LEARNING MATERIAL

FOR

THE LECTURE GROUP

About this Learning Material

You will be given six booklets about 'educational evaluation'. These booklets correspond to the six instructional units of the course on 'Educational Testing and Techniques of Evaluation'. The six units are :

I. Educational Evaluation and Measurement - A

II. Educational Evaluation and Measurement - B

III. Charecteristics of a Good Instrument of Evaluation

IV. Major Tools and Techniques of Evaluation

V. Teacher-made Achievement Tests

VI. Elementary Statistics in Education

All these topics are dealt with, mainly, in the context of school education.

Note that this material is only supplementary to the lectures you have to attend; it is in no way sufficient by itself. Therefore you should attend the lectures regularly. This material is meant to help you remember what you learn through the lectures in an organised manner.

I. Educational Evaluation and Measurement - A

Although, education and evaluation do not appear to be new terms, it may not be known as to what it would exactly mean when the two terms are put together and called educational evaluation. Therefore, first of all, answer would be obtained to the question :

--- What is educational evaluation?

Answer to this question will indicate that educational evaluation refers to a 'process' carried out in schools. As a process, it should obviously involve several steps. Therefore, the second question to be answered would be :

--- What are the steps involved in the process of educational evaluation?

It will also be noted that educational evaluation involves assessing pupils as a part of regular school work. One may again think what is it that is evaluated in pupils; and, also, if evaluation is a part of school work when is it carried out. Because, at schools pupils study, play, mix with other pupils and do various other activities. Also, school work goes on as a daily routine from morning till evening throughout the year. Therefore, in this regard, answer will be obtained to the questions :

--- When do we evaluate pupils?

--- What do we evaluate in pupils?

With all these questions answered one may still be wondering why should this process of evaluation be carried out in schools at all. Thus, lastly, answer would be obtained to the question :

--- What are the uses of educational evaluation?

Meaning of Educational Evaluation and Steps involved in the Process of Evaluation.

Educational evaluation is an integral part of school activities and schools are meant to provide education to the children. Therefore, in order to understand the meaning of educational evaluation, it is necessary to know some details as to how education is provided at the schools.

In order to provide education to the children, various activities like classroom instruction and laboratory training in different subjects, games, sports and such other group activities, extra-curricular activities like science club, debates, dramatics, etc. are organised in the schools. Thus, the purpose of all these activities taken together is to provide education to the children. In other words, it may be said that each activity, at the school, is organised with a view to achieving certain specific objectives and these specific objectives together form the purpose of school education. For instance, students are taught at the schools about Moghul rulers in India with the specific objective that they learn certain important dates, names and events related to Moghul rule in India. Similarly, various topics in arithmetic are taught with the specific objective that students learn to solve arithmetical problems. Different languages are taught at the school with the specific objectives that students learn to express their ideas to others and also understand those of others. Various group activities are organised in the schools with the specific objective that students learn to work with others in a group.

From the above instances, it is clear that various activities are organised in the school with certain objectives specifying what the students are expected to learn. These specific objectives of various activities organised in the schools are called 'educational

objectives' as they together constitute the purpose of school education. This implies that 'to provide education' is 'to achieve the various educational objectives'.

Thus, to achieve the educational objectives, various activities are organised at the school. But, the following crucial question arises at this point. How far can one be sure that the specific educational objectives are being achieved through particular activities? In order to answer this question, it is necessary to determine the extent to which the specific educational objectives have been achieved through the particular activities. This process of determining the extent to which educational objectives have been achieved is called 'educational evaluation'.

Since, educational objectives specify what the students should learn, educational evaluation may also be described as the process of determining the extent to which students have learnt what they should have learnt. Analysing a specific school situation to see how a teacher carries out the process of evaluation would help understand the process better.

Suppose a teacher has taught about the density of materials with the objective that all the students should learn the meaning of density. After teaching, the teacher wants to evaluate i.e., he wants to find out the extent to which he has achieved his objective. In other words, he wants to determine the extent to which students have learnt what they should have learnt. Naturally, this would require him to, first, find out what the students have actually learnt i.e., students' learning. Thus, after determining students' learning, the teacher has, on one hand, students' learning telling him what the students have actually learnt and on the other, educational objectives specifying what the students should have learnt. Therefore, he determines the extent to which his objectives

have been achieved by comparing the objectives (what they should have learnt) with students' learning (what the students have actually learnt). Thus, in the process of evaluation, students' learning is determined and is then, compared with the objectives to determine the extent to which the objectives have been achieved.

From the above description of the process of educational evaluation, it is clear that in order to understand the details involved in the process of evaluation it is necessary to, first, know in detail about the following points.

(1) What do educational objectives specify and the meaning of 'achieving educational objectives'.

(2) The procedure to be followed in determining students' learning i.e. determining what students have actually learnt.

As has been already noted, educational objectives specify why various activities are organised in the school. Therefore, the meaning of achieving educational objectives can be understood by analysing, in terms of what has happened to the students, certain instances from school wherein specific educational objectives have been achieved.

(a) A boy when he joins the school in std.I could not tell his friend how much money he had in his pocket. But, through teaching, changes were brought about in his ability to count and add so that when he came to std.II he could easily count and specify the amount of money with him.

Thus, in this instance, instruction has been provided in the school with the objective of bringing about desirable changes in the boy's computational skills.

(b) Since his childhood, Lakshman used to wonder why a big and heavy log of wood could float on water whereas even a small piece of stone would sink down. The reason became clear to him when his science teacher taught about 'density of materials'. Thus through teaching certain changes took place in his knowledge and understanding of scientific facts and phenomena.

In the above instance, it may be said that instruction has been provided in the school with the objective of bringing about desirable changes in knowledge and understanding.

(c) Raj, when he joined the high school, had a weak physique and poor health. On the advice of his teachers, he participated in games, sports, physical training classes and such other activities organised at the school regularly. Within a year he found himself quite strong and healthy. Thus, desirable changes in his physical abilities took place.

Thus, certain activities are organised in the school with the objective of bringing about desirable changes in physical abilities.

(d) Usha was a problem child for her parents as she frequently got into temper tantrums. She was admitted to a nursery school. Even at the school, she could not adjust herself to other children and used to quarrel with them. However, her teachers tried to understand her. Slowly, they made her mix with other children of her own age and also made her observe the behaviour of other children in different situations. As a result, within a few months, Usha's parents found her manageable. Thus, desirable changes were brought about in

her emotional and social adjustments by organising suitable experiences and activities in the school.

Thus, certain activities are organised in the school with the specific objective of bringing about desirable changes in emotional and social adjustment.

(e) Nutan, after attending classes in Fine Arts for several years, has been able to sing and appreciate music. She has also developed great interest in painting and sculpture.

Thus, suitable activities are organised in the school with the specific objective of bringing about desirable changes in appreciations and interests.

From the above analysis of different instances from school, it is clear that various activities are organised in the school with specific objectives. Further, these objectives specify the desirable changes to be brought about in the students' knowledge, understandings, skills, interests, attitudes, appreciations, adjustments and physical abilities, etc. In fact, it has already been pointed out that the specific objectives of the various activities organised in the schools are called 'educational objectives'. Thus, it may be stated that "Educational objectives are specifications of desirable changes to be brought about in the students' knowledge, understandings, skills, etc." From this description of educational objectives, it may also be stated that "To achieve educational objectives is to bring about desirable changes in the students' knowledge, understandings, etc."

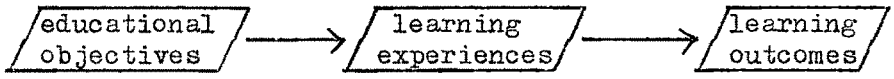
Changes in the students' knowledge, understandings, etc., are the outcomes of their participation in various activities organised in

the school and these changes essentially represent the learning that takes place in the students. Therefore, the changes in the students' knowledge, understandings, etc., may be called "learning outcomes". Further, learning takes place in the students because various activities provide the experiences necessary for the learning to take place. Thus, the various activities organised in the schools can be referred to as "learning experiences".

Relationship between 'learning outcomes' and 'learning experiences' may be represented by the following figure

Also, different learning experiences are chosen on the bases of the educational objectives to be achieved. Again, this relationship may be represented by the following figure.

Observing the above two figures, the following figure may be presented which provides a clear picture of the relationship between objectives, learning experiences and learning outcomes.



Consider the school situation where educational objectives have been specified, appropriate learning experiences have been provided and also, learning outcomes have been determined. Here, educational objectives specify expected changes in the students; learning outcomes specify changes in the students that have actually taken place; and these changes correspond to the educational objectives

that have been achieved. Combining these meanings of educational objectives and learning outcomes, it may be said that educational objectives are expected learning outcomes and learning outcomes represent achieved educational objectives.

Thus, it may be stated that 'to achieve educational objectives is to ensure that the actual learning outcomes correspond to all the expected learning outcomes'.

The next question to be answered is 'what procedure should be followed in determining the learning outcomes?'

Again, answer to this question may be obtained by considering the following specific school situations.

(i) Consider a Physics class in the school. The teacher is teaching about the 'density of materials'. After explaining about the meaning of density, in order to check whether the students have learnt the point (i.e. in order to find out the learning outcomes), teacher asks the students questions on the meaning of density.

Thus, the teacher determines the learning outcomes through the technique of questioning.

(ii) Over several periods, the teacher has taught the students in detail about all the Moghul rulers in India. Then he gives the students a test consisting of questions related to Moghul rulers in India in order to check their knowledge about the topic.

Thus, the teacher determines the learning outcomes through the technique of testing.

(iii) The teacher has trained the students to use different measuring instruments in the chemistry laboratory. Then, he wants to check the skill of the students in measuring a liquid with the help of a burette. In this instance, teacher does not ask questions or gives a test where the students have to write about the use of a burette. Instead, he allows the students to measure a liquid using a burette and observes them when they do it. Thus, the teacher has to select and use the technique of observation. Because, by obtaining answers to questions, teacher can only check the knowledge of the students but he cannot check the skill in using a burette.

Thus, the teacher, first, selects the suitable technique considering the objective to be evaluated and then, uses that technique to determine the learning outcomes.

(iv) The teacher has often been stressing the importance of being cooperative while working in a group. He, then, wants to know to what extent students really cooperate with one another while working together. Again, in order to get this information i.e., in order to determine the learning outcomes, teacher selects the technique of observation.

Again, the teacher selects the suitable technique, and then, he uses that technique to determine the learning outcomes.

From the above instances, it is clear that the teacher has to first, select a particular technique such as oral questioning, testing, observation, etc. on the basis of the objectives with reference to which learning outcomes are to be determined. These techniques are called 'evaluation techniques'. In the second step, teacher actually

uses the technique and determines the outcomes.

Therefore, the procedure to be followed in determining the learning outcomes consists of two specific steps. But, determining learning outcomes constitute the first stage in the process of evaluation. Thus, the first two specific steps in the process of evaluation are :

(i) select the evaluation technique suitable for the educational objectives in consideration,

(ii) use the technique to determine the learning outcomes.

Then, before going to the second stage of the evaluation process viz., comparing the learning outcomes with the educational objectives, it may be observed that the purpose of their comparison is to determine the specific extent of achievement of the objectives. It is clear that both learning outcomes and educational objectives represent learning. In fact, learning outcomes represent the actual amount of learning while educational objectives represent the expected amount of learning. Therefore, to be more accurate in determining the extent of achievement of the objectives, learning outcomes (actual amount of learning) should be first, measured and then, be compared with educational objectives (expected amount of learning). Thus, the third and the fourth steps in the process of evaluation are :

(iii) measure the learning outcomes,

(iv) compare the learning outcomes with the educational objectives to determine the extent of achievement of the objectives.

Finally, it should be observed that a teacher determines the extent of achievement of the objectives in order to judge whether it is satisfactory or not. Thus, the fifth and the final step in the process of educational evaluation is :

(v) judge whether the extent of achievement of the objectives is satisfactory or not.

From the above analysis, now it may be stated that the process of educational evaluation involves five specific steps. The five steps are :

(i) Select the evaluation technique suitable for the educational objective in consideration.

(ii) Use the technique to determine the learning outcomes.

(iii) Measure the learning outcomes.

(iv) Compare the learning outcomes with the educational objectives to determine the extent of achievement of the objectives.

(v) Judge whether the extent of achievement of the objectives is satisfactory or not.

Characteristics of a good evaluation programme.

While learning about the process of educational evaluation and the steps involved in it, it was noted that evaluation is an important activity that goes on in the school. Therefore, it should be interesting to find out when this activity of evaluation is carried

out. In other words, the question to be answered is : When are pupils evaluated?

In order to educate children at the school

1. educational objectives are specified,

2. appropriate learning experiences are provided for achieving the specified objectives, and

3. pupils are evaluated to check whether the objectives have been satisfactorily achieved or not.

This description somewhat gives the idea that evaluation is always done at the end i.e. after specifying the objectives and providing suitable learning experiences. Of course, evaluation is done at the end but not only at the end. However, the above description also indicates the following four possible answers to the question : 'When are pupils evaluated in the schools?'

Evaluation is carried out in the school

(i) at the beginning, i.e. before specifying the objectives.

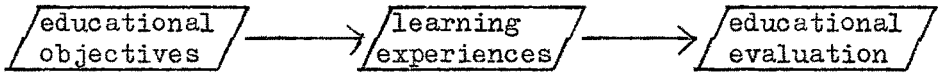
(ii) while providing the learning experiences.

(iii) at the end, i.e., after specifying the objectives and providing suitable learning experiences.

(iv) at all the above stages.

Again, proper answer may be got by analysing certain instances of evaluation at the school and finding out at what stages have these instances taken place.

Final examinations that are conducted at the end of every year is a familiar instance of evaluation. In the final examination, pupils are evaluated in order to see the extent to which each one of them has achieved the objectives specified for that year through the learning experiences provided. Thus, it may be said that final examination is an instance of evaluation that comes at the end. This situation may be represented by the following figure.



However, the same instance of final examination can be viewed from a different point.

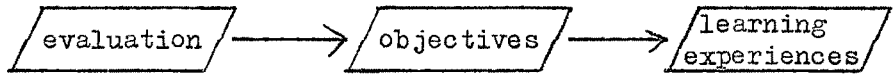
> Suppose that students have just come to std.X after completing std.IX. But, it is found that a few students who were also with these students in std.IX have been made to stay back there only. Obviously, the reason for the above situation is that the latter few students have not passed the final examination and therefore, they have not been promoted to the next higher class.

In other words, it may be said that promotions of students are decided on the basis of the evaluation done through the final examination. And these promotions have further implications for other activities.

> Consider two students A and B. A has been promoted to std.X while B remains in std.IX. Thus, A learns new things while B learns the same things which he has been taught the previous year. This means that new and different objectives are set for A to achieve.

Thus, decision about what objectives different students should follow is taken on the basis of the evaluation done through the final examination.

From the above instance it is clear that objectives to be achieved are specified on the basis of evaluation : and different learning experiences are provided accordingly. This situation may be represented by the following figure.

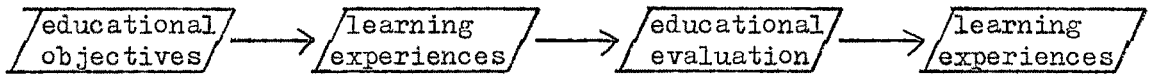


This is, therefore, an instance of evaluation that comes at the beginning, i.e., before setting the objectives.

One more common situation involving evaluation may be considered.

> The teacher started the lesson with the objectives that after his teaching (i) students recognise the cause and effect relationship between heating an object (solid) and its size : (ii) students give two illustrations from their daily life experiences wherein expansion of solids is involved. After teaching the students about the effect of heating on the size of an object (i.e. after providing the learning experiences), teacher asks some questions to check whether the students can recognise the relationship (i.e. he evaluates). But, he finds that the students cannot recognise the relationship between heating and the size of an object. Then, the teacher, as he finds (through evaluation) that students are not learning, changes his method of presentation as well as the content (i.e. presents new learning experiences) in order to enable them to learn.

Thus, in the above instance, teacher started with specific objectives and provided learning experiences and on the basis of his findings through evaluation he again presented new learning experiences. It is clear that, this is an instance wherein evaluation is done while presenting learning experiences to see the suitability of the learning experiences for achieving the objectives. This situation may be represented by the following figure.



It is quite evident from the above discussion of different instances of evaluation that evaluation is done, continuously, at all stages of teaching i.e., while specifying the objectives, while providing learning experiences, and also later on to check whether the objectives have been achieved satisfactorily or not.

The above statement provides a complete answer to the question, "When are pupils evaluated?"

Further, it also states the principle, "Educational evaluation is a continuous process".

It was noted that evaluation is an integral part of school activities and is carried out continuously at all stages of providing education. Another point that needs consideration is regarding what is evaluated in the pupils in the different instances of evaluation.

The definition of educational evaluation does not directly provide any answer to the above question as it refers only to the achievement

of objectives. However, a careful examination of the meaning of educational objectives indicates that it is the extent of the changes in the students' knowledge, understandings, etc., that is evaluated through different instances of evaluation. Thus, answer to the question 'what is evaluated in the pupils?' lies in understanding what do changes in knowledge, understandings, etc., correspond to in the pupils.

It could be easily understood that changes in a child's physical abilities such as strength, physical independence, etc. correspond to physical development of the child. Similarly, changes in interests, attitudes, appreciations, adjustments, etc. correspond to social and emotional development, while changes in knowledge, understandings and skills correspond to intellectual development.

From the above descfiption about changes in the child and their relation to the different aspects of the pupil's development, it is clear that what is evaluated in the pupil is different aspects of development, viz., physical, intellectual, social and emotional. In fact, it may be said that development of all the different aspects, namely physical, intellectual, social and emotional, together represents the development of the whole child. Thus, a good evaluation programme takes into consideration the evaluation of the whole child.

The statement that "A good evaluation programme takes into consideration the whole child" implies that evaluation of only the intellectual aspect of the pupil through certain tests and examination cannot be considered a good evaluation programme unless it also includes evaluation of other aspects of development, viz., emotional, social, physical, etc. In other words, it specifies a characteristic of any good evaluation programme, namely, "A good

evaluation programme should be comprehensive". That is, it should be broad enough to include the evaluation of all the different aspects of development of the child.

The above discussion regarding the questions, 'when are pupils evaluated?' 'what is evaluated in the pupils?' offers two specific principles that should be borne in mind while organising the programme of evaluation in any school.

1. Evaluation should be continuously done at all stages of providing education so that all decisions that are taken regarding the objectives to be specified for different students at different levels and all decisions of the teachers regarding methods and contents that they should adopt, are based on evidences obtained through evaluation.

2. Evaluation programme should be comprehensive so that evaluation is done in all aspects - physical, intellectual, social, emotional, etc. - of development of the children.

Uses of Evaluation :

It was noted that a good evaluation programme should consider all the aspects of development of the pupil. Accordingly, evaluation is carried out in the school at different levels.

Evaluation is done with respect to specific areas of learning through entrance examinations or pretests, through questioning during day-to-day classroom teaching, through periodical tests and terminal examinations. Evaluation is also done which will be schoolwide in character and provide information about the pupils regarding their scholastic aptitudes, record of growth of achievement in different

subjects, personal social development, health, home background and the like.

The sole purpose of conducting evaluation in the schools is that it helps improve education provided at the school. But, improving education through evaluation requires that results of evaluation be used by all those involved in school education, viz., by (1) teachers; (2) pupils; and (3) administrators or school authorities.

Results of evaluation consist of information about the pupils regarding their scholastic aptitudes, personal-social development, health, home background, etc. and regarding the extent to which educational objectives have been achieved. Thus, the question to be answered is : "What are the uses of evaluation for the teachers, pupils and administrators or school authorities?"

Uses for the Teacher :

It is a known fact that although equal facilities are provided in the school, pupils differ in their learning abilities, scholastic aptitudes, home background, social and personal adjustment, etc.

> Suppose there are two teachers 'A' and 'B'. Teacher 'A' knows about aptitudes, home background, social-personal adjustment, learning ability, etc. of every one of his pupils. This information he has in addition to knowing their abilities in the subject which he teaches. But, teacher 'B' is concerned only with his pupils' abilities to learn that subject which he teaches, and he does not know anything about their home background, and adjustment in the school etc. Of these two teachers 'A' and 'B', it is teacher 'A' who would be able to teach his students more

> effectively, because, learning, whether in the classroom or outside, depends upon various factors such as the learner's adjustment in the school and family, his interests and aptitudes, etc.

Thus, a teacher who understands his students in all aspects including their abilities in the particular subject can teach more effectively. And, a teacher can get information regarding his students' home background, interests, etc. from the results of the evaluation. Therefore, "Evaluation is useful to the teacher in understanding his students in terms of their home background, interests, attitudes, etc., and thereby be more effective in his teaching".

Again, it is known that, in each class and in each area of learning students differ in their abilities and achievement. In order to teach effectively, a teacher should know the abilities and achievement of his students in that particular area of learning so that he can adjust his teaching to the differences among them. In other words, the teacher should understand his students in terms of their abilities and achievement in particular areas of learning. Again, this information he can obtain through evaluation. In fact, he can obtain this information about the pupils either (i) from the evaluation records of the pupils during the previous years, or (ii) from the results obtained through a test specifically prepared for this purpose. Thus, "Evaluation is useful to teachers in understanding their students in terms of their abilities and achievement in particular areas of learning."

Thus, in general, it can be said that evaluation is useful to teachers in understanding their students.

Suppose a teacher has taught his students about the two world wars. He asks the students several questions about facts related to the world wars. In other words, he evaluates. During the evaluation he has to come across one of the following situations.

(i) All the students answer his questions satisfactorily.

(ii) Only a few students do not answer satisfactorily.

(iii) Majority of the students do not answer satisfactorily.

Correspondingly, the teacher would proceed in the above situations as follows.

(i) Teacher will proceed with the next topic of teaching.

(ii) Teacher meets those few students and helps them individually.

(iii) Teacher reviews or reteaches the same topic in a new way.

Thus, teacher decides to do different things in different situations on the basis of results of evaluation. That is, "Teachers can use evaluation results to take decisions about teaching at each stage as to what they should do next".

It has been noted that when majority of the students do not answer satisfactorily, the teacher understands that his teaching has not been effective and therefore he changes his method of teaching. Thus, through results of evaluation of pupils, teacher checks the effectiveness of his own teaching. In other words, "evaluation of pupils is useful to the teachers in evaluating their own teaching".

