 40 minutes.

vii) Scoring Procedure

It has already been pointed out previously that the three responses/alternatives represented three degrees of interest in inquiry, namely, most, moderate, and the least. These three alternatives were assigned weightage of 2, 1, and 0 respectively as these alternatives appeared in each item. As a result, the maximum score of the test was 60 and the minimum score was zero. With the above guidelines evolved

for scoring, the responses were scored and the total score of each respondent was arrived at.

viii) Selection of Items

With a view to selecting items from the present pool of items, item analysis of each item was carried out. For this purpose, the methodology adopted in the Likert's method of summated ratings was used. It is because other methods of item analysis like correlational methods do not yield higher accuracy of selection of items. Murphy and Likert (1934) found that the rank ordering of fifteen statements of the magnitude of difference between the mean of high and low groups agreed very well with the ordering of the same statements in terms of the magnitude of correlation between the item response and the total score. Therefore, the procedure of finding out the difference between the means of high and low groups on the individual item as a basis for selecting items for the scale was adopted. According to this procedure, high and low groups had to be formed on the basis of the total score. Hence, all the 72 test scores of the students were arranged in a descending order. From this distribution, the upper 25 per cent and the lower 25 per cent were selected to constitute the high and low groups respectively. For each item, the 't' value was calculated using the standard 't' formula. The obtained 't' values were grouped according to the components to which

they belonged and placed in descending order of 't' values. The 't' values arranged in this order are presented in Table 5.5. While selecting test items, the 't' values above 0.01 level of significance were taken into account. As a result six items from each component were selected for the test. Though, the component on verifying hypothesis had all the items significant at 0.01 level, only six items had to be taken as six items in each of the other two components were found significant.

ix) The Final Test

As a result of the item analysis carried out for item selection, a total number of 18 items were selected for the final form of the test. There were six items under each component. The final form of the test is given in Appendix G .

TABLE 5.5 : Componentwise 't' values of Test Items

Sl.No.	Component	Item No.	't' Value
1.	Sensing the Problems	17	5.66 *
		1	5.23 *
		22	4.82 *
		15	4.60 *
		8	3.69 *
		27	3.52 *
		20	2.32
		11	2.12
		24	2.00
		5	0.88
2.	Formulating hypothesis	16	6.09 *
		28	5.49 *
		30	4.00 *
		13	3.94 *
		4	3.33 *
		25	3.25 *
		9	2.09
		7	1.76
		18	1.83
		2	1.23
3.	Verifying hypothesis	21	8.71 *
		10	7.10 *
		6	6.89 *
		14	5.93 *
		26	5.22 *
		19	4.97 *
		29	4.58
		3	4.44
		12	4.31
		23	2.89

* Items selected for the final test.

x) Psychometric Properties of the Test

Psychometric properties of the test involves computation of reliability and validity of the test. Reliability and validity of the test are given below:

(a) Reliability

One finds various methods of estimating reliability in the literature. For the present test, the split-half method was opted. It is because, the split-half method is employed when it is not feasible to construct parallel forms of the test nor advisable to repeat the test itself. This situation occurs with many performance tests, as well as with questionnaires and inventories dealing with personality variables, attitudes and interests (Garrett, 1966). In order to estimate reliability through split-half method, the correlation between the two sets of scores of the respondents had to be computed. This was done by adding the scores of the respondents on odd items as well as on even items. Correlation between the scores of the respondents on odd items and on even items was computed by using Pearson's product-moment method of correlation. The correlation value obtained through product-moment method was .43. This value is the reliability of the half test. By using Spearman-Brown prohecy formula, the reliability for the whole test was estimated. Thus, the reliability for the whole test was .60. This reliability is considered to be satisfactory.

(b) Validity

Validity of test always depends on the nature of the test. The present test intends to measure students' interest in inquiry. Interest in inquiry is considered to be a psychological trait. Whenever we wish to interpret test performance in terms of some psychological trait or quality we are concerned with construct validity (Gronlund, 1981). There are different methods of obtaining evidence for construct validation. One of them is to analyse the mental process required by the test items. At the outset, the construct, 'interest in inquiry' was operationalised in terms of its three components by analysing the inquiry act. Then, test items were developed on each of the components. The pool of items were subjected to experts' judgement. On the basis of their comments and suggestions, the items were modified and rewritten. The modified items were tried out on a sample and the final test items were selected to constitute the test on interest in inquiry. In this way, the construct validity of the test was established.