Apart from the three uses of evaluation for the teachers, that have been noted so far, there is another important use which is illustrated by the following instance.

> Suppose that a mathematics teacher has taught the unit on 'addition, subtraction, multiplication and division of decimal numbers'. In order to evaluate the pupils, he gives a test constituting of problems on decimals. When he checks the answer of the pupils he finds that many of the pupils have gone wrong on particular items in the test. He further discovers that all these items involve 'multiplication of decimals'. This indicates to the teacher that pupils have difficulty in learning the multiplication of decimals. The teacher, then gives a more detailed test consisting of only multiplication problems involving different types of decimals. He further, finds that pupils have difficulties in multiplying only a particular type of decimals. Thus, the teacher diagnoses the learning difficulty and then, he provides necessary instruction to remedy the difficulty.

It may be said that evaluation results helps the teacher in diagnosing the learning difficulties and providing suitable remedial instruction. Thus, another use of evaluation for the teachers is in diagnosing the learning difficulties of the students.

Four ways in which evaluation is useful to the teachers are

(i) in understanding their students,

(ii) in taking decisions about teaching as to what they should do next,

(iii) in evaluating their own teaching, and

(iv) in diagnosing the learning difficulties of the students.

Uses for the Students :

Results of evaluation tell the students of their own progress in different areas of learning.

> Consider two students 'X' and 'Y'. Suppose 'X' continuously knows about what is expected of him and what has been his progress in different areas of learning. But 'Y', although he is studying with 'X', does not know about his own progress in different areas. Obviously 'X' is likely to be more motivated to learn than 'Y'.

"Evaluation can, therefore, be useful to the students as it motivates them towards learning by providing them the knowledge of their progress".

In educational evaluation, the pupils' actual achievement is compared with what they are expected to achieve. When evaluation results (marks or grades) are provided to the pupils, continuously, pupils themselves try to compare their achievement with their earlier achievements, with the average achievement in the class, and also with what they are expected to achieve. Thus, it may be observed that using the results of continuous evaluation pupils try to evaluate themselves. In other words, "continuous evaluation leads the pupils towards self-evaluation".

Thus, evaluation is useful to the pupils as it (i) motivates them towards learning and (ii) leads them towards self-evaluation.

Uses for the Administrators or School Authorities :

An important decision, that administrators have to take every year, is regarding promotion of students. It has already been noted that promotions are decided on the basis of their performance in tests and examinations, i.e. through evaluation.

Also, it has been observed that admission of pupils to different courses are decided on the basis of their performance measured through either examination conducted in the previous year or an entrance examination. Therefore, it can be said that, evaluation is also useful for the administrators in deciding about admissions of pupils to various courses.

Thus, "evaluation is useful for the administrators in deciding about admissions and promotions".

Lastly, in certain school systems, pupils are expected to take up different branches of study like science, arts, commerce, etc. for specialisation. We know that success in any branch of study depends upon the previous learning of the pupils, their interests and aptitudes. In such situations, school authorities are required to provide guidance to pupils about the suitability of different courses. This guidance is provided to pupils on the basis of evaluation results regarding their previous learning in particular branches of study, their interests and aptitudes.

Therefore, it may be finally said that "Evaluation is useful for providing guidance to pupils".

II. Educational Evaluation and Measurement - B

The last section dealt with the meaning of educational evaluation, the five specific steps involved in the process of evaluation and certain other related facts about educational evaluation. In the present section more details will be presented regarding the different steps in the evaluation process as it may be carried out in school.

It was observed, that the two main stages in the process of evaluation are determining the learning outcomes and comparing them with the corresponding educational objectives. Therefore, one has to first know, clearly, the way, in which learning outcomes represent learning and how should the educational objectives be stated so that learning outcomes can be properly compared with them. In this connection, answers will be obtained to the following specific questions.

--- How do learning outcomes specify learning?

--- How to state educational objectives in terms of students' behaviour?

Further, it was noted that the first specific step in the process of evaluation is to select suitable techniques of evaluation. Therefore, the following specific questions will be answered about the techniques of evaluation.

--- What are the major techniques of evaluation?

--- What are the characteristics of the different techniques?

It is known that 'testing' is the most familiar technique of evaluation used in schools. Sometimes, testing is talked about as though it is synonymous with evaluation. Thus, the role of tests in evaluation will be clarified by answering the question.

--- What is the place of 'tests' in the evaluation process?

It was also noted in the previous section that learning outcomes should be measured in order to make the evaluation accurate. Following three questions will be answered in this regard.

--- What is the meaning of measurement in education?

--- How to measure the learning outcomes?

--- Are all learning outcomes measurable?

It was specified that after measuring learning outcomes, they are, to be compared with educational objectives in order to judge whether the objectives have been achieved satisfactorily or not. But, what should be done, if there are non-measurable outcomes? Therefore, answer will be obtained to the question.

--- How to express evaluation in terms of measurable and non-measurable outcomes?

Specification of Learning Outcomes and Educational Objectives :

As may be easily observed, all discussions about educational evaluation centre, mainly, around three terms, viz., educational objectives, learning experiences and learning outcomes.

Various activities like classroom instruction, laboratory training, games, sports, physical training classes, extra-curricular activities like debates, dramatics, etc., are organised in schools to provide learning experiences to students. These various learning experiences are provided in schools in order to achieve the educational objectives.

Learning or changes in the students' knowledge, understandings, etc., take place when they participate in various learning experiences. In other words, learning or changes that take place in the students at the school are outcomes of the various learning experiences. These outcomes of learning experiences are called learning outcomes.

Thus, learning outcomes specify learning or changes in the students' knowledge, understandings, etc., that take place through the various learning experiences provided at the school.

Or, it may just be said that learning outcomes represent changes in knowledge, understandings, etc. It has been noted that in carrying out the process of evaluation, teacher has to determine the actual learning outcomes. Thus, in order to determine the actual learning outcomes of his instruction, teacher should be able to observe the actual changes in knowledge, understandings, etc. Obviously, one cannot directly observe changes in the students'

knowledge, understandings, etc. Therefore, to determine the actual learning outcomes teacher has to find those changes in the students which are directly observable and which represent changes in the knowledge, understandings, etc.

Thus, certain specific situations of providing learning experiences may be analysed to see what changes does the teacher directly observe in the students after providing the learning experiences. In fact, these observarble changes in the students may be considered as the actual learning outcomes representing changes in knowledge, understandings, etc.

> Consider the student who has not learnt anything about density of materials. The teacher teaches (i.e. provides learning experiences) the student about density of materials'. After teaching, teacher finds that the student
>
> (i) states the meaning of density.
>
> (ii) gives reasons as to why a large log of wood floats on water while even a small piece of stone sinks down.

Obviously, this situation represents changes in the student because after the instruction he states the meaning of density and gives reasons for the particular phenomenon related to it, which he could not have done earlier.

Studying the changes in the students, as mentioned above, it may be observed that changes have been specified by the phrases, 'student states' and 'student gives reasons'. From this it may be simply said that changes in the student are specified in terms of what the student does after instruction. In other words, changes are specified in terms of performance or actions of the student after instruction.

Further, it should be noted that these changes are directly observable. For, one can always observe whether a student states the meaning of density or not, and also, whether he gives reasons for the particular phenomenon or not.

Thus, in the above instance, learning or changes in the student have been specified in terms of what the student does after instruction. And, these changes which are directly observable are the outcomes of the instruction (learning experiences) provided to the student. Therefore, these changes may be considered as the actual learning outcomes.

> In another instance, teacher has taught the student about addition of decimal numbers. He, then, observes that the student solves problems involving addition of decimal numbers.

Again, changes in the student which the teacher could directly observe are in terms of what the student does after the instruction. Further, these changes, representing the learning that has taken place in the student, are outcomes of the instruction (learning experiences) provided to the student. Therefore, these changes may be considered as the actual learning outcomes.

It may be observed that, in the above instances of providing learning experience (or instruction), directly observable changes in the student have been specified in terms of what the student does after instruction. Also, these changes which are the actual outcomes of the instruction provided represent learning, i.e. changes in knowledge, understanding, etc., that has taken place in the student. Therefore, it may be said that in these instances, learning outcomes are specified in terms of what the student does after instruction.

In fact, in any instance of providing learning experience (instruction) learning outcomes may be conveniently specified in terms of what the student does after instruction.

As has been noted earlier, in situations of providing learning experience (instruction), specifications of what the student does after instruction represent changes in his performance. Further, these changes in performance can be considered as specifying changes in the student's behaviour.

Thus, from the above, it may be said that specifications of what the student does after instruction correspond to changes in the student's behaviour. But, these specifications, represent the actual learning outcomes. Therefore, it may be concluded that learning outcomes are specifications of changes in the student's behaviour.

From the above description of learning outcomes, the question, 'How do learning outcomes specify learning?' can be satisfactorily answered as follows :

'Learning outcomes specify learning or changes in knowledge, understandings, etc., in terms of changes in the student's behaviour, i.e., in terms of what the student does after instruction.'

It has been noted earlier that the first main step in the evaluation process is determining the learning outcomes. The foregoing discussion about learning outcomes clarifies as to what one exactly tries to find out in students when he is determining the learning outcomes. In fact, it specifies that to determine the learning outcomes is to find out the changes in the students' behaviour which represent changes in their knowledge, understandings, etc.

In other words, it has been stated that learning outcomes are specifications of changes in the student's behaviour, i.e., what the student does after instruction. Of course, it is true that changes take place in the students' knowledge, understanding, etc., when they participate in various learning experiences. And, therefore, they should be considered as learning outcomes. However, for evaluation purposes, learning outcomes are specified in terms of changes in the students' behaviour which, in turn, represent changes in knowledge, understanding, etc. This is because, changes in behaviour are directly observable where as changes in knowledge, understandings, etc. are not.

It has also been noted that after determining the learning outcomes, the next main step in the evaluation process is to compare them with the educational objectives. The specific extent to which educational objectives have been achieved is determined through this comparison.

If an actual situation of teaching is considered, educational objectives specify changes in knowledge, understandings, etc., that are expected to take place in the students. Thus, in evaluation, one has to actually compare the learning outcomes with the expected changes in knowledge, understandings, etc. From the above it is clear that for evaluation purposes, or more specifically for comparing learning outcomes with educational objectives, it is necessary that educational objectives specify, in clear terms, the expected changes in the students' knowledge, understandings, etc. (i.e. expected learning outcomes).

Therefore, the problem is to find out the proper way in which educational objectives may specify changes in the students such that learning outcomes can be properly compared with them.

It may be said that learning outcomes specify actual changes in the students' knowledge, understanding, etc., while educational objectives specify expected changes in the students. Thus, in order to compare learning outcomes with educational objectives, one has to compare the actual changes in knowledge, understanding, etc., specified by learning outcomes with the expected changes specified by educational objectives. From this, it is clear that the comparison can be done more easily and directly if changes in knowledge, understanding, etc., are specified in similar terms both by learning outcomes and educational objectives. Learning outcomes, specify changes in knowledge, understanding, etc., in terms of changes in the students' behaviour. Therefore, educational objectives should also specify changes in knowledge, understanding, etc., in terms of changes in the students' behaviour, i.e., in terms of what the student does after instruction. Thus, it may be finally concluded that for purposes of evaluation educational objectives should be stated in terms of changes in the students' behaviour. Further, what a student does after instruction may be considered to represent his performance after instruction. Thus, it may be said, in other words, that educational objectives should be stated in terms of students' post-instruction performance.

It has been concluded in the above that, for evaluation purposes, educational objectives should be stated in terms of students' behaviour. However, in order to understand the necessity of stating them in this particular form, one may find out in what other form can they be stated and why that form is not suitable for evaluation purposes.

It may be easily observed that educational objectives can be directly stated in forms of changes in knowledge, understanding, etc.

Suppose a teacher is teaching about world war II. His objective is that through his teaching the students will acquire knowledge about the background of World War II.

Objective in this situation can be stated as follows :

'The students will know the background of World War II.'

Similarly, different educational objectives can be stated in this form. The following are some examples.

1. The students will understand the concept of gravitational force.

2. The students will develop appreciation for music.

3. The students will increase their interest in games and athletics.

Now, an objective of this form may be analysed to see why this form of stating the objective does not help in evaluation.

Let us suppose that two teachers, 'X' and 'Y' are teaching about World War II, independently. Both of them have their objective as follows :

'Students will know the background of World War II'.

It may be easily observed that this educational objective is stated in terms of changes in the students' knowledge.

After teaching, teacher 'X' asks his students to list in writing the three important causes of World War II. On the basis of the students' answers to this, he decides that the objective has been achieved.

Learning outcome expected by the teacher in the above situation may be specified as follows :

'Students list in writing the three important causes of World War II'.

Consider the other teacher 'Y'. This teacher, after his teaching, asks his students to write an essay describing the plitical relationships between the countries involved in World War II and also list the different causes. Again, on the basis of the students' performance he decides that the objective has been achieved.

Learning outcomes expected by the teacher in this situation may be specified as :

(i) Students write an essay describing the political relationship between the countries involved in World War II.

(ii) Students list the different causes.

It is quite clear that the two teacher 'X' and 'Y' expected different learning outcomes.

It may be noted that teachers expect specific learning outcomes after teaching on the basis of the objectives specified. Thus, if educational objectives have been properly stated, for particular educational objectives same learning outcomes should be expected by all the teachers. But, in the above two situations the teachers

had the same objective and yet expected different learning outcomes. This is because, the objective was interpreted or understood by the two teachers in different ways. From this, it is clear that the above objective stated in terms of changes in knowledge does not specify the expected learning or changes in the students in definite and clear terms. In fact, it may be found that all such objectives stated in terms of knowledge, understanding, appreciation, etc., may lead to similar confusion about the expected learning.

Suppose in the illustration discussed above the two teachers had started with the objectives that after their teaching about World War II

(i) student describes, in about 15-20 lines the political conditions that prevailed in Germany during the years 1938-39.

(ii) student lists in writing three causes of World War II.

It can be easily observed that these objectives specify expected learning outcomes in definite and clear terms. Of course, they may seem to be somewhat lengthy but at the same time they are quite clear and definite about the learning outcomes that may be expected by the teacher. Obviously, therefore, these objectives do not allow for different interpretations by different teachers.

Examining these objectives further, it may be noted that in these the phrases, 'student describes' and 'student lists in writing', specify what the student does after instruction. Since, these actions in turn, correspond to the changes in the student's behaviour, it may be finally said that the above objectives have been stated in terms of changes in the students' behaviour.

The above illustration confirms that for purposes of evaluation in specific instructional situations, educational objectives should be stated in terms of students' behaviour. This is because, objectives stated in this form specify the expected learning outcomes in clear and definite terms.

It has also been noted that educational objectives may not be stated using such words as 'knows', 'understands', etc., as they allow for different interpretations regarding the learning outcomes to be expected.

Having understood the guidelines provided above for stating the educational objectives in proper terms, certain illustrations of objectives may be examined to see whether they are properly stated or not. Through this it may also be noted if there is any other point that should be borne in mind while stating the educational objectives. With this purpose in view examine the following objectives to find out whether they are properly stated.

A. The student solves perimeter problems.

B. The student understands perimeter problems.

Of these two, it may be concluded that objective 'B' is not properly stated as 'student's understanding' which may be interpreted in different ways by the teachers and, therefore, cause confusion regarding the expected learning outcomes. On the other hand, objective 'A' is properly stated as it provides the teacher with specific guidance for evaluation. He can actually observe students solving problems and check whether the objective has been achieved.

Again, examine the following three objectives and check whether they specify the expected learning outcomes in definite and clear terms.

(a) The pupil enjoys reading the poem

(b) The pupil recites the poem without any error

(c) The pupil writes the summary of the poem.

Again, objective (a) is stated as 'pupil enjoys' which cannot be observed directly and, therefore, lead to different interpretations regarding the learning to be expected. Objective (b) is specified as 'pupil recites'. This refers to an observable behaviour. Further it clearly specifies that the student recites the poem without any error. Thus the objective specifies the expected learning in proper terms. Objective (c) is also specified by 'pupil writes'. Thus it is stated in terms of pupil's behaviour, viz., writes, which can be observed. Yet, it does not specify the learning outcomes to be expected in proper terms. This point is well clarified in the following illustration.

> Suppose you have taught two pupils, 'X' and 'Y' with this objective, viz., 'Pupils write the summary of the poem'. After teaching you ask them to write the summary. You find that both 'X' and 'Y' have written the summary. But, 'X' has written it in 5-6 lines bringing out all the five ideas in the poem while 'Y' has written it in about 15 lines although he has not referred to all the ideas. Now, you have to decide whether 'X' or 'Y' has achieved the objective specified.

Obviously, in the above situation one cannot decide either way because, summary of a poem can be written in various lengths including all are some of the ideas depending upon expected level of learning or performance. And, the objective in consideration -- does not specify the expected level of performance as to the length or the ideas to be referred to in the summary.

From the above illustration, it may be understood that in order to specify the expected learning outcomes in definite and clear terms it is necessary that the objective is stated in terms of changes in the students' behaviour that can be observed and it should also clearly specify the expected level of performance.

In the light of the above the following two objectives may be examined.

1. Pupil writes the summary of the poem in about 10 lines referring to all the five ideas in the poem.

2. Pupil recites the poem without any error.

It can be easily observed that both these objectives specify the expected learning outcomes in definite and clear terms and therefore they are properly stated.

From this examination of illustrations of educational objectives, it may be concluded that a properly stated educational objective should be in terms of changes in the students' behaviour which can be directly observed. Further, the objective should clearly specify the expected level of performance.

The above discussions about learning outcomes and educational objectives clearly indicates the way in which learning outcomes are specified and the proper form in which educational objectives should be stated for purposes of evaluation in specific instructional situations.

Evaluation Techniques :

It was noted that, at the schools, through the evaluation programme information is obtained about the students' achievement in various school subjects and also regarding various other aspects such as their intelligence, aptitude, interest, attitude, adjustment, etc. And evaluation of all these various aspect of the students is carried out in the schools by using several techniques of evaluation.

By examining these techniques of evaluation, one would find that different techniques adopt different procedures for obtaining information about the students' learning (i.e. for determining the learning outcomes).

Consider the procedure by which teachers usually obtain information about their students' learning or achievement in different school subjects. Teachers generally determine the learning or achievement in different school subjects by giving tests. In other words, evaluation of achievement in different subjects is done using the technique of testing. Therefore testing can be considered as an evaluation technique. The procedure by which information is obtained regarding the students' achievement (i.e., learning outcomes) can be understood by examining a typical testing situation.

Suppose that a teacher has given a test in Physics. He, then, checks the answers written by different students. From their answers he finds out what the different students have learnt. That is, teacher determines the learning outcomes on the basis of the students' answers. It may also be noted that answers are written by the students according to what is required by the questions in the test.

One would, generally find following type of questions in a test.

1. Describe the phenomenon .....
2. Define the term .....
3. Complete the following.
4. Draw the diagram of .....

It can be observed in the above that the words 'Define', 'Describe', 'Complete', etc., refer to certain tasks to be performed by the students. In other words, a test requires the students to perform or do certain tasks. The answers to the questions in the test are, therefore, the products of the performance of the tasks required by the test. It has been earlier noted that in testing, learning outcomes are determined from the students' answers.

Therefore, finally it may be concluded that in testing, learning outcomes are determined from the product of certain tasks performed by the students.

In the above illustration, the product has been in terms of written answers. However, testing technique can also be used in other situations of evaluation as it may be observed in the following illustration.

> Suppose a drawing teacher asks the students to draw a particular object. From what the students have drawn (product) he decides about their learning.

Obviously, in this illustration testing technique has been used although the product is not in terms of written answers.

Tests are most commonly used by teachers for evaluating the achievement of the students in different subjects. However, testing technique is also used in the evaluation of intelligence and aptitude of the students.

Apart from testing, teachers use some other techniques also for purposes of evaluation in the schools.

> Suppose your teacher has trained you in using a particular instrument in the science laboratory. Now, he wants to check whether you have learnt to use the instrument properly (i.e. he wants to determine the learning outcomes). For this, he asks you to actually use the instrument in a proper situation and observe when you do so. Through his observation he gets the required information.

Analysing the above procedure it may be observed that the student actually uses the instrument. In other words, in this case also the student has to perform certain task. But, instead of going to the product of the performance, teacher gets the information through his observation of the performance itself. Since the teacher obtains the required information (i.e. learning outcomes) through his observation, this procedure of evaluation is called the technique of observation.

In the schools, technique of observation is used for the evaluation of achievement of specific skills, and also, in the evaluation of students' social and personal adjustment, etc.

One basic difference may be observed between the two techniques of evaluation, viz., testing and observation.

In testing, information about the student is based on the product of a task performed. On the other hand, in observational technique, it is based on the very performance of the task and not on any product.

However, it may be noted that both in testing and observation, information about the student's achievement in school subjects, intelligence, aptitude, adjustment, etc., is obtained indirectly as it is based either on the student's performance of a task or on its product.

Apart from these indirect methods, information about certain aspects of the student can be obtained directly from the student himself. In other words, for obtaining certain information about the student, the student may be asked to report about himself.

This procedure of obtaining information about the student, where the student is reporting about himself is called self-reporting technique. At the schools, self-reporting technique is used in the evaluation of such aspects as interests, attitudes, etc.

Thus, the following three techniques, which have been discussed above, are the major techniques of evaluation used in the school.

1. Technique of testing
2. Technique of observation
3. Self-reporting technique

Measurement of Learning Outcomes :

While trying to understand the process of evaluation it was noted that there are two main stages in it. The first of these is to

determine the larning outcomes. And, the second is to compare these learning outcomes with the educational objectives in order to determine the extent to which the educational objectives have been achieved.

Considering the second main stage in the process it was observed that in order to make the evaluation more accurate (i.e., to determine the extent of achievement of objectives more accurately) we may measure the learning outcomes before comparing them with educational objectives. Thus, 'measurement of learning outcomes' was specified as one of the five steps in the process of evaluation.

The following description provides answers to the questions : What is meant by measurement of learning outcomes and how is learning measured?

Learning outcomes specify changes in the students' behaviour. Therefore, measurement of learning outcomes would also be related to these changes in the students.

Consider measurement of length, volume, weight, etc. It may be noted that any change in these aspects always exists in certain specific amount or quantity. Therefore, in the case of learning outcomes also it may be said that changes in the students' behaviour exist in certain specific amount or quantity.

Thus, measurement of learning outcomes means the measurment of the specific amount of changes in the students' behaviour.

Again, in the measurement of length, volume, weight, etc., the amounts of changes are specified by assigning specific numbers such as 2 ft., 20 litres, 3 kg., etc. On the same lines, learning outcomes are

measured by assigning specific numbers which indicate the quantity of changes in the students' behaviour.

In simpler terms it may be said that in measurement of learning outcomes specific numbers are assigned to indicate the amount of learning that has taken place in the students.

Further, all are familiar with the procedure of giving marks (assigning numbers) to students based on their answers to the questions in a test. From the technique of testing it is known that learning outcomes are determined on the basis of the answers to the questions in a test. Thus, one would be actually measuring the learning outcomes while giving number or marks to the answers.

Teachers are frequently found saying after a test or examination that a particular student has got 82 marks while another has got only 40 marks and so on. Here the teacher may be considered to have measured the students' learning (i.e. learning outcomes) through the test.

> Consider the situation wherein the teacher has trained the students to operate a particular scientific instrument step by step. Then, he has allowed the student to operate the instrument and observes him when he does so. While observing he notes down the number of steps which the student properly carried out and those in which he goes wrong.

Here, the number of steps properly carried out by the student can be safely taken to represent the amount of learning in the student. Thus, this number gives the teacher a measure of how much has been learnt by the student.

Comparison of Learning Outcomes and Objectives :

After learning outcomes have been measured, they are compared with educational objectives to determine the extent to which objectives have been achieved.

It may be recalled that educational objectives can be considered as expected learning outcomes. Thus, for determining the extent of achievement of the objectives actual amount of learning has to be compared with the expected amount of learning.

Suppose in a testing situation, the learning outcomes have been determined and also been measured. As has been noted, the actual amount of learning that has taken place in the students is represented by specific numbers or marks such as 80, 45, etc. Also, in any test these marks like 80, 45, etc., assigned to the students are always considered as out of a maximum marks of 100, 50, etc.

It may further be observed that a student gets this maximum marks like 100, 50, etc., on a test if all the questions in the test are correctly answered by him. And, the different questions in a test correspond to the different objectives to be achieved, or in other words, to the expected learning outcomes.

Thus, it may be said that the maximum marks like 100, 50, etc., which a student gets by answering all the questions correctly specify the expected amount of learning in the students.

Now, learning outcomes are compared with the educational objectives by comparing the specific number or marks obtained by the students which represent the actual amount of learning with the maximum marks that can be obtained which represent the expected amount of learning.

This comparison of specific numbers or marks which is generally expressed as 80 out of 100, or 45 out of 50, etc., clearly specifies the extent to which the expected learning has taken place, or in other words, the extent to which educational objectives have been achieved.

It may be observed that in the above quantitative descriptions are obtained (i.e. in terms of numbers) for the students' learning (actual learning outcomes) and for the educational objectives (expected learning outcomes). By comparing these, again a quantitative description is obtained for the extent to which educational objectives have been achieved.

After this, the next specific step in the evaluation process represents the final aim of the process. This involves judging the extent of achievement of the objectives to decide whether it is very good, or good, or average, etc.

In other words, in the final step of the evaluation process students are judged by assigning such values as good, average, poor, etc. to the quantitative description of the learning that has taken place in them.

Role of Tests in the Evaluation Process :

Taking a close look at the evaluation programme as it is carried out in the schools, it can be observed that tests constitute the most commonly observable feature of any programme of evaluation. Because of this, it is sometimes talked about as though testing and evaluation mean the same.

But, it is known that evaluation is a process which aims at judging the student's learning as satisfactory or not and one comes across

testing only at a particular stage in carrying out this process.

Testing, as has been already noted, is one of the techniques of evaluation and thus corresponds to a particular procedure of determining the learning outcomes. Moreover, testing is suitable for obtaining information regarding only certain aspects of the students such as their achievement in school subjects, intelligence, aptitude, etc. In other words, tests are useful in the evaluation of only some aspects of development of the student.

Thus, it may be concluded as follows regarding the role of tests in the total evaluation programme.

Tests represent only one of the three techniques of evaluation used for determining the students' learning. And, testing is useful for the evaluation of only some aspects of development of the pupil while a comprehensive evaluation programme has to consider all aspects of development of the pupil.

Place of Measurement in the Evaluation Process :

Sometimes, one would find that evaluation process is considered in such a way as though it is only to assign numbers or marks to represent students' learning.

In other words, sometimes evaluation and measurement are considered to mean the same. However, it is clear that this cannot be the case as measurement is only one of the several steps in the process of evaluation.

In specific terms, it is the process of assigning specific numbers or marks to students' learning. But, evaluation process cannot be

taken to have been completed just by assigning numbers or marks. In order to evaluate students' learning these marks representing the actual amount of learning with the maximum marks representing the expected amount of learning. Further, in order to determine whether the pupil is satisfactorily progressing or not, teacher may compare the present achievement of the student with his previous achievements. He may also compare it with the average achievement of the class as a whole.

Of course, measurement makes all these comparisons more accurate as the comparisons would be between two specific numbers. In other words, measurement makes evaluation more accurate by providing a quantitative description of the students' learning.

From the above discussions it is clear that testing and measurement correspond to two particular steps in the process of evaluation. And, they should not be mistaken to represent the whole process of evaluation.

Measurable and Non-measurable Outcomes :

Learning outcomes are measured by assigning specific numbers to the quantity of change or learning in the students.

Consider the following situation and see whether the teacher can specify the learning in the students in quantitative terms.

> Suppose a teacher is teaching handwriting. Now, he wants to check whether there is any improvement in the student's handwriting. For that he asks the student to write and checks his writing. He finds that the student has improved his handwriting.

But, it is obvious that the teacher cannot assign a specific number to the improvement that has taken place. This is because, improvement is to be seen only in the quality of handwriting and not in terms of quantity.

In other words, with respect to this particular aspect, the change or learning in the student can be expressed only in qualitative terms and not in quantitative terms. Since in this case it is not possible to obtain a quantitative description of students' learning, it may be concluded that the particular learning outcome cannot be measured.

In fact, it may be found that there are other aspects of learning also such as appreciation, adjustment, dramatisation, etc. for which one cannot obtain a quantitative description of the learning. Thus, it may be concluded that some of the learning outcomes are not measurable.

It has been noted earlier that in the final step of the evaluation process the quantitative description of the student's learning is judged and assigned values like good, poor, satisfactory, etc. But, it has been just now discovered that certain learning outcomes are non-measurable for which only a qualitative description of the student's learning can be obtained.

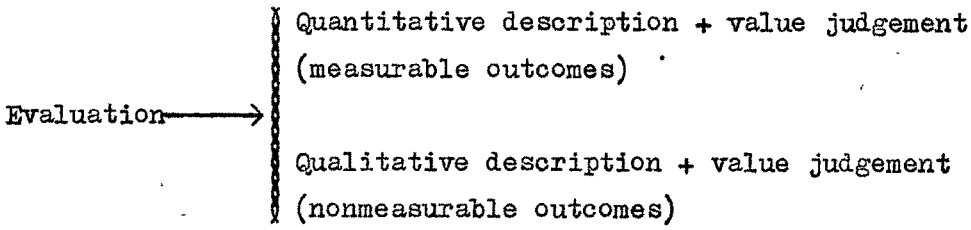
Therefore, in case of non-measurable outcomes value judgement (i.e. judging and assigning values) has to be made on the basis of the qualitative description of the students' learning.

Thus, it has been found that some aspects of learning or development are measurable and some others are nonmeasurable. But, the evaluation programme should consider all aspects of learning.

Therefore, evaluation should take into consideration both measurable and nonmeasurable learning outcomes.

This implies that in evaluation, for some aspects of learning the value judgements would be based on the quantitative description of students' learning and for others it would be based on a qualitative description.

Combining both the above aspects, the evaluation programme may be represented as follows.



III. Characteristics of a Good Instrument of Evaluation

From the discussions in the previous sections it has been learnt that the evaluation programme in the school takes into consideration the pupil as a whole, i.e., all aspects of behaviour or development of the pupil. And, results of evaluation are useful in various ways for teachers, pupils and others involved in school education.

It has also been noted that different evaluation techniques are used depending upon the objectives with reference to which evaluation has to be done. In other words, while selecting a specific evaluation procedure one has to consider its suitability for the particular aspect of development of the pupil that has to be evaluated.

From the above observation, it becomes clear that certain important points have to be looked into while selecting a specific procedure of evaluation. For instance, first of all it has to be checked whether the specific procedure really provides information about that particular aspect which is wanted. Also, one has to see whether the same specific procedure can be adopted for evaluating a particular aspect of the student, at different times. And, even if the above questions are answered satisfactorily, it is necessary to think whether the specific procedure chosen is usable in the school situation and so on. In short, it may be said that while selecting a specific procedure of evaluation it has to be checked to see whether it possesses certain essential characteristics.

Thus, in this section, attempt will be made to answer the following questions.

--- What are the characteristics of a good evaluation procedure?

--- What factors affect these characteristics?

Further, as one proceeds to answer these questions, it would be found that these characteristics are not independent. Thus, attempt will be made to answer the following question also.

---- How are these different characteristics related to each other?

Instruments of Evaluation

Evaluation programme in the school should take into consideration all aspects of development of the pupils. And, in order to evaluate different aspects of development, one has to first obtain information regarding these aspects of the student.

In this context, it was noted that different evaluation techniques are used to obtain information about different aspects of development of the pupil.

Therefore, in the evaluation process, a teacher has to select and apply a specific evaluation technique suitable for the particular aspect of development of the pupil that has to be evaluated. Regarding the selection of the suitable technique it has already been learnt that, mainly, three evaluation techniques are used in the schools. Of these, technique of testing is suitable for evaluation of certain aspects like achievement in various school subjects, intelligence, aptitude, etc. And, the other two techniques, namely, observation technique and self-reporting technique are suitable for evaluation of certain other aspects like personal-social adjustment, attitude, interest, etc.

Thus, the question that remains is how are these techniques applied. This question may be answered by considering how technique of testing is applied in schools.

In fact, it is known that testing technique is applied in schools using certain tests like achievement test, intelligence test, etc. Thus 'tests' represent materials used for applying technique of testing in the evaluation process.

In other words, it may be simply said that 'tests' represent materials used for the purpose of evaluation of certain aspects of development of the students. These 'tests' as they are used for the purpose of evaluation are called tools or instruments of evaluation. Thus, technique of testing is used in the evaluation process with the help of certain instruments of evaluation called 'tests'.

Similarly, under the techniques of observation and self-reporting, certain other instruments of evaluation like rating scales, checklists, questionnaires, inventories, etc., are used.

Therefore, it may be said, in general, that different techniques of evaluation are applied with the help of various instruments of evaluation. In other words, various instruments of evaluation like achievement tests, intelligence test, rating scales, questionnaires, inventories, etc., are used in the school evaluation programme for obtaining information about different aspects of development of the pupils.

Further, it is known that the final results of evaluation are based on the information obtained through these instruments of evaluation. Therefore, in order to get accurate results of evaluation one should select and use a proper instrument of evaluation. This brings up the question : In order to get accurate results, what essential characteristics should be looked for in it while choosing an instrument of evaluation?

Validity

It has been noted that evaluation results are used for different purposes in schools; and, that these results are based on the

information obtained through different instruments of evaluation. Thus, different instruments are used for different purposes of evaluation at schools. In other words, this implies that specific instruments of evaluation have to be selected for serving specific purposes of evaluation.

From the above it is clear that while selecting an instrument of evaluation the first characteristic one has to look for is whether it serves the purpose for which it is being used.

In daily life, certain things like certificates, licence cards, entrance tickets, etc., are used for certain specific purposes. And they are considered as valid if they serve the specific purpose for which they are being used, and as not valid if they do not.

Similarly, an instrument of evaluation is considered to be valid if it serves the purpose for which it is being used. Thus, the first characteristic an instrument of evaluation should possess is that it should be valid. An instrument which is valid is said to possess the characteristic of validity.

Types of Validity

Since evaluation instruments are used for different purposes in the schools, one comes across different types of validity. In fact, generally four types of validity are identified in educational measurement.

These four types of validity will be discussed in the following. Each type will be discussed as it relates to testing procedures in schools. However, these four types of validity are applicable to all instruments of evaluation.

Content Validity

A particular purpose for which tests are used is to obtain information which indicate the extent to which students have learnt specific content material in different school subjects. Thus, when a test serves this particular purpose, it may be said that the test has validity with reference to the purpose of determining students' learning of specific content material. It may be observed that this particular type of validity is mainly related to the content area that is being considered by the test. Accordingly, a test having this type of validity is said to possess content validity.

Thus, a test may be said to possess content validity to the extent the information obtained through it indicates student's achievement with respect to specific content areas of learning.

Predictive Validity

Sometimes, teachers use test performance of students for the purpose of drawing conclusions about their future learning ability in the particular area tested. That is, they expect information obtained through the particular test to serve the purpose of predicting future performance of the students.

Thus, when information obtained through a test serves this particular purpose of predicting future performance it may be said that the test has validity with reference to this particular purpose. Since this type of validity of a test is based on the predictive value of the information obtained through it, it is referred to as predictive validity.

Thus, a test possess predictive validity to the extent the information obtained through it serves the purpose of predicting the future performance of the students in a particular area of learning.

In fact, a specific type of tests called aptitude tests are specially meant for the purpose of predicting the future performance of the students in specific areas of learning. For instance, the results obtained through an aptitude test in Mathematics should enable the teacher to predict, somewhat accurately, the future success of the student in learning Mathematics.

Concurrent Validity

In daily life, quite often, time shown by a watch is compared to that given over the radio or by another watch which is known to show correct time. And, the time shown by that watch is taken to be correct or valid to the extent it is same as that given over the radio or by the watch which is known to show the correct time. In other words, if there are two instruments measuring the same aspect, say time, then, one naturally expects the time shown by them to be same or concurrent.

Similarly, suppose a teacher has prepared a test for the purpose of measuring a particular aspect. And, he finds that there is already another test which is known to measure the same aspect accurately. Then, he naturally, expects the results obtained through the test prepared by him to be concurrent with those obtained through the other. Therefore, a test measuring a particular aspect is considered valid to the extent it provides same information as that obtained through another test which is known to measure the same aspect accurately.

This refers to another type of validity of a test. And, this type of validity is based on the extent to which the results obtained through a test are concurrent with those obtained through another test in current use for measuring the same aspect. This type of validity is called concurrent validity.

Thus, a test may be said to possess concurrent validity to the extent the information obtained through it is concurrent with that obtained through other tests measuring the same aspect.

Construct Validity

Tests are also used to measure certain psychological characteristics of the student such as his intelligence, creativity, abstract thinking, motivation to learn, anxiety, persistence, etc. In these cases, test results are taken to indicate the degree to which a student possesses these psychological characteristics. This, therefore, represents another purpose for which test results are used in the schools.

These characteristics are abstract in nature and are defined in terms of certain psychological concepts or principles. Therefore, these characteristics are also referred to as psychological constructs. Thus, it may be said that certain tests are used with the purpose of finding out the degree to which a student possesses certain psychological constructs. This may be considered to represent another type of validity wherein the purpose corresponds to the use of test results for measuring certain psychological constructs. This type of validity is called construct validity.

But, it has been noted that these constructs are themselves abstract and they have to be understood in terms of certain psychological concepts or principles.

Consider for instance the construct, intelligence. Regarding this construct, a principle that is known, is that the degree of intelligence possessed by a child increases as his age. Therefore, if an intelligence test is administered to a boy at different age levels and the results are noted, these test results should indicate that intelligence level increases as his age level increases. In fact, there are many other psychological concepts and principles underlying the construct intelligence. And, therefore, the results of an intelligence test should be explainable in terms of all these psychological concepts and principles.

Thus, in general, the result of a test of any psychological construct should be explainable in terms of the psychological concepts underlying that particular construct.

Therefore, a test can be considered as valid for measuring a psychological construct if its results can be interpreted in terms of the concept and principles underlying that construct. In other words, it may be said that a test of any psychological construct possesses construct validity to the extent results obtained through it can be interpreted in terms of the known psychological concepts and principles about the particular construct.

The four types of validity identified in the above discussions are : 1. Content Validity 2. Predictive Validity 3. Concurrent Validity and 4. Construct Validity.

It may be observed that in the above each of the four types of validity has been identified with reference to a particular purpose for which results of evaluation are used. Thus, whenever the validity of an instrument is discussed, it must be with reference to a specific purpose for which the information obtained through it has to be used.

In other words, information obtained through an instrument, say, an achievement test in Mathematics, may very well serve the purpose of indicating a student's achievement in certain content areas of Mathematics. But, it may be of little use for predicting his future performance in Mathematics. That is, an evaluation instrument may possess high content validity, but, its predictive validity may be very low. Similarly, results of a numerical aptitude test may have predictive value for predicting a student's success in mathematical computation. But, they may be of limited value for predicting the student's mechanical ability. Thus, in selecting a particular evaluation instrument, one should mainly look for that particular type of validity which refers to the particular purpose for which the results obtained through it are to be used.

One often comes across selection tests being conducted for selecting persons for new courses of study or for new vocations (jobs). Based on the results of these tests those persons who will succeed well in the new courses will be selected. In choosing such selection tests one has to mainly consider their predictive validity. It is quite obvious that while selecting a person for a new course or job the main interest would be in predicting his future performance. And, therefore, results of the selection tests have to serve this purpose of prediction.

Periodical tests and examinations are conducted in various school subjects. The main purpose of these tests is to measure student's achievement in specific content areas of learning. Therefore, these tests and examinations are expected to possess content validity. And, they may or may not be suitable for predicting the future performance of the student.

However, at schools and colleges, quite often students are selected for different branches of study mainly on the basis of their performance in previous achievement tests. Thus, in these situations, achievement tests are assumed to be possessing predictive validity although it is known that they may not serve the purpose of predicting future performance.

Another important point may be observed regarding different types of validity. An evaluation instrument is considered to possess a particular type of validity if the results obtained through it serve the purpose for which they are to be used. For instance, an achievement has to, mainly, possess content validity. And it will be considered to be possessing this type of validity if results obtained through it serve the purpose of evaluating students' achievement in specific ontent areas of learning.

But a question that arises at this point is : What should be the nature of a test, if the results obtained through it should serve the purpose of evaluating student achievement?

Answer tto this question is discussed in the following.

At the school, specific oontent areas of a subject are taught with a view to achieving certain specific educational objectives. Therefore, results of achievement tests are used to determine the extent to which all these specific objectives have been achieved. In other words, an achievement test in a particular content area may be considered to possess content validity if it serves the purpose of determining the extent of achievement of all the specific objectives related to that particular content area.

> Suppose your teacher has taught you a particular unit in science, say on water, with some fifteen specific objectives. Then, he prepares and uses a test containing questions related to all these fifteen specific objectives. Obviously, this test can be considered to have content validity.

Thus, a test of a particular content material will have content validity if it contains questions relating to all the specific objectives of that particular content material.

> On the other hand, suppose you, as a teacher, have to conduct a comprehensive test for a complete course containing several units. In this case, since the content material is large, corresponding specific objectives will also be large in number. And, obviously, you cannot have questions covering all the objectives. In such a situation, you would naturally prepare a test with questions referring to a sample of objectives selected from the total set of objectives for the complete course.

Thus, in preparing a comprehensive test for a course with large content material one has to select a sample of objectives from the set of objectives for the complete course. In fact, it may be observed that the question papers for the final examinations have to be of this type.

> Suppose a teacher has taught an elementary science course consisting of four units. Then, he prepares a comprehensive test for the complete course. But,

> by studying the questions included in the test you find that they correspond only to a sample of objectives related to the first three content units of the course. And, there is no question related to the last unit. Obviously, this cannot be a valid test as the sample of objectives taken for testing does not represent the last content unit of the course.

Thus, for preparing a valid test, the teacher has to first choose a sample of objectives which represent all the content units of the course.

Many times, students would be found complaining that the traditional examinations test only their memory and not their learning. This, in fact, indicates that these examinations do not have content validity.

> At schools, several facts, principles, etc., are taught to the students. But of these, students are expected merely to know certain points which they can reproduce by memory. Certain other points are meant to be understood by them; also, they are expected to be able to apply certain other aspects of their knowledge; and so on. Accordingly, the total set of objectives for a course would contain objectives of all levels, viz., knowledge, understanding, application, etc.

Thus, the examinations which require the students only to reproduce certain points from memory contain questions representing only knowledge objectives.

From the above, it is clear that a valid examination of student's achievement should include questions related to objectives of all levels included in the course. In other words, in order to prepare a valid question paper for a comprehensive examination one has to choose a sample which includes objectives of all levels.

> Consider another situation where a comprehensive test has been prepared with questions relating to all content units of the course. But, you may find that all the questions in the test are very easy. Or, you may find that all of them are very difficult. Obviously, this cannot be a valid test as the results of such a test would not evaluate student's achievement of topics of all difficulty levels.

Thus, it may be said that a valid test should include questions of all difficulty levels.

The points discussed above specify the qualities that are essential for an achievement test to be valid.

Reliability

Various instruments of evaluation are used to measure different aspects of the students. For instance, an intelligence test is used to measure the intelligence of a student. Also, it has been noted that measuring the intelligence of a student involves obtaining a number or score which is taken to indicate the intelligence of the student.

In other words, one relies upon the score assigned to a student's performance on the test as the indicator of this intelligence.

Similarly, suppose an achievement test in Mathematics is administered and it is found that the student gets a score of 60. Then, again, this score will be considered as a reliable indicator of the student's actual test performance.

In fact, it is known that it is based on this score that the final evaluation of the student is done. Thus, for proper evaluation it is necessary that the scores obtained by the students are reliable.

When the scores obtained through a test are reliable, the test is said to be reliable, or that it has reliability.

Or, in general, an instrument of evaluation is considered to have reliability if the results obtained through it are reliable for indicating actual amount of change in the particular aspect of development of the student.

Thus, another essential characteristic of a good evaluation instrument is that it should have reliability.

When will a test be reliable?