5.5 TEST ON HABITS OF PRECISE THINKING

i) Rationale

Joyce and Weil (1980) have hypothesised on development of habits of precise thinking as a nurturant effect of the Advance Organizer Model of teaching. But, they have not

given any operational definition of the concept 'precise thinking'. However, an analysis of the syntax of the model of teaching, particularly at the phase three, reflects that the authors advocate development of critical approach to subject matter or knowledge. According to them, a critical approach to knowledge is fostered by asking students to recognise assumptions or inferences that may have been made in the learning material, to judge and challenge the assumptions and inferences, and to reconcile contradictions among them. Ausubel (1968) also emphasised an independent and critical approach to the understanding of subject matter. To him, this involves, in part, the encouragement of motivations for and self-critical attitudes towards acquiring precise and integrated meanings, as well as the use of other techniques directed towards the same end. Therefore, Ausubel mentions about precise thinking as acquiring precise and integrated meanings.

According to Ausubel (1968, 1985), precise and integrated understandings are, presumably, more likely to develop if :

1. the central unifying ideas of discipline are learned before more concepts and information introduced;
2. the limiting conditions of general developmental readiness are observed;
3. precise and accurate definition is stressed, and emphasis is placed on delineating similarities and differences between related concepts;

4. learners are required to reformulate new propositions in their own words.

All of these devices come under the heading of pedagogic techniques that promote an active type of meaningful reception learning. Teachers can help foster the related objective of assimilating subject matter critically by encouraging students to recognize and challenge the assumptions underlying new propositions, and to distinguish between facts and hypothesis and between warranted and unwarranted inferences (Ausubel, 1968, 1985).

So, based on Ausubel's definition of precise thinking, components for the test on habits of precise thinking were selected. Five components constituted the test on habits of precise thinking. These are:

1. Distinguishing between facts and hypothesis;
2. Delineating similarities and differences between related concepts;
3. Distinguishing between warranted and unwarranted inferences ;
4. Identifying the central idea of theme ; and
5. Recognising the assumption underlying a new proposition.

ii) Format and Nature of Test Items

Evaluation of habits of precise thinking which is a nurturant effect, required measurement device to be content free. However, the measurement device should be able to measure the attribute in content related area. Therefore, the researcher, while selecting themes to construct items for the various components of the test, chose contents pertaining to different branches of social studies as well as social events which students are acquainted with. The nature of the test items varied from component to component as each of the components demanded different kinds of test items. The nature of the test items under each component is discussed in the coming sections.

Component-1 : Distinguishing Between Facts and Hypotheses

Under this component, the students had to distinguish between facts and hypotheses. Hence, a few statements of facts and a few statements of hypotheses were developed. Statements of facts described the events, ideas which exist or have happened in the past and statements of hypothesis described a hypothetical proposition which can be proved or disproved.

Component-2 : Delineating Similarities and Differences Between Related Concepts

Students had to delineate similarities and differences between related concepts under this component. Therefore,

pairs of related concepts were taken and three similarities and one difference between them were pointed out. These three similarities and one difference constituted the four alternatives to the two concepts. Hence, two concepts and the four alternatives constituted a multiple-choice type item. While presenting the four alternatives, care was taken to present the alternative i.e. the difference between the two concepts, in the positive form so that it would have equal distraction power as other three alternatives had.

Component-3 : Distinguishing Between Warranted and Unwarranted Inferences

This component required the students to distinguish between warranted and unwarranted inferences. Therefore, two passages from a text book prescribed for eighth grade students - one dealing with ancient Takhshila University and another dealing with Einstein's visit to India were selected. These passages were analysed and a few statements were written on the basis of analysis. Some of these statements were directly inferred from the information given in the paragraph and therefore, called true or warranted inferences. Some of them were indirectly inferred from the information given in the passage and hence, called false or unwarranted inferences. About the rest of the statements, no inferences could be drawn from the paragraph, and thus, called 'No inference' statements. All the statements were presented in mixed up manner at the end

of the passages.

Component-4 : Identifying the Central Ideas of a Theme

Under this component, students had to identify the central idea of a theme. Therefore, two passages from a text book prescribed for eighth grade students - one dealing with Ramakrishna Paramhansa's attitude towards religion and the other describing the power of writing were selected. The central ideas of the passages were identified. Two more ideas close to the central idea were also identified as distracters. These two ideas alongwith the central idea of a passage were presented at the end of the passages in the form of multiple-choice type item.