It has been noted that the student's answer to the questions in a test are taken to represent the actual learning outcomes. These outcomes are then measured by assigning numbers or scores to them. And, the sum of these scores is the total score for the student which is considered to specify the actual amount of learning.

Thus, the total score of a student will represent the actual amount of learning if their answers can be taken to represent actual learning outcomes.

Therefore, a test or a student's score on a test is reliable if the answers given by the students to the test questions represent actual learning in them.

> Suppose in a test several of the questions are such that the students may simply answer them by guessing. And, by chance a student gets almost all his guesses correct and thereby he gets a total score of 65.

Obviously, this test cannot be a reliable one as a part of the student's score does not represent actual learning in him. In fact, this is because of the error made during the measurement of learning by giving marks to answers which do not represent actual learning in the student.

> Suppose in another case a student knows answers to all the questions. But, the language used in the questions is such that he cannot understand what he has to do. Thus, he writes wrong answers and gets a low score on the test.

Again, the student's score does not represent the actual amount of learning as it does not include the amount of learning corresponding to those questions which he did not answer properly although he knew the answers. Here, also, it is because of the errors involved in the measurement of learning. But, it may be noted that in both the cases, the errors of measurement were brought in by the nature of the test itself.

Thus, a test is said to possess reliability to the extent the scores obtained through it represent actual amount of learning in the students and they are not due to the errors of measurement brought in by the nature of the test.

How to check the reliability of a test?

Suppose that the total score of a student on a test has been obtained. Now, it has to be checked whether the score is reliable or not i.e. it has to be checked whether there is any error of measurement involved in obtaining the scores or they represent the actual amount of learning in the students.

Thus, one has to check whether he has gone wrong in the measurement of learning.

Test-Retest Method

> Consider the following instance. Suppose you have multiplied two large numbers, and got their product. Now, you want to check whether you have gone wrong in the multiplication i.e. whether the number obtained by multiplication is the actual product. Obviously, a simple way is that you do the multiplication again. You may consider that there was no error involved in multiplication if the product obtained through remultiplication is the same as the one obtained earlier.

On the same lines, in order to check whether any error of measurement is involved one may measure the learning again. That is, if one has got a set of scores on the test, it may be checked whether it is reliable or not by getting another set of scores by giving the test again.

And, the test is considered to be reliable if the scores obtained through the test and the retest situation are the same.

In other words, a test is considered to be reliable if the scores obtained through it are consistent when administered at two different times.

Thus, the reliability of a test is specified by the consistency of the scores obtained through it over a period of time.

As the reliability of a test is here checked by comparing the scores obtained through test and a retest situation, it is referred to as test-retest reliability.

Equivalent Forms Method

Consider an achievement test in Mathematics consisting of several problems. Suppose one of the problems is : "Add 283 and 386" which corresponds to the objective : Student adds three digit numbers correctly! Now if the student's answer represents the actual achievement of the objective then the student must also be able to solve an alternate problem of the same type.

Similarly, suppose a test has been developed with an alternate problem to each problem on the first test. That is, an equivalent form of a test is derived. Then, one would expect the scores of the students on the two forms to be same.

In other words, if the student's score on a test represent actual learning in him then he should get a similar score on an equivalent form of the test.

Thus, a test may be considered as reliable if the scores obtained are consistent over two equivalent forms of the test. Here, reliability is specified in terms of consistency of scores over two equivalent forms of the test.

Split-half Method

As it has already been noted, all the questions in a test together try to measure learning related to a particular aspect. Therefore, if the test is split into two halves, they would represent two equivalent forms. And, each of them will still be measuring the same aspect.

Thus, if the score on the two halves of the test are separately taken and compared, one should find that each score is representing the same amount of learning related to the particular aspect being measured. This implies that the total scores of the students on each half should be consistent.

Thus, here, the reliability of the test is specified in terms of the consistency of scores over two halves of the test.

As, in this case, reliability is checked by splitting the test into two and considering the score on each half, it is referred to as split-half reliability.

It may be noted that in all the three approaches of checking the reliability, a test's reliability is specified in terms of the consistency of scores obtained through it.

Thus, reliability is another essential characteristic of a good instrument of evaluation.

An instrument of evaluation is considered to possess reliability if the information obtained through it represents actual amount of the particular aspect that is being measured through it and is not affected by the errors of measurement.

Thus, the reliability of an instrument of evaluation decreases as the information obtained through it becomes less reliable due to the errors of measurement.

Objectivity

As it has been noted earlier a student's total score on a test should be based on his answers which represent the actual learning outcomes. Under reliability, it was discussed as to how, sometimes, this score may not be a reliable indicator of the actual amount of learning in the student if his answers do not represent actual learning in him. In other words, when the students' answers do not represent actual learning outcomes, the amount of learning indicated by the student's total score may be greater or less than the actual amount of learning.

Suppose a teacher has conducted a test and he is checking the answers and assigning marks to them.

Now, although it is the teacher who assigns the marks, a student's marks on the test should depend only upon the answers he has given and not on the scorer who is scoring them.

When the student's score on a test is based only on the student's answers and not the screen who is scoring, then the scoring is considered to be objective.

It has already been noted that the total score should always be based only on the student's answers. Therefore, if a student's score has to represent the actual amount of learning then the scoring of answers on the test should be objective. That is, objectivity in scoring is essential for proper evaluation.

Suppose a test has been administered to a student. And, several competent teachers have checked and scored the answers, independently. Also, suppose that the scoring done is objective in all the cases. Then, in all the cases the total score assigned to the student should remain the same. That is, the total scores assigned by different scorers should be consistent.

Thus, objectivity in scoring may be considered as consistency in scoring by different scorers.

Now, consider the following instance of testing. Let one of the questions in a test be : 'Discuss in detail the meaning of the term 'Reliability of a test'. Suppose, several equally competent teachers have independently scored a student's answer which contains several important points about the meaning of 'Reliability of a test'.

Obviously, experience would suggest that one cannot expect the different teachers to give same marks for an answer to such a question. That is, scoring of answers to such a question may not be objective.

'Because, teachers differ in their opinion about the correctness or suitability of the answer given by the student.'

Thus, it may be observed that in this case the scoring is not objective because, it is based not only on the student's answers but also on the opinion of the teacher who is scoring them.

Suppose, the same set of teachers are scoring the same answer in another situation. Here, the same answer has been given with reference to the following question : 'Explain in about 8-10 lines the meaning of 'Reliability of a test'. What are the different types of validity? Specify each type in 4-5 lines'.

Obviously, in this case, the scores assigned by different teachers would not differ so much as in the earlier situation.

Thus, the same answer has been scored more objectively in this situation.

'Because, the question, in this case, clearly specifies the correct or suitable answer that should be expected.'

It may be observed that the same answer was not scored objectively, earlier, as in that case, the question did not specify the exact answer to be expected.

Thus objectivity in scoring depends upon the nature of the question in the test.

When the questions in a test are such that the scoring of the answers can be objective the test is said to be objective; or, that the test has objectivity.

Thus, the objectivity of a test may be defined as the extent to which a student's score on it is based on his actual answer or performance on the test and not on the opinion of the scorer.

It was noted that for proper evaluation, scoring should be done objectively. It was also found that objectivity in scoring depends upon the nature of the test. Thus, it may be said that objectivity is an essential characteristic of a good test.

Or, it may be said, in general, that objectivity is an essential characteristic of a good instrument of evaluation.

An instrument of evaluation is considered to be objective to the extent results obtained through it depend upon the actual information provided by the student and not on the opinion of the person who is using it.

In fact, it may be noted that results obtained through certain instruments such as rating scale, checklist, etc., are always influenced by the opinion of the observer. Even under testing, one cannot expect objectivity in scoring of answers to essay type questions.

Thus, certain instruments of evaluation are less objective by their very nature. And, they may be selected for use only if a more objective instrument is not available or if no other instrument serves the specific purpose.

It was observed that objectivity in scoring refers to the consistency in the scores assigned by different teachers.

Thus, in other words, objectivity of a test is the consistency in the scores assigned to the answers on it by different scorers.

It may be recalled that reliability is defined in terms of the consistency of scores. And, objectivity refers to the consistency among different scorers in assigning scores to answers on a test. Thus, objectivity is also referred to as inter-scorer reliability.

Relationship between Validity, Reliability and Objectivity

Validity, Reliability and Objectivity are the three essential characteristics of a good instrument of evaluation.

Although these characteristics have been identified separately, it should be noted that all the three characteristics have to be looked for in every instrument of evaluation. Therefore, it is essential to know whether these characteristics exist in an evaluation instrument independently or they are related to each other. In this connection certain specific questions such as the following may be answered. Can there be an instrument of evaluation which is perfectly valid, perfectly reliable, and perfectly objective? Can a tool be valid without being reliable and objective?

Various evaluation tools can be used in the schools for obtaining information about changes in different aspects of the students. Some of these like intelligence test, personality inventory, etc., may be readily available. Certain others such as achievement tests, checklists for observation, questionnaires for obtaining special information, etc., have to be prepared by the teacher himself.

In any case, before using a specific instrument for evaluating the student, the teacher has to obtain satisfactory answers to certain basic questions about the results that would be obtained through that instrument. These questions are :

1. Will they (results) serve the purpose for which he wants to use them?

2. Will they tell him the actual change in the particular aspect of the student?

3. Will they be influenced by his personal opinion?

These questions, in fact, correspond exactly to the three characteristics one should look for in the instrument of evaluation, viz., validity, reliability and objectivity.

It may be observed from the definition of these characteristics, that each of the three terms has been defined as the extent to which the respective characteristic is possessed by a particular evaluation instrument.

Thus when a teacher tries to know about an instrument's validity, reliability and objectivity, he would, in fact, be finding out the extent to which the particular instrument is valid, reliable and objective.

The above fact indicates that any instrument of evaluation would be valid, reliable and objective only to a particular extent which may be high or low.

For instance, suppose that an achievement test has been prepared with much care and the teacher wants to use the results of the test for determining the extent to which specific educational objectives of the course have been achieved. But, as has been noted earlier, the questions in the test would correspond to only a sample of the total set of objectives. Thus, the test results do not serve the purpose as regards to those objectives which are not included in the sample of objectives.

Therefore, it may be said that if proper care is taken the teacher will be able to prepare a highly valid test but not a perfectly valid one.

Thus, when one looks for these characteristics in an instrument of evaluation, he has only to check the degrees to which they are in it; and, not whether the characteristics are possessed by the instrument to a perfect extent.

Suppose that a test has been prepared for the purpose of determining student's achievement in a particular area. But, it is found that several of the questions in it are such that students may not be able to understand what they should write. That is, student's score on the test cannot be considered to represent actual learning in him. In other words, reliability of the test is very low.

Obviously, results of such a test do not serve the particular purpose for which they are to be used. Since the test results do not serve the purpose for which they are to be used, the test may be considered to have low validity.

In fact, when the test is less reliable the student's score on it is also less reliable and therefore, it does not represent actual learning in the student. Therefore, naturally, the score cannot be taken as a valid indicator of student's achievement.

Thus, if the reliability of a test is low then its validity will also be low. Also, if a test has high validity it would imply that it has high reliability also.

On the other hand, suppose that the reliability of an achievement test is found to be high. But, it is not known whether the questions in the test correspond to a representative sample of objectives of the complete course or not. That is, nothing is known about the test's validity.

It should be noted that high reliability of a test does not ensure anything about the test's validity. This is because, it is quite possible that the test questions are small in number and do not represent certain content areas at all.

Thus, if a test is highly valid it implies that it is highly reliable also. But, the high reliability of a test does not ensure high validity for it.

Again, suppose that in a test, many of the questions are such that the answers cannot be scored objectively. That is, the test's objectivity is very low. In this case also it is obvious that the results obtained through such a test cannot serve the purpose of measuring student's achievement satisfactorily.

This indicates that if the objectivity of a test is low then its validity will also be low.

From the above observations it is clear that for a test to be valid it must be reliable and objective. But, reliability and objectivity of a test do not tell anything about its validity.

While discussing about objectivity it was noted that both reliability and objectivity of a test refer to the extent to which the total score of a student on the test remains consistent under differing conditions.

Thus, in a way, objectivity may also be considered as a particular type of reliability which refers to the consistency of scores assigned by different but equally competent persons.

From the above discussions it is clear that the three essential characteristics of a good instrument, viz., validity, reliability and objectivity are not seperate qualities; and, they are related to each other.

Usability

Suppose that a teacher has chosen a tool of evaluation taking into consideration all these three characteristics. But, an important question would be whether the tool is usable in the school?

Thus, while selecting a particular evaluation instrument it is necessary to consider the usability of the particular tool in the school situation.

Usability of a tool refers to various practical aspects involved in the use of the particular tool. Particularly in schools with limited resources, consideration of this factor becomes essential.

It is known that some tests would require the user to have got proper training in administering them. If the school does not have properly trained people administration of such tests may not be possible.

Thus, ease of administration in using a tool is a practical factor that should be taken into consideration in selecting a tool of evaluation.

Also, certain tools such as those involving projective techniques, and certain personality measures involve special and somewhat complicated scheme of scoring. This again shows that ease of scoring is another practical aspect that should be considered in selecting a particular instrument of evaluation.

It is known that except for achievement testing evaluation is generally not considered as a part of the day to day school schedule. Thus, time requirement for the preparation and use of the particular instruments have to be borne in mind while selecting them for using in the school.

Another aspect that should be considered is the cost involved in using the particular tool.

Thus, while selecting specific instruments of evaluation, apart from validity, reliability and objectivity, it is also necessary to consider certain practical aspects involved in using them.

These aspects are 'ease of administration', 'ease of scoring', 'time requirement' and 'cost involved'. These practical aspects refer to the usability of a tool of evaluation in school situation.

Therefore, it may be concluded that the essential characteristics one should look for while choosing a particular instrument of evaluation are : (1) validity, (2) reliability, (3) objectivity, and (4) usability.

IV. Major Evaluation Tools and Their Uses

It has been learnt in the previous sections that a variety of tools are available to the teacher for the purpose of evaluating the students. And, it is through these tools that different techniques of evaluation are applied in the evaluation process.

This section will be devoted to know about the major tools of evaluation that come under the three main techniques of evaluation used in the schools. Attempt shall be made to answer the question.

--- What are the major tools of evaluation used in the schools?

It was learnt in the last section that these tools of evaluation differ in their characteristics and the teacher has to choose the proper ones which satisfy his requirements. In fact, it was noted that certain tools, by their very nature, are less objective. Thus, attempt will be made to know, in brief, about the nature of the different tools of evaluation. In other words, answer will be obtained, in brief, to the question.

--- What are the characteristics of the different tools of evaluation?

It was also noted that different instruments of evaluation have to be used for evaluating different aspects of development of the student. And, the teacher has to select or prepare tools suitable for evaluating specific aspects of development. In this connection the suitability of specific tools of evaluation for the evaluation of particular aspects will be discussed. Thus, attempt will be made to answer the question.

--- What are the uses of the different tools of evaluation?

The different tools of evaluation can be categorised under the three main techniques of evaluation. They will be discussed accordingly under three main heads, namely, testing techniques, observational techniques, and self-reporting techniques.

Evaluation of any aspect of development of the student requires that we obtain relevant and accurate information about the particular aspect. In fact, it is based on this information that it is finally judged whether changes in the particular aspect of development are satisfactory or not.

It has also been noted that three main procedures are adopted in schools for obtaining this information. These procedures are represented by the three techniques of evaluation used in the schools.

Further, these techniques of evaluation are applied in the evaluation process with the help of different instruments of evaluation.

Thus, different tools of evaluation may be classified under these three techniques of evaluation. However, it should be noted that this classification is not very rigid. Certain instruments may be used under more than one technique of evaluation.

Another important point is that the tools need not be always represented by materials like tests, rating scales, or questionnaires, etc. Rather each tool represents a specific procedure of obtaining information under the particular technique. Thus, a tool and a technique may not be considered as two seperate things.

For instance, conducting an interview for undestanding the students problems is a specific procedure used under the category of self-reporting techniques. But, in the interview you may not use any material like a test, a rating scale or a questionnaire. Thus, it may be said that in certain situations, techniques of evaluation may be directly applied without the help of any instrument such as test, questionnaire, etc.

In the following discussion about tools under different techniques of evaluation, such specific procedures shall also be discussed about.

As it was said earlier one may discuss about the different instruments of evaluation, as they fall under the three techniques of evaluation which represent the main procedures of obtaining information for the evaluation of the different aspects of development of the students.

Testing Techniques

It is known that information for evaluation may be obtained by requiring the student to perform a given set of tasks. Products of the tasks performed would provide the information needed for evaluation. This procedure of obtaining information for evaluation represents the technique of testing.

One may observe that in this technique the first step is to get the student perform certain task, and then to obtain the needed information for evaluation from the product of the task performed.

Tasks can be of different types. Accordingly, the products of different types of tasks will also be in different forms.

For instance, a boy is asked in the craft class to prepare a card board box. Here, it may be said that the student is required to perform the task of preparing the box. It may be noted that this task involves the manipulation of certain objects like cardboard pieces. Further, products of the task carried out will also be in the form of an object, namely, cardboard boxes.

Thus, in this instance, the task is of a particular type which involves manipulation of objects. This results in a product which is also, generally, in the form of an object.

Consider another instance where the teacher asks the students to write an essay. In this case, the students are required to perform the task of writing an essay. And, the product of the task will be in the form of written material. In this case, the task is of a particular type which involves writing certain things resulting in a product which is also in the written form.

Suppose a teacher in the elementary school is teaching spelling. He calls out a word and asks a student to spell out the word orally.

Again, it may be said that the student is asked to perform a task. The task to be performed is spelling out the word orally. It may be observed that in this case the task is carried out orally. Also, the student's answer, which is the product of the task is in the oral form. Thus, this instance represents a type of task which is carried out orally and also results in a product in the oral form.

From the above illustrations, one finds that at schools the student may be required to perform different types of tasks resulting in products of different form.

As can be seen the testing technique only specifies that certain tasks should be performed resulting in certain products which (the product) may be, then, used as the basis for evaluation. Therefore, one can have different types of testing procedures involving different types of tasks resulting in products of different forms. In other words, different types of tests can be

classified on the basis of the nature of the task to be performed and the nature of the resulting product.

The three illustrations discussed above represent three types of tasks and corresponding product forms. Accordingly, the tests may be classified into three types.

The three illustrations may again be examined to see the three types of tests represented by them.

Consider the illustration about teaching of spelling. Here, it has been noted that the student is required to give oral answers. Suppose, in this case, the teacher evaluates the student's proficiency in spelling by considering the oral answers (i.e. the product). It may be said that the teacher is applying the testing technique. Obviously, it is testing as it involves the performance of a task resulting in certain product on the basis of which the student is evaluated.

Tests of this type where the students are required to give oral answers (i.e. to perform an oral task) and evaluation is done on the basis of these oral answers (products) are called oral tests or examinations.

Consider, now, the illustration about writing an essay. Here the student is required to give written answers. Again, suppose that the teacher evaluates the student's knowledge related to the topic of the essay on the basis of his written answer (the product). Then, the teacher is using the testing technique for the evaluation of the student.

Tests of this type where the students are required to given written answers (i.e. to perform the task of writing answers) and the student is evaluated on the basis of these written answers (products) are called written tests or examinations. These are also called paper-pencil tests.

Now, consider the other illustration which has been noted, earlier. It was observed that, in this case, the task to be performed by the students involves the manipulation of certain objects resulting in products which are also in the form of objects, namely, card-board boxes. Here, also, if the teacher evaluates the student's ability on the basis of the cardboard boxes prepared by him, it may be said that the teacher is using the testing technique.

Tests of this type are known as 'performance tests'.

It may be noted that in the other two situations, viz., oral tests and written tests, the task involves the manipulation of words. In other words, it involves manipulation of verbal material. On the other hand, in the present situation, the task involves the manipulation of objects, i.e., the manipulation of nonverbal material.

Also, the products in the situations of oral and written tests are in terms of words, i.e. the producfs are verbal material. On the other hand, product in the present situation is in terms of objects, i.e., in terms of non-verbal material.

Further, in oral and written tests, the tasks do not involve much of physical or motor activities. But, in the case of performance tests the tasks, mainly, involve physical or motor activities.

Thus, performance tests are those wherein the tasks to be performed by the students require manipulation of objects, i.e., non-verbal material and involves motor activities to a great extent.

It may be noted that these three types of tests are not specific instruments of evaluation. Rather, they simply represent three procedures by which technique of testing may be adopted in different evaluation instruments.

Further, these procedures are specified in terms of the nature of the tasks involved in the test. Thus, any test in which all tasks are to carried out orally may be called an oral test.

In other words, an achievement test in a particular subject which adopts oral testing procedure may be considered as an oral achievement test. Or, by changing the procedure (i.e., nature of the task involved) to writing one may have a written test of achievement. Similarly, a test of intelligence in which all the items require the manipulation of certain objects (i.e., all are performance items) may be called a performance test of intelligence.

However, this does not mean that a test should always consist of tasks of only one type. Instead, depending upon the suitability of particular procedure for the specific aspect, a test may involve more than one procedure of testing. For instance, certain tests of intelligence consist of items involving tasks of all the three types, namely, oral, written, and performance.

Also, one may come across two tests for evaluating the same aspect but adopting two different procedure of testing. For instance, for questions in achievement tests answers may be obtained from the students both in the oral and in the written forms. And,

the teacher has to select the appropriate procedure by considering the purpose and certain practical aspects.

As it can be seen, written examination requires certain level of mastery in reading and writing on the part of the students. This, naturally, one cannot expect for children at the early stages of education. Therefore, for evaluating students in lower classes oral testing procedures cannot be adopted.

Another use of oral testing is that during the course of classroom work the teacher can conveniently evaluate the students through oral testing.

While examining orally, there is possibility for changing or rewarding the questions. Therefore, through further questioning, teacher can follow the thinking process in a student when he fails to answer certain questions. This helps the teacher in diagnosing students problems in learning. Thus, oral tests have special value as tools for the diagnosis of learning difficulties of the students.

For evaluating through oral testing, each individual will have to be tested seperately. Therefore, it is much time consuming and less practicable. Thus, for evaluating large groups of students written tests are more suitable than oral tests.

Moreover, by the very nature of the procedure, when oral tests are used as tests of achievement, scoring of students' answers is easily influenced by the examiners' opinions. Thus, oral tests are, in general, less objective.

Apart from being less objective, in the procedure of conducting oral examinations, several important aspects such as the questions to be asked, time for which each student should be tested, etc., are not

specified before hand. The examiner is free to change these things for different students. Because of this non-specific nature of the procedure, the results of oral tests become less reliable. Thus, in general, oral examinations are considered to have low reliability.

It may be, therefore, be concluded that unless the procedure of conducting oral examinations, and the procedure of scoring students' answer in them are made highly specific, oral examination results are quite unreliable.

Thus, in general oral tests are considered to be less reliable and less objective. But, that does not imply that written tests of achievement are always more objective and reliable. In fact, written tests with only essay questions may also be less reliable and less objective to an equal extent as the oral tests.

The above discussions regarding tests was only in terms of different testing procedures and not in terms of tools of evaluation.

But, it has been noted earlier that different tests (which are the tools of evaluation under testing technique) have to be used for different purposes of evaluation. Therefore, tests may be classified into specific types, considering specific purposes for which they are used in schools.

Achievement Tests

As has already been discussed, tests are most commonly used in schools for the purpose of determining students' achievement in specific content areas of different school subjects. Tests used for this purpose are, therefore, called achievement tests.

Teacher-made and Standardised Tests

Generally, achievement in school subjects is considered in relation to the instruction provided at the school. Therefore, a test of achievement may also be considered in relation to the instruction provided. Accordingly, an achievement test used in a particular school is generally made by the concerned teacher himself. These tests prepared by the teacher himself are called teacher-made achievement tests. Results of these teacher-made tests, therefore, indicate students' achievement in relation to the instruction provided in the particular schools.

Evaluation comes not merely at the end of instruction. It is a continuous process appearing at all stages of providing instruction. Thus, teachers test students' achievement, periodically, at various stages of providing instruction.

Results of these tests provide information regarding the extent to which all the specific objectives of a particular content area have been achieved. Results of these tests are of great help in improving the instruction. As it has already been noted, they are useful to teacher and students in various ways. They help the teachers in understanding their students, in taking proper decisions about teaching, and testing the effectiveness of their own teaching. They motivate the students towards further learning and lead them towards self-evaluation.

In addition to this type of tests, teachers, also prepare and use comprehensive tests of achievement. This type of test is used to measure students' achievement as related to a complete course. Marks obtained by students on these tests are interpreted by the teachers to give ranks and grades to the students as A, B, C, etc.,

i.e. consider them as good, average, etc. which will be the basis for deciding about promotions.

Consider an illustration and see what these grades and ranks really imply.

Suppose a test in history has been administered by a teacher. After scoring the answers, he finds that majority of the students have got between 50 and 60 marks out of a total of 100. Only a few, say 7 or 8 students, have scored above 60, with 68 as the highest score. Naturally, he would consider those students who have scored above 60 marks as good or above average, those with marks between 50 and 60 as average and so on. The student with 68 marks would be ranked as the best student.

It may be observed that the teacher considered the level of achievement of the student with 68 marks as the best as it is the highest score obtained in the group. Similarly, students with more than 60 marks were considered as good by considering the general performance of students of the particular group or class. Grades or ranks as average, good, best, etc., are considered in relation to the performance of the particular groups of students for whom it has been prepared. In other words, interpretation of scores on a teacher made test as good, very good, average, etc., is done with reference to the performance of the particular group of students for whom it has been prepared.

Now, another aspect of achievement testing may be considered. Although, instruction provided differs from school to school, one, generally expects students of a particular age group or class to possess some general level of achievement in different school subjects, such as languages, arithmetic, science, history, etc. Further, one would expect to find students of different achievement

levels, who can be considered as good, average, etc., in relation to the general achievement level of the whole group of students belonging to the particular age group or class.

For instance, one may like to know the average achievement level in Mathematics of all students of age 8 to 9 years. Also, he may like to know whether a particular student's achievement in Mathematics is average or above average, etc., in relation to the average achievement level of the whole group. For this, one has to, first, obtain scores for all the students of the particular age group which would indicate their achievement in Mathematics. He may, then, compare these scores to determine the average achievement level for the whole group, and also to grade students as good, average, etc.

But, an important question is whether one can use results of teacher-made tests for the above purpose?

Suppose scores and grades have been collected in Mathematics of a large number of students of the age 8-9 years on different teacher-made tests administered in the respective schools. It was noted that for the present purpose one has to compare these scores and grades. But, these cannot be compared as they are results obtained through different teacher-made tests.

This indicates that teacher-made tests cannot serve the purpose of studying achievement of students in large groups as the one that is being considered.

In the following, it is clarified as to why results of different teacher made tests cannot be compared and what should be the nature of the test if it should serve the present purpose of measuring the general achievement level of students of particular age group or class, etc.

First of all, it may be noted that results of teacher made tests have to be considered in relation to the instruction provided at particular schools. Also, instruction provided may differ from school to school.

It is also known that instruction provided in any subject is based on the set of objectives specified for the subject. Naturally, different teachers may consider different sets of objectives for the same course.

Thus, a test prepared by particular teacher would represent a particular set of objectives. In other words, results of different teacher made tests represent achievement related to different sets of objectives.

Thus, teacher-made tests cannot serve the present purpose, since results of these tests cannot be compared as they do not correspond to a standard set of objectives. This indicates that for the present purpose of studying the achievement level of students in relation to large groups, one needs a test measuring achievement related to a standard set of objectives which are, generally, covered by the instruction provided in all schools.

Another aspect that has to be considered in this connection is differences in the accuracy of results obtained through different teacher-made tests.

And, it is known that accuracy of test results depend upon the extent to which the test possesses the three essential characteristics, namely, validity, reliability, and objectivity.

It is also known that these qualities in a test depend upon the testing procedures adopted in it such as the procedure of administration, selection of specific items to be included in it,

procedure of scoring, etc. And, different teacher-made tests may adopt different testing procedures.

Thus, it may be said that results obtained through different teacher made tests do not serve the present purpose as these tests do not adopt standard testing procedures.

This indicates that for measuring achievement of students in large groups as the one that is being considered, a test which adopts standard testing procedures is needed.

Suppose one has got teacher-made tests which are based on standard set of objectives and adopt standard testing procedures. And, he has collected the grades of large number of students of age 8 - 9 years obtained through these tests in the respective schools. But, one cannot consider these grades to indicate the achievement level of the students in relation to the whole group of students of age 8 - 9 years. This is because, these grades have been given considering the average achievement of the students of particular groups in different schools.

That is, different procedures may have been adopted in the interpretation of scores for different groups of students. For instance, in one group a score of 60 may be considered as good, while in another group it may be in the average level. Thus, grades on different teacher-made tests cannot be compared as they may not adopt a standard procedure for interpretation of scores.

From this it is clear that one has to adopt a standard procedure of interpretation based on the information obtained about the general performance of all or a large number of students of the particular age group, viz., 8 - 9 years. That is, it is necessary to have a test which adopts a standard procedure for interpretation of scores.

Thus, in order to study the general achievement level in various school subjects for large groups of students, such as all those belonging to a particular age group or class, a test with the following qualities is needed.

1. It should be prepared with reference to a standard set of objectives.

2. It should adopt standard procedures of testing.

3. It should adopt standard procedures for interpretation of scores.

Tests of this type are generally known as standardised tests of achievement.

Results of standardised tests of achievement provide a common basis for comparing the achievement of students of large groups without considering the instruction provided at the particular school. Also, from a student's performance on a standardised test of achievement one can determine whether his achievement level is satisfactory or average, or above average, etc. with respect to the average achievement level of the whole group being considered.

Further, one can also identify children at the extremes, that is, those who are very bright and those who are very dull. Thus, it helps the teacher and school authorities adjust the instructional programme to the needs of all the students.

Diagnostic Tests

Achievement tests, about which it has been discussed, are tests used for the purpose of determining achievement of students in relation to specific content areas.

Now, tests with purposes other than measurement of achievement, may be considered. Consider the following illustration of testing and see the purpose for which testing has been used.

Suppose a mathematics teacher has taught the unit on 'addition, substraction, multiplication and division of decimal numbers'. In order to evaluate the pupils he gives a test consisting of problems on decimals. When he checks the answers of the pupils he finds that many of the pupils have gone wrong on particular items in the test. He further discovers that all these items involve 'multiplication of decimals'. This indicates that pupils have difficulty in learning the 'multiplication of decimals'.

The teacher then gives a more detailed test consisting of only multiplication problems involving different types of decimals. Thus, the teacher is diagnosing the learning difficulties of the students. Here, testing has been used for the purpose of diagnosing the learning difficulties of the students.

Tests used for this purpose, namely, for diagnosing the learning difficulties of the students are called diagnostic tests.

It may be observed that achievement tests provide information about student's learning in different subjects. And, diagnostic tests tell the teacher about the difficulties of the student in learning different subjects.

It may be observed that there is an important difference between general achievement tests and diagnostic tests, although both are related to classroom learning.

From the illustration discussed above it is clear that through the general achievement test the teacher could identify only the

general area, viz., multiplication of decimals, where the student's had difficulty. But, to identify the specific point where the difficulty was, he needed a diagnostic test which was more detailed and dealt with only that particular topic. Thus, diagnostic tests of achievement are more detailed but they cover only particular topics of instruction. On the other hand, general achievement tests are not detailed and they cover all the topics of instruction.

Merely diagnosing the illness of a person does not help. He should also be provided remedial treatment to bring him to normal health. Similarly, after diagnosing the difficulties of a student, the teacher has to provide remedial instruction to bring him to the general level of achievement. Thus, it may be concluded that using a diagnostic test to identify the learning difficulties of the student should always be followed by remedial instruction.

Intelligence Tests

Purpose of general achievement testing as well as of diagnostic testing is directly related to classroom learning in different subjects. In the following, two types of tests will be discussed whose purpose is not directly related to classroom learning.

In classroom and outside, one generally finds some students who always learn better and more, adjust themselves to new situations faster, and can think and reason out abstract ideas more easily than other students. Teachers, generally, consider these students as intelligent or bright students.

Also, one may find some students who are always poor in the subjects and cannot adjust to the home or school properly. These students are, generally, considered to be dull or unintelligent. And, the remaining large majority of students may be of average level in all the abilities.

The fact that students are different in their learning abilities, which can be easily seen to influence all aspects of development, implies that they need different types of programmes at the school.

Thus, a proper understanding of the above characteristic of the student is necessary both for teachers and administrators. With this knowledge they can adjust the classroom instruction and also activities in the school so that they suit the abilities of all the students.

It was noted that during casual observation one may find some students to be highly intelligent, some to be unintelligent and the remaining to be of average intelligence. However, the teacher cannot completely rely upon information obtained through such observation. Moreover, through such observations he cannot know how intelligent a student is; that is, he cannot specify the actual amount of intelligence of the student.

To know the actual amount or quantity of intelligence of a student would mean that he has to measure the intelligence. Tests used to measure intelligence are called intelligence tests.

It may be found that some tests of intelligence involve tasks which require the manipulation of words or language. Or, one may simply say that some tests of intelligence involve the use of language or verbal material. And, some others involve the use of non-verbal material. Accordingly, intelligence tests may be classified into two types namely,

1. Verbal tests of intelligence
2. Non-verbal tests of intelligence

Also, one may come across certain tests of intelligence, which, mainly, require manipulation of objects and involve physical or motor activities on the part of student being tested. These tests which involve only performance items may be called performance tests of intelligence.

Thus, on the basis of the nature of the tasks involved one may classify intelligence tests into the following three types :

1. Verbal tests
2. Non-verbal tests
3. Performance tests.

It may also be noted that certain tests of intelligence are to be administered to each individual separately. And, certain others tests can be administered in a group.

Thus, on the basis of administration procedure, we can classify tests of intelligence into individual tests of intelligence and group test of intelligence.

Aptitude Tests

Through intelligence tests the teacher may measure students' general ability to learn. But, he also comes across in schools certain students who may not be bright in all aspects, but are good in certain specific or special abilities such as musical ability, mechanical ability, etc. or in certain, specific subjects of study as languages, science, mathematics, etc.

In the same way, as in the case of general ability he finds that different students have these special abilities to different extent. These special abilities of learning are called aptitudes. And,

tests used for the purpose of measuring the different special abilities are called aptitude tests.

There are various special abilities; they may be musical ability, numerical ability, mechanical ability, etc. or may be ability in learning specific subjects such as language, mathematics, etc. Accordingly, there are various aptitude tests such as musical aptitude test, mechanical aptitude test, numerical aptitude test, etc.

It should be noted that these aptitudes are not directly dependent on the classroom learning. Therefore, by knowing the aptitude of a student one can know the particular work or subject of study which the student can learn well at present and also, will succeed in the future when he takes up a job or higher studies.

This indicates that through results of aptitude tests one cannot only know the present abilities of the student but he can also predict his success in future learning or work.

Many times students take up certain special subjects of study in the lower classes, but find later on that they cannot learn them well. Similarly, some get trained for particular jobs based on the subjects they have studied and find that they cannot do well when they actually take up the job or vocation. In order to avoid these situations, the teacher may guide the students in the schools by knowing their aptitudes which tell them the particular subjects or vocations in which they would succeed in the future. Thus, aptitude tests have special value for providing educational and vocational guidance to the students.

Thus, under testing there are three types of tests considering the nature of the task involved i.e. on the basis of whether the task

is carried out orally, or whether it requires written response, or whether it involves manipulation of objects and motor activities. These types are :

1. oral tests
2. written tests
3. performance tests.

Also, there are four types of tests considering the purposes for which they are used, such as for measuring achievement, for diagnosing learning difficulties, for measuring general ability of learning or intelligence, and for measuring special abilities. These tests, respectively are :

a. achievement tests
b. diagnostic tests
c. intelligence tests
d. aptitude tests.

Observational Techniques

In testing, information for evaluation is based on the product of a task performed by the student. Thus, it is the product of the performance that is considered and not the performance itself.

However, in many situations, apart from knowing about the product, the teacher may also have to know how the student arrived at that product. That is, it may also be important to obtain information about the performance itself which has resulted in the particular product.

Also, performance in certain aspects such as social behaviour, singing, dramatising, etc. may not result in any product at all. In such cases also one has to evaluate the student based on the performance of the task, only.

Information about the performance of a task can be collected by actually observing the student when he performs. And, this procedure of obtaining information for evaluation through observation is called observational technique.

While using the testing technique, one can, generally, assign specific numbers to the answers of the students (i.e. products). In other words, generally, one can specify the actual amount of learning or changes by measuring the learning outcomes determined through testing.

Consider the following situation of using observational technique and see whether one can measure the learning that has taken place.

Suppose the teacher has taught the students to carry out certain experiment in the science laboratory. Now, in order to evaluate, he observes the students when they are doing the experiment.

Here, the teacher is using the technique of observation for evaluating the students. Through his observation he finds that all could complete the experiment. But, some could do it very well, some could not do well, some did it very badly, and so on. Thus, the quality of performance differed from student to student.

Now, for measuring the learning outcome teacher has to assign specific numbers which should indicate the actual amount of learning that has taken place. But, he cannot assign specific numbers to the learning which he has noted through observation of the performance. Because, he finds that all the students could do the same experiment completely, eventhough there was difference in the quality of performance.

Thus, in this case, learning has to be noted in terms of quality and not in terms of quantity.

In other words, in this case, one cannot obtain a quantitative description. Only a qualitative description of learning can be obtained.

It may be seen as to how the teacher may record this qualitative description of the information obtained through observation.

Teacher observed that students differed in the quality of performance. And, he could consider different students as good, average, poor, etc., considering the quality of performance. Thus, he could classify them into different groups as good, average, poor, etc., based on the quality of performance.

It may be observed that such classification actually involves rating the quality of performance as good, poor, etc. Thus, through observation one can obtain ratings for each student which indicate the quality of performance.

Quality can be considered as something which increases from very poor to very good or in some similar way. Thus, the teacher can present the ratings on a scale of the following type which represent increase in quality.

Very poor : poor : average : good : very good.

Thus, the teacher can present qualitative description of information through observation through a scale of ratings.

Rating Scale

While actually observing the performance of a task, teacher not only looks to the task as a whole, but also in terms of various aspects of carrying out the task. Therefore, before observation,

he may list the different aspects involved in performing the particular task. And, during the observation he may record his rating for each aspect on a seperate scale. Such an instrument, specifying the different aspects being observed and providing for rating each aspect on a seperate scale is called a rating scale.

The scale that is presented contains five classifications or scale points. Any how, teacher may prepare scales with more or less number of scale position depending upon the aspect being measured.

In certain rating scales the ratings are specified by numbers as 1, 2, 3, 4, 5, in order, indicating the increase in quality. And, the teacher may thus assign these numbers to each aspect of performance. The total of these numbers is sometimes taken as specifying the quantity of learning i.e. as a measure of learning. Thus, in certain rating scales specific numbers are assigned to the ratings, in order to measure the characteristic or aspect of learning that is being evaluated.

Rating scales are mainly used in the evaluation of different skills and also several aspects of personal-social behaviour.

Check Lists

Another, approach to record the observation is to, again, prepare a list of specific actions or behaviours to be observed. And, during observation, you may check those behaviours or actions when they actually take place. Such a tool is called a check list.

Check lists may be used for observing the performance of a specific task, or, it may also be used to record the observations of a student's behaviour during several instances.

Anecdotal Records

Certain aspects of behaviour cannot be evaluated by observing the student at a particular time. They have to be evaluated through certain incidents or anecdotes as they take place in the usual course of the day.

In such cases, one may record the incident or anecdote when it happens. From these recordings he may finally, evaluate the particular aspect. This type of tools where information is provided in terms of records of anecdotes are called anecdotal records.

Suppose, you find that a pupil in the school always remains seperate from others and tries to avoid taking part in any group activities. In other words, you find the pupil socially not well adjusted. In order to make him better adjusted you have organised certain programmes through which the student may try to change his behaviour. Suppose you want to evaluate the changes in behaviour of this student through observation.

Here, you cannot say anything by observing him just once. Because, changes in such aspects do not take place all of a sudden.

Instead, at the school the pupil gets several situations where he may either mix with others or stay alone. Therefore, in order to determine the changes in the pupil's behaviour the teacher may record the several incidents or anecdotes which describe whether the student remained alone or he made any attempt to mix with others. This record may be considered as an anecdotal record.

The anecdotal record gives just description of certain incidents. The teacher has to study the incidents and interpret whether they indicate desirable changes in the behaviour of the students.

Thus, in using anecdotal record as a tool of evaluation, the teacher has to first record several incidents that he observes over a fairly long period. And, then he has to interpret these incidents to determine the extent to which changes have taken place in the behaviour of the pupil.

The above illustration of the use of anecdotal records indicates that they are of special value for studying and evaluating certain unusual aspects of emotional and social behaviour.

So far, three tools of evaluation used under observation technique, have been noted, namely, (i) rating scales, (ii) check lists and (iii) anecdotal records. In all the three tools, the observer has to carefully see and decide whether a particular aspect behaviour is there in particular incidents or performances. Thus, in using these tools of evaluation, accuracy of results, mainly, depend upon the ability of the observer.

Further, many times, during observation the observer may bring in his own opinion, instead of objectively recording the observations. Thus, generally, these tools under observational technique are less objective.

Sociometric Techniques

In the classroom and outside certain students always stay together, certain students are more liked by all the students, certain of them are, generally, not liked by any one else, and so on. These social relationships between different students, naturally, influence all aspects of their development. Thus, in order to properly understand the students, teacher has to systematically evaluate these social relationships among the students.

A method for studying these relationships is to create specific situations wherein the students' indicate their preferences or choices or likings through their actions. One may determine the social relationships among the students through observation of the students' actions which indicate their choices, preferences, likings, etc.

For instance, in the play ground the teacher may ask the students 'whom do you want to play with?' Or, for choosing leaders for different games he may ask 'whom do you want to be your leaders?' And, then he may actually allow them to act according to their choices. Several situations of this type can be presented to the students in the classroom as well as outside. It may be noted that each such instance represents a social situation. And, the choices or actions of the students indicate what type of relationships they have with others. Therefore, by studying the choices i.e. observing their action in these situations one can determine the social relationships among the students. This method of evaluating the social relationships is called 'sociometric technique'.

Thus, sociometric techniques are methods of studying and evaluating social relationships among students. These methods involve presentation of specific social situations in which the students' actions indicate their choices, preferences for particular students, etc. Social relationships among them are determined by observing choices, preferences, etc. made by them.

These techniques are specially useful in identifying those who are isolated i.e. not preferred by any other student, and helping them to develop proper relationships. It is also useful for identifying those who are liked by many others and who can be good leaders.

Thus, four types of observational techniques have been discussed, which are useful for evaluation of specific skills like reading, drawing, etc., and also social and personal behaviour.

These four types are :
1. Rating scales
2. Check lists
3. Anecdotal records
4. Sociometric techniques.

Self-Reporting Techniques

Both under testing and observation information about the student is obtained indirectly from the product of a performance, or the performance, itself. However, we can also obtain certain information directly from the student himself. This procedure of obtaining information for evaluation is called self-reporting technique.

Interview

In order to make the student report about himself, a way would be to ask him several questions related to a specific aspect. And, the teacher may obtain information for evaluation from his answers to these questions.

Thus, the technique mainly involves questioning and answering. This can be carried out face to face, i.e. one may ask the questions personally. When this procedure is adopted, i.e. when you meet a person and get his answers to the questions, it may be said that you are having an interview with the person.

Thus, 'interview' is a specific procedure under self-reporting technique. In this procedure the teacher has to meet the student

personally and obtain answers directly from the student for questions which are related to the particular aspect being evaluated.

For instance, a teacher may ask an individual student about his personal or social problems. Then, it may be said that the teacher is using interview as the procedure of evaluation.

In conducting interviews one may use a set of questions specified before hand i.e. he may have structured the interview process. Or, in some cases one may not specify the questions to be asked before hand i.e. it may be unstructured. Thus, interviews may be of two types, namely, 1) structured interviews
2) unstructured interviews.

Questionnaires and Inventories

Instead of asking the questions personally the teacher may present the questions to be asked in a printed form and obtain answers for them. These tools consisting of questions in printed form, through which he obtains information for evaluation are called questionnaires.

These tools, namely, questionnaires are also referred to as inventories.

Certain questionnaires may consist of all open-end questions, where the student gives his answers in his own words as in an 'interview'. Such questionnaires are referred to as open-end questionnaires.