Component-5 : Recognising the Underlying Assumption Behind a Proposition

This component required the students to recognise the underlying assumption behind a proposition. Therefore, a few propositional statements related to social events were developed. One assumption and three non-assumptions related to the proposition were also formulated. These four - one assumption and three non-assumptions, comprised the four alternatives to the proposition. These four alternatives had equal distraction power. The items in this section were presented in the multiple-choice type form.

iii) Response Pattern

Response pattern in the test differed from component to component due to difference in the nature of items under different components. (a) Under Component-1, students had to write (F) against the statement which is a fact and (H) against the statement which is a hypothesis. (b) They had to put a circle around the alternative which points out the difference between the related concepts under the Component-2. (c) Component-3 required the students to write (T) for the statement which was true inference, (F) for the statement which was false inference and (N) for no inference. (d) Students under Component-4 had to tick mark the alternative which described the central idea of a theme (e) They had to encircle the alternative which was an assumption to the proposition under the Component-5. All the components contained the necessary guidelines for the students to respond to the items.

iv) Development of Test Items

After the nature and response pattern of items to be developed clarified, an initial pool of 63 test items were written. These items were scrutinised and edited by the researcher. The test containing the pool of items were given to eight experts to examine the contents of the items in terms of their validity, ambiguity of words used, length,

appropriateness of the alternatives in case of multiple-choice items. Five experts returned the tests with comments and suggestions to modify some of the items. However, nobody suggested deletion of any item. On the basis of experts' suggestions, all those items requiring modifications were rewritten. As a result, all the 63 items were retained for the tryout purpose. The test was given to an English editor for language editing. Then, it was given to two eighth grade students of English medium school who were average learners for their comprehensibility. It was found that the items in the test were comprehensive for them. The distribution of items to different components of the test was unequal. This was done taking into account the nature of the components. For example, under the Component-4, there were 20 items as students had to distinguish between true, false and no inferences, whereas under Component-3 there were only two items as students had to identify only the central idea of the two passages. Thus, distribution of items to different components on the tryout form of the test on habits of precise thinking is presented in Table 5.6.

TABLE 5.6 : Distribution of Items to Different Components on the Tryout Form of the Test on Habits of Precise Thinking

Components	No. of items
One	15
Two	17
Three	20
Four	2
Five	9
Total	63

There was no need of arranging 63 items in the test as items belonging to a particular component were kept under that component only. However, under Component-1, items belonging to fact and hypothesis were spread over the component equally. Similarly, under Component-3, items belonging to true, false and no inference categories were spread over the component equally. In the test form, these five components were called as five sections namely I, II, III, IV and V in order to avoid confusion among the students regarding components. The tryout form of the test is given in Appendix H .

v) Tryout of the Test

To make a selection from the pool of sixty three items, tryout was conducted. This required standard eighth grade students of English medium schools in the urban setting other than the school where the experimentation was carried out. Therefore, Laxman Public School, New Delhi, whose consent could be obtained, was selected for the tryout purpose. The selection of the school was done due to certain reasons which have been discussed under 5.4(vi) i.e. the tryout of the test on interest in inquiry.

The sample for the tryout consisted of 62 VIIIth grade students of Laxman Public School. During the tryout time, students were told about the purpose of the test, namely, to measure their precise thinking ability. Students were placed in a comfortable situation during the testing time. Students were allowed to take their own time to complete the test. On an average, the time taken by students was about an hour.

vi) Scoring Procedure

While scoring the responses of the students on the items in the test, 1-0 scoring procedure was followed. This means one mark was awarded to each correct answer and zero was given to each wrong answer. The scoring was

done with the help of handmade scoring key.

vii) Selection of Items

In order to select suitable items from the present pool of items, item analysis of the items had to be carried out. For item analysis, the procedure prescribed by Kelley (1939) was followed. According to this procedure, upper and lower 27 per cent of the group in terms of the performance are taken into consideration. By taking 27 per cent of the total group, one can say with the greatest confidence that those in the upper group are superior in the ability measured by the test to those in the lower group (Kelley,1939). Applying this procedure, all the test papers were rank ordered from the highest to the lowest score. Twenty seven per cent of the test papers i.e. 16.74 or 17 test papers were taken from the top and the bottom. Thus, the upper and lower groups constituted 17 respondents each. Then, difficulty value and discrimination index for each item were computed using the prescribed formula. The difficulty value and discrimination index for each item are presented in Table 5.7. While selecting items for the final test, both difficulty value and discrimination index had to be taken into account. According to Psychometricians, items having average difficulty value and high discrimination index should be selected for a test. This point was considered while selecting items for the test. However, in some cases items having low

TABLE 5.7: Difficulty value and Discrimination Index of Items Under Various Components:

Sl.No.	Component	Items	Difficulty value	Discrimination Index
I	Distinguishing between fact and hypothesis.	1	67.64	. 17
		2	88.23	. 23
		3	85.29	. 29
		4	76.47	. 11
		5	85.29	. 29
		6	64.70	. 35
		7	64.70	. 35
		8	76.47	. 35 *
		9	64.70	. 47 *
		10	85.29	. 29
		11	70.58	. 47 *
		12	76.47	. 35
		13	85.29	. 29
		14	79.41	. 41 *
		15	82.35	. 35 *
II	Delineating similarities and differences between the related concepts.	1	52.94	. 94 *
		2	50.00	. 88 *
		3	41.17	. 70
		4	58.82	. 82
		5	55.88	. 72
		6	44.11	. 76
		7	61.76	. 76
		8	61.76	. 76

Table 5.7(contd.)

Sl.No.	Components	Items	Difficulty value	Discrimination Index.
		9	55.88	. 76
		10	58.82	. 70
		11	58.82	. 70
		12	29.41	. 23
		13	52,94	. 94 *
		14	52.94	. 82 *
		15	32.35	. 52
		16	55.88	. 88 *
		17	52.94	. 58
		Paragraph-A		
III	Distinguishing between warranted and unwarranted inferences.	1	58.82	. 11
		2	79.41	. 14
		3	23.52	0
		4	55.88	. 14
		5	52.94	. 11
		6	29.41	. 11
		7	70.58	. 35 *
		8	41.17	. 23
		9	55.88	. 05
		10	70.58	. 23
		11	67.64	. 05
		12	35.29	0
		Paragraph-B		
		1	70.58	. 35 *
		2	58.82	. 05

Table 5.7 (contd.)

Sl.No	Components	Items	Difficulty value	Discrimination Index
		3	58.82	. 35 *
		4	61.76	. 52 *
		5	26.47	. 05
		6	79.41	. 17
		7	52.94	. 58 *
		8	79.49	. 29
		Passage-1		
IV	Identifying the central idea of a theme.	1	11.76	. 11
		Passage-2		
		2	35.29	. 23 *
V	Recognising the underlying assumption behind a proposition.	1	91.17	. 05
		2	47.05	. 35 *
		3	64.70	. 35 *
		4	61.76	. 64 *
		5	44.71	-0.17
		6	23.52	. 47 *
		7	17.64	. 11
		8	20.58	. 17
		9	29.41	. 58 *

* Items selected for the final test.

discrimination index had to be selected because their inclusion was required e.g. item No.2 under Section-IV.

A glance at the Table 5.7 provides that whereas most of the items under Section II have good difficulty value and discrimination index, all the items under Section-III except five items have very low discrimination index. Therefore, a policy had to be taken while choosing items from different sections for the final test. It was decided to select equal number of items from each section except Section IV. So, the maximum number of items to be selected from Sections I, II, III and V was five as the maximum number of items under Section III qualifying discrimination index wise was five. Out of five items under Section III, two items each were statements of true and false inference and one was statement of no inference. Similarly, three statements of hypothesis and two statements of facts were selected from Section-I. From SectionsII and V, five items each having highest discrimination index were selected. Thus, discrimination index of the items selected ranged from .23 to .94 and the difficulty value of the items ranged from 23.52 to 82.35.

viii) The Final Form of the Test

Through the process of item analysis and selection, the final form of the test consisted of a total of twenty one items. SectionsI, II, III and V consisted of five items each, whereas Section IV comprised only one item. The final form of the test and the scoring key is given in the Appendix I .

ix) Psychometric Properties of the Test

To fulfil this purpose, reliability and validity of the test had to be calculated. Reliability and validity of the test are given below.

(a) Reliability

There have been mentions of various methods of estimating the reliability of a test. From among them, the method providing a measure of internal consistency was followed for the present test. Out of two methods measuring internal consistency, namely, split-half and Kuder-Richardson, Kuder-Richardson method was selected. Kuder-Richardson method was preferred owing to its simplicity in computation. There are two formulae of computing reliability through Kuder-Richardson method - KR-20 and KR-21. The researcher opted for the formula KR-21 as it is less cumbersome than the formula KR-20. By using the formula, the reliability of the test was worked out. The reliability of the test was found to be•80 which is considered to be a good estimate of reliability for a test.