On the other hand, one may also have closed questions where the possible answers are already given, generally as 'Yes' or 'No',

'undecided', etc., and the student only has to indicate his choice. Such questionnaires are called closed questionnaires.

Also, in certain questionnaires, instead of questions, only statements will be given. And, the student may be required to express his 'agreement' or 'disagreement' to the ideas presented in each statement. In such a questionnaire, the student is generally made to specify or rate his extent of agreement or disagreement on a rating scale with numbers representing complete disagreement to complete agreement. As was noted under rating scales, earlier, a total score may be obtained representing the amount of the characteristic being measured. Thus, in certain questionnaires, instead of questions, there may be statements. Also, the characteristic being evaluated through the questionnaire may be measured by obtaining the answers on a rating scale.

Questionnaires and interviews are mainly, used in the evaluation of certain personal aspects as interest, aptitude, opinion, etc. and also, in understanding personal problems of the students.

Thus, under this section various tools used under different techniques of evaluation have been discussed. But, as it was noted earlier, certain tools classified under a particular technique may be used under other techniques also.

For instance, in certain questionnaires, which come under self-reporting, the students may be required to express their agreement or disagreement on a rating scale. Thus, rating scales are used under self-reporting technique, also.

Suppose the teacher has asked the students to draw a figure and he evaluates on the basis of their drawings (i.e. product). That is,

he uses the testing technique. But naturally, he cannot exactly measure the learning in such a case. He, therefore, rates the drawings as good, poor, etc. Thus, here, rating scale has been used under testing. Also, in evaluating the same specific aspect more than one technique of evaluation may be used. For instance, consider the illustration about preparing cardboard boxes. Teacher may evaluate the final product, which is a cardboard box that has been prepared. Thus, he may use testing. But, he may also evaluate the student's skill in preparing it by observing him when he prepares it. Thus, here, teacher has employed both testing and observational techniques.

Thus, it may be concluded that the teacher has a variety of tools available for use. But, he has to select and use that tool which is appropriate for the particular aspect being evaluated and also suitable for the students for whom it is being administered.

V. Teacher-made Achievement Tests

In the last section, the discussions were about various tools of evaluation that can be used by the teacher in schools. And, it was noted that achievement tests are the most commonly used tools of evaluation at the school.

Even otherwise, one may easily observe that major part of the school day is always devoted for classroom learning. And, it is natural that more time and effort is spent in evaluating classroom learning. Tests used for the purpose of evaluating learning are called achievement tests. These tests may be further specified as teacher-made achievement tests as they refer to achievement or learning due to classroom instruction.

In the present section, more details will be learnt about teacher-made achievement tests. Attempt shall be made to know about different types of achievement tests that a teacher may prepare and use and also to know about the advantages and limitations of these different types of tests. Further, certain points will be noted which may act as guidelines in preparing these tests. Thus, the following specific questions will be answered.

- What are the different types of teacher made tests of achievement?
- What are the advantages and limitations of these different types of tests?
- What guidelines may be followed in preparing these tests?

It is known that each type of items have their own value in measuring particular learning outcomes. Therefore, question paper

for a comprehensive test may be prepared including all types of items in such a way that it measures achievement in a comprehensive as well as accurate way. In connection with this, the following question will be answered.

--- How to set a good question paper?

Achievement Tests

In the previous section, certain facts have already been learnt about achievement tests, in general, and teacher-made achievement tests, in particular.

The purpose of an achievement test is to determine the extent to which educational objectives have been achieved. Further, teachers prepare tests at the schools, always, in relation to the classroom instruction. Thus, teacher-made tests evaluate learning in relation to the instruction provided in the classroom.

It is known that a major part of time and energy in the school is devoted to classroom learning. This fact, further emphasises the need for having teacher-made tests of achievement.

It has been noted that different types of tests may possess to different extent the characteristics of validity, reliability and objectivity. Also, certain tests may be less objective by their very nature. Thus, teacher-made tests may be classified into tests which are objective and those which are less objective. Accordingly, those tests which are objective are called objective type tests; also, those which are not objective are called essay type tests.

Objectivity of a test is essentially objectivity in the procedure of scoring answers written by the students. Further, objectivity in scoring is dependent upon the nature of the questions asked in the test.

If the questions are such that they do not require specific answers, then the scorers may also differ in their opinion regarding the value of the answers given. Thus, objective tests contain questions which require specific answers.

Under objective type tests one may have such questions which would require the student to give his own answers. But these answers should consist of only one or two words. This is essential. Because, even if the answer is in one sentence, the scorer while reading the answers may be influenced by his own opinion.

There may also be a different procedure of obtaining answers. Instead of asking the students tosupply the answers the teacher himself may give a few possible answers and require the student only to select the right one.

Thus, in objective type tests student has to answer only in one or two words, which the student himself has to supply. Or, there may also be questions which require the student only to select the right answer.

An important point that should be noted is that in an objective type test when the students are required to write their own answers and not select from a given set of alternatives, it is essential that the students are required to write only some specific words. That is, in objective type tests the students' answers can be in terms of words and not in terms of complete sentences.

When the answers are to be in sentences, naturally, different forms of sentence can be written expressing the same idea. But, here, scorers may differ in their opinions as to the particular form of the sentences which expresses the idea in the right way. And, this may, in turn, influence their scoring. This implies that when students' answers are obtained in the form of sentences, one cannot expect objectivity in scoring.

In other words, when the questions in a test require the students to write their answers in the form of sentences, the test becomes less objective.

But, expressing the ideas in the best way is itself an ability that one may expect the student to possess. Thus, though they may not be objective one should also test the students' ability to express their ideas. And, therefore, there should be questions which require answers in the form of sentences. Tests with this type of questions which require answers consisting of one more sentences are called essay type tests.

Essay Type Tests

An essay type question does not clearly guide the student as to the particular material that should be written as the answer. For instance, when a question requires the students to discuss the effects of World War II, the students may be knowing many ideas related to World War II. But, here, of all that he knows, he has to select those ideas which, he thinks, should be included in the answer. Thus, an essay type test requires the student, first of all, to select the particular matter that forms the answer for the particular question.

Selecting the particular ideas that may be included in the answer does not complete the work. Because, before writing those ideas on the paper, they have to be properly organised. He has to decide which ideas should appear in the beginning, which later on, etc. in the essay to be written. Then, he may actually write his answer in an essay form. Thus, in essay type tests, after the student selects the matter which forms the answer, he has to organise the matter properly, and then write his answer in an essay form. Thus, the student has to select the relevant matter, organise the matter and then write the essay, we may see that each of the above acts, namely, selection and organisation of the

relevant ideas, and expressing these ideas in written form, represents an important learning outcome that may be expected from the students. Also, it can be seen that students differ in these abilities. Therefore, through an essay type test one can know the extent to which students possess these abilities (learning outcomes) of selection, organisation and expression of ideas in writing. Thus, essay type tests are useful for measuring the outcomes specifying the ability to select, to organise, and to express in writing.

An essay type test contains only few questions and, generally, preparation of essay type test is considered to be easy, when one considers the difficulty involved in preparing objective type tests.

An important characteristic one has to essentially look for in any type of achievement test is whether it can cover all the content areas of instruction. But, in an essay type test, since it would consist of only few questions, coverage of content will always be low.

Further, accuracy of results obtained by any test depends, mainly, upon whether it is possible to control or guide the students to write only specific answers. Otherwise, student's answers may not clearly indicate actual learning outcomes. From this point of view, there is practically no control over the nature and actual content of students' answers in essay type tests. In fact, students are free to write the answers in their own way.

Also, by the very nature of the answers required it is essential for the student to have the ability to write his answers properly. Sometimes, it is possible that a student may know the correct answer but is not able to present his answer in writing properly. And,

this may, naturally, affect the scoring of his answers, and he may get less marks although he knows the answers, correctly. Thus, in essay type tests writing ability of the student influences the scoring of his answers.

In tests where the students are required to answer by just selecting the correct one from a given set of alternatives, they can select the answer by guessing. And, by chance their guessing may be correct and they may get high marks which may not represent the actual amount of learning in them.

But, it can be easily seen that, in essay type tests the possibility of guessing is minimised.

Another factor involved in the way student answers may also influence his total score in the same way as guessing does. It is that, sometimes, when a student does not know the actual answer, he may still write some matter which seems to be relevant to the question but does not represent the answer that is actually required. This is, generally, called, 'bluffing'. Bluffing may misguide the scorer and make him award marks to the answer given, although it is not correct. This type of error due to bluffing can be avoided if the question is specific and the nature and content of the students' answer is controlled. But, obviously, no control is brought over students' answers in essay type tests. In fact, students are completely free to select and write their answers in the way they want. Therefore, it may be concluded that in essay type tests there is possibility of bluffing by the students which would influence the scoring of answers. And, as a result of this, the students' marks on the test will not represent the actual amount of learning in them.

Considering the different aspects of essay type tests, discussed above, one may try to know about their validity, reliability and objectivity which are the essential characteristics of any good instrument of evaluation.

It was noted that essay type questions do not clearly specify the exact answer required. Also, same ideas can be organised and presented in different forms. Because of this lack of specificity in the content and nature of the answers required, scoring will be influenced to a great extent by the opinions of the scorers. Thus, essay type tests, by their very nature, are not objective.

In addition to this lack of objectivity, there is possibility of bluffing and also writing ability of the student may interfere in giving the correct answer. Because of these errors in measurement, the total marks obtained by the student may not represent the actual amount of learning in him. This implies that essay type tests have less reliability.

Thus, generally, essay type tests are not objective and are also less reliable. This indicates that validity of essay type tests has to be low as they are less reliable and not objective.

Moreover, essay type tests consist of only a few questions and therefore content coverage will be low. This also indicates that essay type tests have low content validity.

Thus, essay type tests have many draw backs. They have no objectivity and they are less reliable and less valid.

Although essay type questions have so many limitations, it is essential to note that measurement of certain higher level objectives such as application, organisation, interpretation, etc.

can be properly done only through essay type tests. This indicates that in order to measure all types of learning outcomes, in particular, for measuring certain higher level objectives, it is essential to have certain essay type questions also in the examinations.

Thus, essay type questions are essential. But, it should also be noted that, generally, they are very poor in all the three essential characteristics of a good instrument. This suggests that essay type tests, should be used only after making proper improvement.

How to Improve Essay Type Tests?

Consider the following question : 'Describe the three essential characteristics of a good instrument of evaluation and explain why essay type tests are not objective.' It may be seen that the first part of the question refers to a particular objective, which is different from the objective referred to in the second part. Here the difficulty may arise as to which should be given more importance. To avoid this situation the question may not refer to more than one objective. In other words, each question in an essay type test should be planned to measure only one specific objective of instruction.

Certain objectives like mere recalling certain names, dates, specific tests, etc., can be measured through objective tests. Of course, they can be asked in the form of essay questions also, in which case, naturally, the scoring will be less objective. In such a case, objective type questions should be asked.

Thus, essay type questions should be used only to measure those outcomes which cannot be measured through objective type questions.

For instance, consider the following two forms of a question on techniques of evaluation.

A. What are the three major techniques of evaluation used in schools?

B. The three major techniques of evaluation used in schools are :

1. _____ technique
2. _____ technique
3. _____ technique.

In both the forms, answer expected from the student is to specify the names of the three techniques. But, form A. is an essay type question while form B. is an objective type question.

Also, through both A. and B. just a knowledge outcome is being measured. Student has to recall the names of the three techniques and reproduce them as the answer. Of the two forms, being an objective type question form B. is specific and clearly indicates the answer required, viz., names of the three techniques. But form A. is not clear and specific. For, it does not clearly indicate whether only the name of the technique should be written or some explanation should be given for each and so on. Here, form B should be used because it is completely objective. But form A. is not clear and specific about the exact answer required and therefore, it is not objective. From the above instance it is further clear that essay questions should not be used to measure those outcomes of learning which can be measured through objective type questions.

A very common but, serious defect of essay questions is that they do not indicate in definite and clear terms to the student what he is required to write. For instance, frequently questions of the type : 'Write a half page on 'validity of tests' are found.' About validity of tests one can write many things, such as, meaning of validity, types of validity, methods to find out validity of test, relationship between validity and other characteristics and so on. Therefore, this essay question does not indicate to the student the exact answer he should give.

When such questions are asked, the pupil naturally has to make guesses concerning what it is that the teacher desires. As a result, the student's answer will not be definite and restricted to the particular aspect being tested. Therefore, essay questions should be properly formulated such that they require definite and restricted answer.

When the question does not require definite and restricted answer, content of the students' answers will vary so widely that the scorer may not be able to decide which answer is more appropriate. And, therefore, answers will be scored according to the personal opinion of the scorer and not their actual merit.

Let us consider a question which does not require definite and restricted answer, and, understand how to improve it to make it more definite and clear.

Suppose, you have been asked to answer the following question.

'Compare objective and essay type tests'

For this question, one student may compare the two type of tests, mainly, with reference to the learning outcomes which can be properly measured through the two types, and with reference to the nature of students' answers in the two types.

But another student, in answer to the same question, may compare the two types with reference to preparation of questions in the two types and objectivity of scoring of anwers.

In fact, both these answers should be considered as partial answers. Yet, one cannot say anything definitely regarding the appropriateness of the two answers.

Suppose the same question is written in the following form :

'Compare objective and essay type tests with respect to (i) learning outcomes measurable (ii) preparation of questions (iii) coverage of course content (iv) control over pupil's answer (v) scoring.'

It can be easily observed that in this case, the question requires definite and restricted answers from the students.

So far, two important principles have been discussed about which should be followed in preparing essay type tests, in order to make the test results more reliable.

One of the principles referred to the appropriate purpose for which essay type questions should used. Essay type questions should be used to measure only those learning outcomes for which no objective test is available.

The other principle referred to the proper formulation of essay type questions. It was observed that essay questions should be so formulated as to require definite and restricted answers from the students.

Further, it is known that the greatest drawback of essay type tests is that scorers will be influenced by their personal opinions in

scoring essay type answers. This may be a personal bias toward the styles of writing, the content of the answer itself, or certain other factors such as legibility, spelling, grammatical usage etc. Thus, the main limitation of essay type tests is that scoring of answers in these tests cannot be objective.

Inspite of this serious limitation it is obvious that one cannot totally stop using essay type tests. The main reason is, that it is only through this form that it is possible to test students for the achievement of certain higher level objectives like application, interpretation, evaluation, logical organisation, etc. Therefore, essay type tests have to be used in the evaluation of classroom learning. But, certain methods may be followed which would increase the objectivity in scoring essay answers.

Lack of objectivity of a test refers to the inconsistency in the scoring procedure that may be adopted in scoring the same answer. But, it may be noted that this inconsistency in the procedure is not peculiar to evaluation process. In all walks of life one can find activities which are influenced by the personal opinions of those who carry them out. Also, when several people are involved in an activity, one quite often finds that there is no consistency or uniformity in the procedure followed by different people in carrying out the same activity. However, when consistency in the procedure is essential, a standard procedure that may be followed by all, should be formulated. Similarly, to bring about consistency in scoring answers one should formulate a standard scoring procedure which can be followed by all the scores.

While scoring an essay type answer inconsistency comes at the very first stage of deciding as to what matter constitutes the correct answer. Therefore, the first step in formulating a standard scoring procedure is to formulate a standard answer to each question.

This standard answer will guide the scorer as to the content he should look for in the answer.

It should be borne in mind that the primary purpose of testing is determining actual learning outcomes in the students. Therefore, both in formulating the standard answer and in checking the answer given by the student, the teacher should check whether it measures the particular learning outcomes.

Many times an answer would be scored low because of such external factors as style of writing, spelling, grammatical usage, etc. But, it should be noted that here writing is only a medium and these factors do not represent the achievement of learning outcomes. This suggests that while checking an answer for its content with the standard answer such external factors as legibility, style of writing, errors in spelling, etc. should not be considered. And, even if these factors should be considered then separate scores may be assigned to them, but these should not be taken as part of the scores which represent the actual learning outcomes.

Sometimes, different correct answers can be given to the same question with each representing the same learning outcomes. For instance, you may come across such questions as 'Give atleast two reasons for _____' or 'Discuss about two important uses of _____'. It can be easily seen that several choices are there and each will be an acceptable answer. Thus, when there are certain choices among acceptable answers, then all such acceptable answers should be included in the standard answer.

Thus, in order to avoid inconsistency in scoring arising out of differences in the expectations among the scorers as to content of the correct answer the teacher should formulate a standard answer; when several possible choices are there among acceptable

answers then all such possible choices should be included in the standard answer; further, while checking an answer for its content with the standard answer, he should consider the content only in terms of the particular learning outcomes that are being measured and he should not consider such external factors as legibility, spelling, grammatical usage, etc.

Another discrepancy in the scoring procedure arises in awarding marks (or credits) to different parts or aspects of an answer. This is a major point to be considered as quite often one finds only partial answers to the question. For instance, consider that students have been asked to describe a phenomenon with atleast four illustrations. For this, you may get a variety of answers. Some students may give only one, or two, or three illustrations. Also, the illustrations may vary in their degree of relevance to the phenomenon. Or, the illustrations may be good but they may not have been properly used in describing the particular phenomenon. This indicates that there is a need to have a standard scoring key which divides the total credit of mark for the question into partial credits clearly specifying the marks that can be allowed to different parts or aspects of the answer.

Thus, it may be concluded that in order to bring about objectivity in scoring essay type answers a standard answer should be formulated for each question which should also include a standard scoring key clearly specifying the credits that can be allowed to different parts of the answer.

In the above certain steps that may be taken to improve essay type tests were noted. However, it is quite clear that essay type tests can not be completely objective even after these improvements. In fact, as the first principle, it was stated that essay type

tests should be used to measure only those learning outcomes for which objective tests are not available. Thus, teacher should use objective tests wherever he can and use the essay tests to the minimum.

Objective Tests

It was found that one of the major reasons for the lack of objectivity of essay type tests is that they do not require definite and restricted answers. Unlike this, objective type tests require definite and restricted answers from the students.

Also, it was noted that some objective type tests require the students to write the answers on their own just as in essay type tests. But, here the answers consist of one or two words only. Since in this type students have to supply their answers tests of this type are called supply type tests.

In the other type of objective tests students are not required to supply the answers. Instead, they have to only select the correct answer from a given set of alternatives. This type of objective tests where the students have to only select the correct answer are referred to as selection type tests.

Thus, objective tests can be broadly classified into two types. These two types are - 1. supply type tests
2. selection type tests

Supply Type Tests

Under this type one may ask direct questions as in essay type tests. But, these questions would require only short answers

consisting of only one or two specific words. This type of questions may be referred to as short answer tests.

Another method adopted under the supply type is that students are presented with a statement in which one or two words are missing. And, they are to complete the statement by supplying the missing words. Tests consisting of this type of items are known as completion tests.

It was noted that short answer items consist of a direct question requiring a short answer from the students. Consider the following questions :

(a) What is the name of the prime minister of India?

(b) What is the chemical formula for hydrochloric acid?

(c) In which year did India become independent?

It may be observed that all these are direct questions and require answers of one or two words, symbols or numbers.

As it can be observed in all the examples the student's answers will be some information which he has already memorised and he only has to recall that information. Thus, it may be said that short answer questions are suitable for testing the ability to recall memorised information.

Consider certain other illustrations.

(i) In a triangle ABC, if $\angle A = 30^{\circ}$ and $\angle B = 60^{\circ}$, what will $\angle C$ be equal to?

(ii) Suppose the price of a book is 12 Rs. What will be the price of 10 copies of the book?

It may be observed that here student's answers do not represent information that has to be merely recalled. Instead, in order to arrive at the answer, the student has to solve certain simple mathematical problems. Thus, short answer tests are also useful for measuring simple problem solving abilities.

Therefore, one may say that short answer tests are suitable for measuring outcomes related to recall of memorised information and also, outcomes related to certain simple problem solving abilities in mathematics.

Consider the following items which come under the completion test.

A. The name of the prime minister of India is _____.

B. The chemical formula for hydrochloric acid is _____.

C. In a triangle ABC, if $\angle A = 30^\circ$ and $\angle B = 60^\circ$, then $\angle C =$ _____.

D. If the price of a book is 12 Rs., then the price of 10 copies of the book will be _____ Rs.

It may be observed that completion statements A and B require the student only to recall some information which he already knows. But, completing statements C and D require some problem solving on the part of the student.

It was earlier specified that short-answer and completion tests are two types under the supply type tests. The examples given above for the two types may be studied to see how the two types differ.

It can be easily seen that same questions have come under both the types but have been presented in different forms. Thus, it may be

said that short answer and completion tests are essentially the same, differing only in the form of presenting the questions i.e. either in the form of direct question or in the form of incomplete statement.

An important advantage of these tests is that the student must either recall the information required or make the necessary calculations to solve the problems. Thus, in these tests student can not get at the correct answer by guessing, which may be possible when alternative answers are presented and he knows that one of the alternatives must be correct.

Since in this type the students have to write the answers, sometimes the answers may be wrongly spelt.

Suppose, you have given the following question.

What is the name of the prime minister of India?
or The name of the Prime minister of India is _____.

You expect the students to write 'Indira Gandhi'. But, you find that some students have written 'Indra Gandhi' or 'Indra Gandi'.

That is, you find that some students have spelt the words wrongly. This may lead to inconsistency in scoring. Some may ignore the error in spelling while some teachers may think that wrong spelling indicates that the student does not know the required answer.

Thus, sometimes scoring on the short answer test will not be objective.

Consider the following.

Who is the prime minister of India? or The prime minister of India is _____. The teacher may expect to get the name of the prime minister. But, he may find a variety of answers, such as 'a woman', 'a lady', etc.

Suppose the same is written as follows.

What is the name of the prime minister of India? Obviously, teacher would not find such differences in the students' answers because the question requires a specific answer.

This suggests that the question or the blank in the incomplete statement should require a specific answer.

It was noted that direct question form and incomplete statement form are essentially the same differing only in the method of presentation. However, study the following forms of a question.

A. _____ is the capital of Gujarat State.

B. What is the capital of Gujarat State?

Suppose this question has to be presented to children, particularly at the lower classes.

Naturally, it is form B., which they will understand more easily. This can be easily seen, because first of all students are quite familiar with direct questions. In the other form they have to first understand the structure of the sentence and then the position of the blank. Only in certain situations, an incomplete sentence may make it more concise. Thus, direct question form is, generally, more desirable than incomplete statement form except where the other form makes the idea more concise.

Sometimes teachers may take a statement directly from the textbook and remove a word to make it incomplete. But, such questions may not require definite and restricted answers. Suppose the teacher wants the answer 'halogen'. For this, he may write the textbook sentence as below :

'Chlorine is a _____.'

But, he finds that many students give 'gas' as the answer.

He then changes the question as follows :

'Chlorine belongs to a group of elements that combine with metals to form salt. It is therefore a _____.'

One can easily see that in this form there are not chances of getting such answers as 'gas'.

This suggests that one should not take statements directly from textbooks to use as a basis of short answer items.

Thus, three important guidelines have to be followed in constructing supply type items, i.e. short answer or completion type. These are

(1) The question or the blank in the statement should require a specific answer from the student.

(2) In the lower classes, direct question form is more desirable than incomplete statement form except where it makes the idea more concise.

(3) Statements should not be taken directly from textbooks to be used as a basis for supply type items.

Selection Type Tests

It was already noted that in selection type tests, students do not have to write their answers. They only have to recognise the correct answer from a given set of alternatives. This given set of alternatives may contain two to any number of choices. Consider the case where there are only two choices.

True-False or Alternative Response Test :

Objective type questions with only two laternatives are obtained by giving statements which the student is required to mark true or false, right or wrong, yes or no, and the like.

It may be observed that there are only two possible answers. Since the true-false alternative is the most common one a test with items of this type is, generally, referred to as True-False type test.

Study the following examples and find out what learning outcomes are determined through these tests.

1. Wax is soluble in water.
2. Photosynthesis is the process by which leaves make the food for a plant.
3. The earth revolves around the sun.
4. 25% is less than 12.
5. Particles of negative electricity are called neutrons.

In the above set of examples, one may observe that items 1, 3, & 4 are statements of facts, which may be correct or incorrect. And, items 2 and 5 are definitions of terms 'photosynthesis' and 'neutrons', respectively, which may be correct or incorrect. The student is required to identify these as correct or incorrect.

Thus, true-false test items consist of specific statements. These may be statements of facts, definition of terms, and the like. The student is required to identify the correctness of these statements.

From the above, it is clear that True-False tests are useful for measuring the ability to identify the correctness of statements of facts, definitions of terms, statements of principles and the like.

The above indicates that True-False items are suitable for measuring certain simple learning outcomes which involve recall of learnt material.

Another important point about these items should be noted.

Consider the following item :-

T.F. Particles of negative electricity are called neutrons.

A student may identify that negative electricity particles are not called neutrons. So, he marks the statement as (F).

And, this indicates that the student knows that this definition of neutrons is incorrect. But, this does not mean that the students knows the correct definition of neutrons. Therefore, it may be said that this is a crude measure of knowledge of the students.

An advantage of these items is that scoring of these items is quite easy and it is also objective.

Although, scoring of these tests is objective, reliability of these tests is very low. Suppose a student has got 25 items correct on a test of 50 items.

Can one check whether he has really identified the correctness or he has merely guessed?

Obviously, one cannot know whether the students' score is due to guessing or it represents actual learning.

Preparation of True-False items is considered to be, generally, very easy. However, in practice, preparing items which really measure learning is a difficult job.

In the following, certain errors which should be avoided in preparing these items will be discussed.

Consider the following statement.

'Observational techniques are not objective'

It may be noted that the above statement refers to a broad generalisation about the objectivity of observational techniques. This implies that no observational technique is objective.

But, this broad generalisation cannot be accepted.

Suppose, the broad generalisation is avoided and the items is written as 'Observational techniques are generally, not objective'. This is certainly acceptable.

Thus, while constructing True-False items one should avoid broad generalisations.

In order to avoid broad generalization, sometimes, very trivial statements are given. For instance,

T F 'During the World War II, United States declared war on Japan on December 7, 1941.'

It is, naturally, not necessary for the student to know the exact date of such events.

Thus, it may also be noted that in writing True-False items one has to avoid trivial statements.

Consider the following item.

T F 'Despite the difficulties of determining the exact pH value of a solution, it is possible to determine if a solution is acid by the red colour formed on litmus paper when it is converted into solution.'

Such long sentences would, invariably come in the way of some students' answering as they may not clearly understand the question.

Thus, it may be noted that long sentences should be avoided.

Thus, it has been noted that although preparation of True-False items looks to be easy, it is not so in practice. And, certain common situations of mistake should be avoided while preparing the item.

Matching Test

In another type of selection type item, instead of one statement several statements, phrases or words will be given. And, instead of two alternatives a set of several alternatives will be placed in a column in front of the first list. Thus, all the items, will

have several alternatives. Student in this case is required to match the words or phrases in the two sets. And, this type of tests are called matching tests.

Study the following example.

Column A.	Column B.
1. Name of the answer in addition problems	A. Difference B. Multiplicand
2. Name of the answer in subtraction problems	C. Dividend D. Product
3. Name of the answer in multiplication problems	E. Quotient F. Subtrand
4. Name of the answer in division problems	G. Sum

It may be observed that in these items the student has to actually identify the relationship between the two things. Thus, matching tests are useful for measuring student's ability to identify the relationship between two things.

However, the typical matching exercise as the one given above, is suitable only for measuring actual information based on simple associations.

Consider the following example.

A	B
1. Square	(a) Figure in which all the sides are equal.
2. Diameter	
3. Isosceles triangle	(b) Figure in which two angles are equal.
4. Perimeter	
5. Rectangle	(c) Figure in which the diagonals bisect each other but not at right angles.
6. Radius	

In this example, it can be observed that under A, items 1, 3, 5 represent names of figures while other three items are names of nere lengths. Thus, one finds that there is no uniformity in the set. And, in fact, under B, all the three are properties of figures. Therefore, for a student who differentiates lengths and figures there is no multiple choice. Moreover, items under B are also not uniform. A student will only be confused with such matching exercises.

Thus, it may be noted that in matching tests there should be uniformity of property among things or items coming under any of the two column or lists.

Multiple Choice Tests

It was noted in the very beginning that supply type items require the student to write his answers. And, selection type tests require him only to recognise the correct answer. Accordingly, one may have a selection type test which consists of a direct question or an incomplete statement. This will be followed by multiple number of choices from which the student has to select the correct one. This type of tests are called multiple choice tests.

Completion tests were found to be, somewhat, subjective, which is caused by the freedom given to the student to write his own answer.

In multiple choice item student is not given any freedom to write his answer. Instead, he selects one, which he thinks to be correct, from a set of more than two choices. Since, there is complete control over the student's response, the multiple choice tests are objective.

It was noted that True-False items consist of two choices as answers to the question given. And, because of this there is great possibility for guessing.

But, in multiple choice tests since there are several choices which the student can choose, the possibility of guessing is much reduced. Thus, multiple choice tests are more reliable.

Therefore, from the point of view of objectivity and reliability, multiple choice tests are definitely better than the completion and True-False tests.

In multiple choice tests student has to choose from a set of multiple number of choices while in True-False tests only two possible choices are given. Thus, in a way True-False tests may be considered as a simplified type of multiple choice test.

However, in True-False type only one alternative response is given to the correct response in each case. That is why, sometimes, True-False tests are also referred to as alternative response tests.

In matching tests, one may consider one list of words, or phrases, or symbols, etc. to form a set of multiple choices for each item in the other list. For instance, consider the following.

(a)	(b)
1. Square	A. Diagonals bisect but are not at right angles.
2. Rhombus	B. Diagonals are at right angles but only one is bisected by the other.
3. Parallelogram	C. Diagonals bisect, are at right angles but unequal.
	D. Diagonals bisect and are at right angles.

In this, for each 1, 2, and 3 in (a) one can take A, B, C, and D under (b) to be four alternatives or multiple choices.

Thus, amatching test can be considered as a special form of a multiple choice test; it may also be considered as a combination of several multiple choice items.

A multiple choice item consists of a direct question or an incomplete statement. This is called the stem of the item. It is followed by a list of suggested answers which are called alternatives.

Consider the following item

The capital of Gujarat is

(a) Baroda (c) Ahmedabad
(b) Surat (d) Rajkot

Here, the stem is : 'The capital of Gujarat is'

The stem is followed by a set of possible alternatives. Of these, one is the answer to the question while the remaining are placed merely to distract the attention of a student who does not know the answer. These alternatives other than the answer called distracters.

Consider the following example

The capital of Gujarat is

(a) Baroda (c) Ahmedabad
(b) Surat (d) Rajkot

Here, there is only one correct answer and the student has to identify it. This type is called correct answer type. But, in certain other questions there may not be only one correct answer, the student has to choose the best answer from the alternatives given. This type is called best answer type.

Unlike the other objective type tests multiple choice tests are also useful for measuring certain higher level learning outcomes, viz., understanding, application.

Consider the following example

India is (i) ruled by the Britishers.
(ii) situated in the African continent.
(iii) a peninsula
(iv) in the western hemisphere.

Here, one may see that the stem does not represent a meaningul problem. The distracters refer to four widely seperate ideas. This, therefore, does not present a clear problem to the student.

Thus, it may be noted that in a multiple choice item the stem should always represent a meaningful problem.

Further, in the above example, since the alternatives are not homogeneous in meaning, they may not really distract the students attention at all. A student who has partial knowledge may also be able to exclude those which do not seem related to the idea and arrive at the correct answer.

Thus, it may also be noted that distracters should be homogeneous in a multiple choice item.

Apart from these, several of the suggestions given for other objective type tests also should be borne in mind while preparing multiple choice items.

Thus, under selection type tests there are three specific types of tests. They are :

1. True-False tests
2. Matching tests
3. Multiple choice tests

However, it was noted that inspite of their advantage regarding ease of scoring, objectivity and such other aspects, objective tests can measure only simple learning outcomes, mainly, at the knowledge level.

Therefore, it should be noted that in a good test which proposes to measure all levels of objectives one should have different types of tests and also different types of questions.

How to prepare a Good Question Paper?

In preparing a good question paper teacher should consider certain important principles. Some of them are discussed in the following.

It is known that the most important characteristic of an achievement test is content validity, that is, the test should represent all content areas. For this, the teacher may first of all divide the course into several units and decide about the weightage to be given to different content units.

Thus, the first step in preparing a good question paper is to analyse the course content into different content units and decide about the weightage to be given to each one of them in the test.

Also, instruction in different content units correspond to specific educational objectives. And, these specific objectives will be at all levels, knowledge, understanding, application, etc. Therefore,

after analysing the course into content units the teacher has to specify the weightage to be given for different objectives of each level under each content unit. Thus, apart from analysing into content units, the teacher should also specify the weightage to be given to objectives of different levels like, knowledge, understanding, etc., under each content.

Thus, the teacher has to make specifications of two types. First, he has to specify the different content units of the course and the weightage to each of them. Then he has to specify the objectives of different levels and the weightage to be given to them under each content units.

On the basis of the specifications made regarding the content units and the objectives of different levels the teacher may actually prepare a table of the type given in the following. Such a table representing the specifications of content units and objectives is generally known as 'table of specifications'.

Study the following table of specifications for a social studies test.

Content Area	Objectives				
	Know-ledge	Under-standing	Applica-tion	Interpre-tation	Total
Food	12	3	-	-	15
Clothing	16	4	5	-	25
Transportation	8	5	2	5	20
Communication	8	5	2	5	20
Housing	10	5	5	-	20
Total	54	22	14	10	100

The numbers like 12, 3, 16, 4, 5 etc. in the different cells are in percentages. For instance 12 in the first cell indicates that 12 percent of the items in the test, related to the content area 'Food', will be concerned with knowledge. Similarly, 3 percent of this content unit will be concerned with objectives of understanding level. Thus, such a table of specifications clearly indicates to the teacher, how much weightage should be given for objectives of different levels under each content unit.

While discussing about different types of objective tests and also essay tests it was noted that different types of test items would be suitable for evaluating different learning outcomes or objectives. From the table, one can easily make out what type of questions would be framed under the different content units so that they are appropriate for evaluating the corresponding objectives at the specified level.

The table can also help the teacher in specifying the number of questions to be asked under the different content units, so that the test will possess content validity.

This type of table actually serves the teacher the same purpose as a blue print of the design of a building to be constructed would do for an engineer. That is why, a table of specifications is also referred to as a blue print for the test.

Thus, with the help of a blue print teacher will be able to see that the question paper covers all content units and also objectives of all levels.

Another aspect that should be taken care of is regarding the difficulty level of the questions. Generally, in a final examination

question paper, about 25% of the questions should be of high difficulty level, and 25% of low difficulty level, and remaining being at the average difficulty level. Thus, in a good question paper there will be questions of all difficulty levels. Because, all easy questions, or all difficult questions would fail to measure the students with differemt degrees of achievement.

Another point can be directly inferred from the use of a blue print. Since different types of tests are suitable for evaluating different learning outcomes, it is clear that the question papers should contain different types of objective type as well as essay type questions.

Many times, in question papers one comes across broad choices between questions on different content areas. This, naturally decreases the content validity of the examination. Thus, in a good question paper there should not be broad choices between questions on different content areas.

Another important point which has been noted under the reliability of a test is regarding the clarity of language of the questions. If the language used in the questions are not clear and easily understandable to all the students their answers may not represent actual learning in them. Therefore, there should be clarity of language used in the questions.

Apart from the above points which are directly related to questions the teacher should also take care that the students are given specific directions as to what they should do under each question.

Thus, it may be concluded that the teacher has a variety of test item types which he can use for evaluating the achievement of the students. And, it is left to the teacher to adopt proper test items in proper proportion in the question paper he would use to evaluate the students' achievement in different school subjects.

VI. Elementary Statistics in Education

During the discussions in the previous sections it has been learnt about various tests, ways to prepare them and use them. It was also noted as to how various aspects of development of the students such as achievement in school subjects, intelligence, aptitude, etc. are measured with the help of various tests. Measurement involves assigning numbers. Thus, through measurement one obtains sets of scores for the students. And, these scores get meaning when they are compared with numbers which represent the achievement of specific educational objectives.

Thus, the teacher interprets the numbers obtained through testing with numbers representing the expected amount of learning specified in educational objectives. This comparison tells the teacher regarding the student's position as an individual. But, teacher has to teach the student as a member of a group of students namely, class. And, therefore, he has to understand his position in relation to the group as a whole.

For this, the teacher, generally, compares the individual student's performance with the performance of the class as a whole on different tests. But, this requires the teacher to obtain some numbers to represent the nature of the group's performance.

This section is about the methods by which one may understand the nature of performance of groups so that the individual's position can be interpreted meaningfully.

About the performance of a class as a whole one can try to know what is the average performance. Or one may also try to know what is the general nature of the students in the class in terms of how much do they vary from each other. And, also, one may understand whether one can have a way of comparing scores on different subjects.

Thus, attempt will be made to answer the questions

—— How to obtain numbers representing the average nature of a class?

—— How to obtain numbers representing the variability among the members of a class?

—— How to obtain numbers with which one can compare a student's position in different subjects?

It was noted that mere scores by themselves do not carry any meaning. A way of presenting the numbers so that they would carry meaning is to represent them in a graphical manner. Accordingly answer will also be obtained to the question

—— How to represent scores obtained on a test graphically?

Thus, the main purpose of this section is to learn different ways of analysing and understanding the scores of students of a group, taken as a whole.

Raw Scores

Consider a set of scores obtained by students on an achievement test. This set of scores may be analysed to understand the different characteristics of the set of scores such as average, variability, etc.

Suppose the teacher has got the scores of twenty students of a particular group on an achievement test.

This set of twenty scores together may be considered to represent the performance on the achievement test of the group as a whole.

This indicates that with the help of the scores of students belonging to a particular group one should be able to know about the performance of the group.

Table 1 : Scores of 33 students on an achievement test

15	13	14	13	15	17	16
22	20	24	24	2	24	21
8	22	14	8	23	12	15
23	18	24	15	6	25	
9	25	23	25	7	14	

It may be observed that the scores have been presented in the above table just as they have been obtained after scoring the test answer papers. One may say that these scores are like the raw materials which have to be processed before they can be used for different purposes in daily life. Here, also, the teacher cannot use these scores directly for the purpose of understanding the nature of the performance

of the group. Therefore, one may consider these scores which are just as they have been obtained by testing as raw scores.

As it is known, the simplest way to arrange a set of numbers (i.e. scores) is to arrange them in the increasing or decreasing order.

Accordingly the 33 scores in Table 1 can be arranged in the increasing order as in Table 2 below.

Table 2 : Scores on an achievement test arranged in the increasing order.

6	12	15	17	21	22	25
7	13	15	18	22	24	25
8	13	15	19	22	24	25
8	14	15	20	23	24	
9	14	16	20	23	24	

In the above table, scores have been arranged in increasing order, with 6 as the first score and 25 as the last score.

Generally, after the scores of all the students have been noted, teachers announce the ranks of the students as first, second, third, etc. The student who has scored highest is considered to be the first, the student with next highest score as the second and so on. Now, the teacher can directly find out the ranks of the students from Table 2.

Thus, by rearranging the raw scores in the increasing or decreasing order the teacher can directly know about the ranks of the students. In the set of scores presented in Table 2, 25 and 24 represent the scores of first and second rank students, respectively.

When one talks of prices of different material, in daily life, it is quite often said that the prices of particular material ranges from say Rs. 12 to Rs. 18 and so on. Here, the material with price Rs. 12 is considered as the material with the lowest price and that with Rs. 18 as the one with highest price. Also, in the set of scores presented in Table 2, one comes across lowest and highest scores, viz., 6 and 25 respectively. Therefore, it may be said that the scores of the corresponding group of students ranges from 6 to 25.

The difference between the highest and the lowest score is called the range of scores.

Suppose a teacher has administered a test to some seven students and the scores are 18, 23, 14, 28, 21, 20, 19, here,

The lowest score = 14
The highest score = 28
and, the range = highest score - lowest score
= 14

Thus, from raw scores the teacher can rank the students as first, second, etc., and also, calculate the range of scores, which is the difference between the lowest score and the highest score.

Grouped Data

The range of scores tells the teacher as to how much the two students, viz., the first rank student and the last rank student, vary in their achievement. Thus, the teacher identifies the student with lowest score as very poor and offers him help. Similarly, he comes to know about the best student and his abilities. But,

the teacher is required to consider all the students, and not merely the first and the last student. So that he can properly adjust his teaching to the whole class. But, from the raw scores and the range, the teacher cannot decide anything definite about the other students.

We may observe that if the teacher considers the student with a score of 6 as very poor and needing help, perhaps, he should consider those with score 7, also in the same category. Similarly, he should consider the students with scores 23 and 24 to be equally good as the one with score 25.

Further, he will have to consider some students, say those with scores 8, 9 and 10 also to be poor and needing help, but better than those with score 7 and below. Similarly, he may consider those with scores 20, 21 and 22 to be good but not so good as those with score 23 and above. In the same way he may consider students with scores from 11 to 19 as belonging to different groups as those with scores from 11 to 13, those with 14 to 16 and those with 17 to 19.

Thus, it may be said that the teacher has to understand the achievement levels of not merely those at the extremes but of all the students in the class. For this purpose, he may group the students on the basis of their scores, as those with scores 8, 9 and 10, those with 11, 12 and 13 and so on.

Thus, when the raw scores of students of a whole class are given, in order to understand the nature of performance of all the students, one may form several group of raw scores, each group representing scores of particular range like 8 to 10, 11 to 13, etc. Scores of an achievement test presented in Table 2 may be grouped and presented as in Table 3 below.

Table 3 : Grouped Scores

5 - 7	In a way, the teacher
8 - 10	may consider all students
11 - 13	in a particular group like
14 - 16	5-7, 8-10, etc. as of
17 - 19	approximately same
20 - 22	achievement level.
23 - 25	

Generally, anything given in the form of numbers, or statements etc. which may be used for further calculation and understanding is called the data. Therefore, the scores used for further calculation and understanding may be called the data. Further, the scores, when grouped as in Table 3 may be called grouped data.

In the grouped data presented in Table 3, each group such as 8-10, 11-13 etc. may be considered to represent the scores, of a small class of students, which lie in intervals of scores, viz. 8 to 10, 11 to 13 etc.

Therefore, the different groups such as 8 - 10, 11 - 13, 14 - 16, etc. in the grouped data may be called class intervals.

Each class interval represents a group of scores. For instance, class interval 8 - 10 reprsents the group of 3 scores, viz., 8, 9, and 10. Similarly, a class interval of 10 - 14 represents the group of 5 scores, 10, 11, 12, 13 and 14. In the first case, therefore, 3 may be called the size of the group. Also, in the second case the group is of the size 5. Since, these groups are the class intervals we may also consider 3, 5, etc. as sizes of the class intervals 8 - 10, 10 - 14, etc.

The lowest and the highest scores represented in a class interval are called the limits of the class interval. And, the lowest score included in the interval is taken as the lower limit and the highest score included as the upper limit of the class interval.

In a class interval 8 - 10, the lowest score included will be 8 and the highest score will be 10. Thus,

Lower limit of the class interval 8 - 10 = 8
Upper limit of the class interval 8 - 10 = 10

Thus, class intervals are formed by dividing the range of scores into small units, including a particular number of scores in each unit. The number of scores included to be in a particular class interval is called the size, or length of the class interval.

An obvious question that arises is, how many scores can be included in a class interval? In other words, this refers to the size of the class interval.

Consider the two tables 4 & 5 given below which present in the form of class intervals, the scores given (i.e. the ungrouped data) in Table 2.

Table 4	Table 5
(1) 23 - 25	(1) 24 - 25
(2) 20 - 22	(2) 22 - 23
(3) 17 - 19	(3) 20 - 21
(4) 14 - 16	(4) 18 - 19
(5) 11 - 13	(5) 16 - 17
(6) 8 - 10	(6) 14 - 15
(7) 5 - 7	(7) 12 - 13
	(8) 10 - 11
	(9) 8 - 9
	(10) 6 - 7

Size of each class interval in Table 4 = 3
Size of each class interval in Table 5 = 2

Thus, size of class intervals in Table 5 is smaller than the size of those in Table 4. It may also be observed that the number of class intervals in Table 4 is 7. But, it has increased to 10 in Table 5. Thus, as the number of class intervals has increased the size of the class interval has decreased.