(b) Validity

The present test intends to measure the precise thinking ability of the students, which is a psychological trait. Whenever we wish to interpret test performance in terms of psychological trait or quality, we are concerned with

construct validity (Gronlund, 1981). There are different methods of obtaining evidence for construct validation. One method is to analyse the mental process required by the test items. At the outset, the construct "Precise thinking" was operationalised in terms of its components by analysing the definition given by Ausubel (1985) himself. Then, test items were developed on each of the components. The pool of items were subjected to experts' judgement in terms of the validity of the construct. On the basis of their comments and suggestions, the items were modified. The pool of items thus arrived at was tried out and from the tryout, the final test form was developed. In this way, the construct validity of the test was taken care of.

5.6 RETENTION TEST ON MEANINGFUL ASSIMILATION OF INFORMATION AND IDEAS

1) Rationale

Retention of meaningful learning is an important aspect of Ausubel's theory of meaningful verbal learning (Ausubel, 1963). Ausubel asserts that advance organizers enhance retention by increasing the discriminability of learning material from analogous and conflicting ideas in the learner's cognitive structure, thus reducing the possibility of the new material being represented by these ideas in memory. Quoting studies by Kastrinos (1965), Kuhn (1967), and Ring (1969), Ausubel et al. (1978) claim that meaningfully

learned concepts and propositions may be retained for a period of years and that they continue to function in the facilitation of new meaningful learning. Retention of meaningful learning has been studied by a few researchers in the past (Lawton, 1977; Holzman et al. 1982; and Healy, 1989). But many researchers on advance organizer have not given importance to this significant aspect of meaningful learning (Clawson & Barnes, 1973, Kahle and Rastovac, 1976). Therefore, the researcher thought it relevant to study the retention of meaningful learning.

ii) Retention Test

There are two different approaches to evaluate the retention of meaningful learning. One of them is to develop a retention test on meaningful learning which is different from the test on meaningful learning given at the end of the experiment (Ghosh, 1986). The other approach is the use of the post-test on meaningful learning as the retention test. (Holzman et al, 1982; Healy, 1989). Out of these two approaches, the researcher opted for the second one as it was thought that by using the post-test on meaningful learning as retention test, it would be easier to know whether assimilated ideas and information have been retained. Therefore, the comprehensive test on

meaningful assimilation of information and ideas was used as the retention test on meaningful assimilation of information and ideas. As a result, no separate test for measuring retention of meaningful learning was developed. The detailed discussions on comprehensive test on meaningful assimilation of information and ideas have been given under the section v of development of the testson meaningful assimilation of information and ideas.

5.7 REACTION SCALE TO STUDY STUDENTS' REACTIONS TOWARDS ADVANCE ORGANIZER MODEL

One of the objectives of the study was to measure the reactionsof the students toward the Advance Organizer Model of teaching. This required students to give their opinions about the model of teaching. The measurement of this could have been done by asking the students individually about their reactions towards the model. However, the researcher thought that the students might not react voluntarily towards the model if they are asked verbally to do so. It was, therefore, decided to construct a test consisting of 10 items to measure their reactions. These items were in the form of declarative sentences. Six of them were in the positive form and four of them were in the negative form. Items 1, 2, 4, 6, 8 and 9 were in the positive form and items 3, 5, 7 and 10 were in the negative form. The students were expected to write (yes) against a statement if they agreed to that statement and (No) if they disagreed to that statement.

The researcher did not include various degrees of agreement or disagreement as the response mode because it was thought that such a response pattern might confuse the students in reacting to the model. All the statements, thus developed, were validated by five experts. All the items in the test were preceded by appropriate guidelines as to how to respond to them. Language suitability and ambiguity of words used in the statements were taken into account while framing the statements. The final form of the test is given in the Appendix J.

5.8 SUMMARY

The present Chapter dealt with the development of tools to assess different behavioural outcomes arising out of the treatments. Tools were developed to measure the instructional and the nurturant effects of the Advance Organizer Model. Reaction scale was also developed to assess the reactions of the students of experimental group towards Advance Organizer Model. Four unit tests and one comprehensive test were developed to study the students' ability to develop conceptual structures or concept maps. Similarly, four unit tests and one comprehensive test were developed to assess students' assimilation of information and ideas in a meaningful way. All these tests were content specific. Test on interest in inquiry was developed and tried out on a sample of 72 students. The split-half reliability of the test was .60. Similarly,

test on habits of precise thinking was also developed. The KR-21 reliability of the test was .80. Both these tests were constructed to measure nurturant effects of the model. The comprehensive test on meaningful assimilation of information and ideas was used as retention test. A reaction scale conisting of 6 positive items and 4 negative items was constructed to study students' reactions toward AOM. The next Chapter deals with analysis and interpretation of data collected with the help of the above discussed tools.