In practice, also, size of the class interval has to be decided on the basis of the number of class intervals. Generally, for satisfactory grouping, number of class intervals should be between 10 and 15. Thus, the number of class intervals in Table 5 is more satisfactory than the number of class intervals in Table 4.

Generally, the convenient size to use for the class interval is determined by dividing the range of scores by 15 and taking the nearest number. In the case of scores presented in Table 2, the range is : 25 - 6 = 19.

Therefore, satisfactory length or size of the class interval may be calculated as follows :

$$\frac{\text{Range}}{15} = \frac{19}{15} = 1.266$$

Although, 1 is the nearest number, we cannot have 1 as size, the next nearest number has to be taken. Thus, the size of the class interval can be taken as 2.

But, this is not a rule that should be strictly followed. In fact, it is always convenient to take an odd number as the size, so that the midpoint of the class interval will be a whole number.

Consider the following set of raw scores.

40	56	67	71	82	86	95
43	59	67	75	82	88	102
43	62	68	75	82	90	127
46	64	68	76	82	90	
46	64	69	76	82	91	
46	66	69	78	83	91	
54	66	69	80	84	92	

Highest Score = 127

Lowest Score = 40

Range = highest score - lowest score

= 127 - 40 = 87

$$\frac{\text{Range}}{15} = \frac{87}{15} = 5.8$$

Therefore, the size of the class interval can be = 5.

Although 6 is the nearest whole number, 5 may be taken as the size of the class interval since 5 is an odd number.

With 5 as the size of the class interval, the first and the last class intervals will be, respectively, 40 - 44 and 125 - 129. A convenient way of writing down the class intervals is to separately add the size of the class interval (which is 5 in the present case) to the lower limits and the upper limits. For instance, in the present example 40 + 5 = 45 is the lower limit, and 44 + 5 = 49 is the upper limit of the second class interval. Thus, there will be '18' class intervals.

Another point to be noted about class intervals is regarding their actual limits.

It was noted that for a class interval, 40 - 44 one may consider 40 as the lower limit and 44 as the upper limit. But, a score is generally considered as the midpoint of a distance extending from half a unit below to half unit above this number. Thus, score 40 is actually the mid-point of 39.5 to 40.5. Thus, when one wants the lower limit of a class interval he should substract 0.5 and then, take similarly for the upper limit he should add 0.5. Therefore, the actual limits of the class interval 40 - 44 will be 40 - 0.5 = 39.5 and 44 + 0.5 = 44.5.

Therefore, one may write the class intervals in terms of actual limits also.

For instance, the class intervals may be written in the following two ways for the same data.

Class Intervals	Class Intervals
23 - 25	22.5 - 25.5
20 - 22	19.5 - 22.5
17 - 19	16.5 - 19.5
14 - 16	13.5 - 16.5
11 - 13	10.5 - 13.5
8 - 10	7.5 - 10.5
5 - 7	4.5 - 7.5

Of these, in the second the class intervals have been written in terms of actual lower and upper limits.

Frequency Distribution

First of all it may be found out from the set of raw scores, how many of them are in the different class intervals.

Consider the following set of raw scores

6	12	15	17	21	23	25
7	13	15	18	22	24	25
8	13	15	19	22	24	25
8	14	15	20	23	24	
9	14	16	20	23	24	

Number of scores in each class interval may be noted by examining the set of raw scores in an order and making a tally in front of the class interval every time a score is found in that class interval. Then, these tallies may be added to get the number of scores in the class interval.

C.I.	Tallies	f
23 - 25	𝍸 𝍸	10
20 - 22	𝍸	5
17 - 19	///	3
14 - 16	𝍸 //	7
11 - 13	///	3
8 - 10	///	3
5 - 7	//	2

In the second column, above, each tally is written by a mark '/' and 5 tallies are, generally, noted in the form 𝍸.

Now, numbers 10, 5, 3, 7 etc. represent the total number of students with scores between 23 - 25, 20 - 22, 17 - 19, etc. respectively.

These numbers in the third column actually tell how frequent are the scores appearing in the corresponding class intervals.

Thus, number indicating how frequent are the scores appearing in the particular class interval is called the frequency of the class interval and is, generally, denoted by 'f'.

Further a table as the one given above showing the distribution of frequencies into different class intervals is called a frequency distribution.

From a frequency distribution, as was noted earlier, one can clearly understand the nature of the distribution of scores. In the frequency distribution the total set of scores is divided into several small groups called class intervals and also, it specifies the exact number of students belonging to each group.

Consider the frequency distribution presented above.

There are 33 scores in all, distributed into different class intervals. This Number 33 is denoted by 'N'. Therefore, in the present case $N = 33$.

Add all the frequencies, i.e. all the numbers in Column 3. The sum obtained may be represented by Σf, which means sum of all 'f' i.e. frequencies. In the present case $\Sigma f = 33$.

It may be observed that $\Sigma f = N$.

Graphical Representation

Take the distribution presented above, to understand how it may be graphically represented.

Consider the following figure :

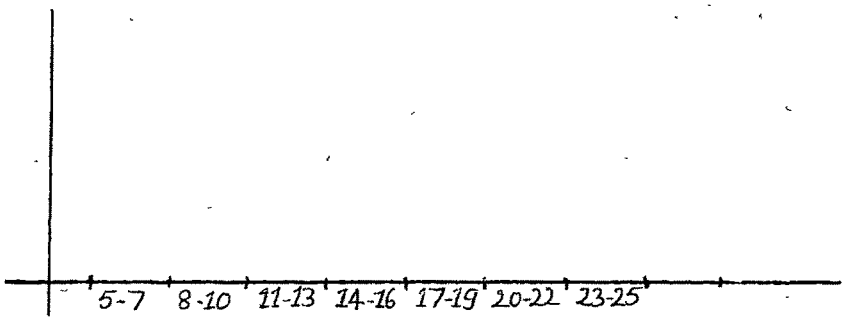
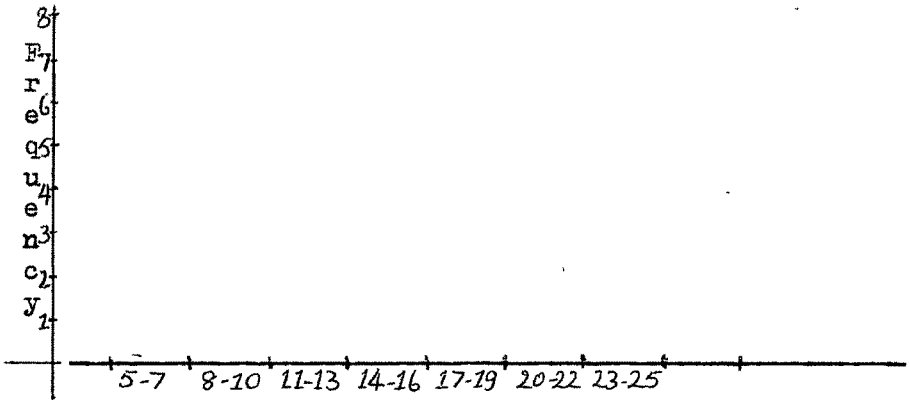


Figure 1

Here X axis has been divided into seven blocks of equal width. Each of these is taken to represent a class interval.

Now, study the following figure :



Class Interval

Figure 2

Here, markings on the Y axies represent the frequency.

Finally, the frequency distribution may be represented as follows :

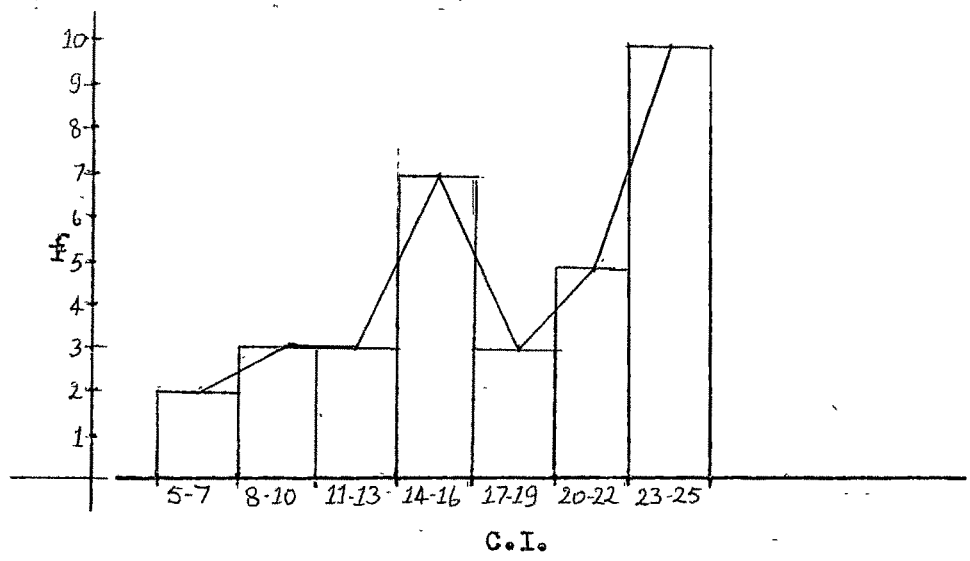


Figure 3

In the above figure, there are several rectangles. This graph is called a 'histogram'. Each rectangle has for its width the size of a class interval. Height of all the different rectangles are not the same. This is because the height of each rectangle represents the frequency of scores in the corresponding class interval.

From the above histogram the teacher can understand several aspects about the nature of performance of the group.

For instance, as was noted the height of the rectangles represent the number of students whose scores lie in the particular class interval. From this, he can compare the height of the rectangles and know the number of students in the class interval representing the maximum score or the least score, or the average score etc.

In Figure 3, there are lines joining points A, B, C, D, etc. one may easily see that these points are actually the mid-points of the different class intervals. Also, each of these points is at a height which represents the frequency of the particular class interval. Thus, lines joining A, B, C, D, etc. from a polygon (a figure with more than four sides) joining points which represent the frequency of scores in the class interval. The figure formed by these lines is, therefore, referred to as a frequency polygon.

Like the histogram, the frequency polygon also helps the teacher in making comparisons of different groups of students within the class. Also, he can know the position of each individual in relation to the position of others in the class.

Measures of Central Tendency

A way to find out the average of a set of numbers is already known. Suppose, the teacher has to administer a test to four students and has obtained the following scores.. 65, 68, 70 and 73.

The average of these four scores

$$= \frac{65 + 68 + 70 + 73}{4} = \frac{276}{4} = 69$$

Average determined by the above method is referred to as Mean and represented by the letter **M**. Thus, the arithmetic mean for the scores presented above is 69.

It may be noted that this 'mean' represents the typical value of a set of scores.

Apart from this arithmetical average, one also comes across two more types of average in statistics.

Consider the following table of scores.

Scores	
40	
36	
33	
33	
32	
30	
29	- Median
28	
28	
28	
27	
25	
20	

Both above and below the mark one finds equal number of scores. This indicates that the mark is at the mid point of the set of scores.

This mid-point of a set of scores is referred to as the median of the set of scores. In the above case, score 29 can be taken as the median. It may be observed that median divides the set in such a way that 50% of the scores lie above as well as below that point. Thus, one may consider median as another typical value representing the set of scores.

In daily life situations, one finds people doing certain things most frequently. We say that it is a fashion or the mode of life. In test scores also some score would appear most frequently. Such a score which is most frequent is called the mode of the scores.

For instance, consider the set of scores presented above. Here, 28 is the most frequent score. Therefore one may consider 28 as the mode of the scores.

Thus, three numbers have been noted which would represent the typical value of the scores of the group. These values indicate the nature of the particular group in terms of how much do students tend to be nearer the centrel point of the whole group. Accordingly, these are called measures of central tendency.

The above methods for finding out the three measures of central tendency are suitable when the scores are small in number and they are not grouped.

Consider the following set of Grouped scores.

Class Interval C.I.	Frequency f
10 - 19	15
20 - 29	12
30 - 39	10
40 - 49	3

$\sum f = N = 140$

It may be noted that, here, since actual scores are not there one has to consider the mid-point of the class interval to represent the scores in it.

Now examine the following.

C.I.	x	f	fx
10 - 19	14.5	15	217.5
20 - 29	24.5	12	294
30 - 39	34.5	10	345
40 - 49	44.5	3	133.5

In this table, one may note that 'x' i.e. 14.5, 24.5, etc. correspond to the mid-points of the class intervals. And, fx is to be considered as the product of frequency and midpoint.

Further, the mid-point of the class interval may be taken to represent all the scores in that C.I.

Therefore 'fx' should represent the total of all the scores lying in the particular interval. And, when such 'fx' for all the C.I. are added it gives Σfx which represents the sum of all the scores in the distribution.

The arithmetica average i.e. mean is given by

$$M = \frac{\text{Sum of all the scores}}{\text{Total number of scores}} = \frac{\Sigma fx}{N}$$

$$\text{Therefore, here, } M = \frac{\Sigma fx}{N} = \frac{990}{40} = 24.75$$

It may be observed that adding after multiplication makes the calculations complicated and also very long. In order to avoid this, there is a method which is shorter than the above.

Consider the same grouped data and see how one may proceed to calculate mean by the shorter method.

C.I.	x	f	x'	fx'
10 - 19	14.5	15	-2	-30
20 - 29	24.5	12	-1	-12
30 - 39	34.5	10	0	0
40 - 49	44.5	3	1	3

Total number of frequencies, N = 40.

In this method one has to assume some number to be the mean for the whole group. And, then the Mean (M) is obtained by adding a correction to the number chosen. This method is referred to as assumed mean method.

Examine the table given above.

$$\text{Here, the correction, } c = \frac{\Sigma fx'}{N}$$

$$= \frac{-39}{40}$$

$$\text{Mean} = \text{A.M.} + ci$$

$$= 34.5 + \frac{-39}{40} \times 10$$

$$= 34.5 - 9.75$$

$$= 24.75$$

It should be noted that in the problem that has been solved, the term, 'x' is the deviation of different 'x' from the mid-point of the class interval where the A.M. is assumed to lie.

Consider the following data and obtain the mean

C.I.	x	f	x'	fx'
23 - 25	24	10	3	30
20 - 22	21	5	2	10
17 - 19	18	3	1	3
14 - 16	15	7	0	0
11 - 13	12	3	-1	-3
8 - 10	9	3	-2	-6
5 - 7	6	2	-3	-6

Let the Assumed Mean = 15

$N = \sum f = 33$

$c = \frac{\sum fx'}{N} = \frac{28}{33}$

Now,

$$M = A.M. + ci$$
$$= 15 + \left(\frac{28}{33} \times 3\right)$$
$$= 17.54$$

One point may be noted regarding the particular class interval whose mid-point may be taken as the A.M.

Generally, the interval which is at the middle and which has greater frequency is chosen. However, Assumed Mean can be taken in any class interval. You may change the place of A.M. and try to find out the mean in the above problem.

Median

Suppose, in the table of frequencies considered in a seperate column you start from below and go on adding the frequencies and writing them in that column. For instance, in the last C.I. you have, f = 2. Suppose, you add this to the frequency in the next C.I. Then, you get 2 + 3 = 5 in the next C.I.

Here, one should remember that 5, yet represents frequency, but of not merely the second C.I. Instead, it is cumulative of frequencies upto that C.I. Thus, the frequencies you obtain in that column by adding to the frequency of a particular C.I. all frequencies below that C.I. are called cumulative frequencies.

Consider the following table of frequencies.

C.I.	f	Cumulative frequency
23 - 25	10	33
20 - 22	5	23
17 - 19	3	18
14 - 16	7	15
11 - 13	3	8
8 - 10	3	5
5 - 7	2	2

It may be noted that in the highest C.I. cumulative frequency will become equal to the total number of scores i.e. N.

As it is known Median refers to the particular score, which represents the mid-point of the distribution. In other words, it is the point, above and below which one finds half the number of the total number of scores in the distribution.

Since, one should find half the total number of scores both above and below, Median refers to that point represented by half of the N i.e. N/2.

Therefore, if N = 33 then,
Median refers to the point
where the score value is $= \frac{N}{2} = \frac{33}{2} = 16.5$

Consider the following distribution of frequencies.

C.I.	f	Cum.f.
23 - 25	10	33
20 - 22	5	23
17 - 19	3	18
14 - 16	7	15
11 - 13	3	8
8 - 10	3	5
5 - 7	2	2

Here, N = 33

Length of C.I., i = 3

Also, N/2 = 33/2 = 16.5

Since, all the scores in the particular range are included in the above distribution, there should also be a score corresponding to N/2 = 16.5.

While looking out for the particular C.I. one may have in mind that the particular C.I. would have a cumulative frequency value of 16.5 (N/2) or more than that but not less than N/2 in any case.

Since, the C.I. in which the median lies extends from 16.5 to 19.5, one has to find out the exact score which corresponds to 16.5.

This is determined using the formula

$$Mdn = l + \frac{N/2 - F}{f} \times i$$

where l = actual lower limit of the C.I. in which the Mdn lies.

N = Total number of scores

F = Sum of all frequencies below that interval (i.e. cum. f. below that C.I.)

f = frequency of the C.I. in which the Mdn. lies

i = length of the C.I.

$N = 33$ and $i = 3$.

Mdn. lies in the C.I. 17 - 19

Therefore, l = the actual lower limit of this C.I. = 16.5

F = cumulative frequency below this C.I. = 15

f = frequency of the particular C.I. = 3

Now, substituting in the formula the value of the Median is obtained :

$$\text{Thus, } Mdn = l + \left\{ \frac{N/2 - F}{f} \right\} \times i$$

$$= 16.5 + \left\{ \frac{16.5 - 15}{3} \right\} \times 3$$

$$= 16.5 + 1.5 = 18$$

One main difference in these two measures may be noted. In calculating the mean one has to consider all the scores. Therefore, scores at the extreme influence the value of M. But, this is not so in the case of median.

For instance, if one has the following seven scores
27, 11, 10, 11, 12, 11, 10
Then M = 91/7 = 13 and Mdn. = 10.
Suppose, instead of 27 the first score is 14 then M = 78/7 = 11.14
and Mdn. = 10

Thus, Mdn. remained the same and was not affected by the extreme score. But, Mean had gone up to 13, mainly because of the extreme score of 27.

Thus, M is affected by extreme scores while Mdn. is not affected.

Measures of Variability

As in case of central tendency, here also one finds different measures of variability.

Range :

As has been already noted one measure indicating the variability of the group is the difference between the highest and the lowest score in any distribution.

Highest score - Lowest score = Range

Refer to the set of scores presented in Table 2 at the beginning.

The Range = 19

Quartile Deviation

Obtaining this measure of variability is very similar to the procedure of calculating the Median.

Here, the score corresponding to the point above which 3/4 of the scores would lie and below which 1/4 of the scores would lie, is determined. This represents one quartile of the total number of scores. And, accordingly this point is called the first quartile point, denoted by Q_1. Similarly, the score corresponding to the point above which 1/4 of the scores lie and 3/4 of them lie below

it is determined. And, this point is called the third quartile point denoted by Q_3. Then, the difference between Q_3 and Q_1 which infact, represent variability in the middle $\frac{1}{2}$ of the scores, is calculated. Finally, this difference i.e. $Q_3 - Q_1$ is divided by '2' such that it gives the variability or deviation in the quartile (i.e. $\frac{1}{4}$ of the total length). This measure of variability is called quartile deviation, denoted by $Q = \frac{Q_3 - Q_1}{2}$

It may be noted that Median is the point above and below which $\frac{1}{2}$ of the total scores lie. Thus, it may be considered as Q_2 and therefore, the same procedure of calculation may be applied for calculating Q_1 and Q_3.

Consider the following distribution

	C.I.	F	Cum.f
Q_3	23 - 25	10	33
	20 - 22	5	23
	17 - 19	3	18
Q_1	14 - 16	7	15
	11 - 13	3	8
	8 - 10	3	5
	5 - 7	2	2

Here, Q_1 and Q_3 are calculated from the following formulae.

$$Q_1 = 1 + \left\{ \frac{N/4 - F}{f} \right\} i \qquad Q_3 = 1 + \left\{ \frac{3N/4 - F}{f} \right\} i$$

$N = 33$ and $i = 3$

$N/4 = 33/4 = 8.25$

$\therefore$ Q_1 lies in the C.I. : 14 - 16

Also, l = actual lower limit

$= 13.5$

$F = 8$

$f = 7$

Similarly $3N/4 = 24.75$

And Q_3 lies in the C.I. : 23 - 25

$\therefore$ $l = 22.5$

$F = 23$

$f = 10$

Substituting these values in the corresponding formulae

$$Q_1 = 13.5 + \left\{ \frac{8.25 - 8}{7} \right\} \times 3$$

$$= 13.5 + 0.107$$

$$= 13.067$$

$$Q_3 = 33.5 + \left\{ \frac{24.75 - 23}{10} \right\} \times 3$$

$$= 22.5 + 0.525 = 23.025$$

Now, Quartile deviation,

$$Q = \frac{Q_3 - Q_1}{2}$$

$Q_3 = 23.025$

$Q_1 = 13.607$

Therefore, $Q = \frac{23.025 - 13.607}{2}$

$= 4.079$

Standard Deviation

It may be noted that Q considers the variability in only the middle 50% of the scores. However, the next measure of variability that would be learnt considers the middle 68.26% of the scores and is known as 'Standard Deviation'.

For this, one may actually start with the details that one would have for finding out the Mean. Then, proceed to do some more calculations.

C.I.	f	midpoint x'	Deviation x'	fx'	fx'^2
23–25	10	24	3	30	90
20–22	5	21	2	10	20
17–19	3	18	1	3	3
14–16	7	15	0	0	0
11–13	3	12	–1	–3	3
8–10	3	9	–2	–6	12
5– 7	2	6	–3	–6	18

$$\Sigma fx' = 28 \qquad \Sigma fx'^2 = 146$$

Just as it was done in the case of Mean, here also one has to start with an Assumed Mean. And, here, as can be noted from the above Table, A.M. = midpoint of the C.I. 14–16 i.e. A.M. = 15.

The last column fx'^2 is obtained by multiplying the values in the previous two columns.

Thus, $fx'^2 = fx' \times x'$

One may calculate the Standard Deviation (S.D. or σ) by using the following formula

$$\text{S.D.} = i\sqrt{\frac{\Sigma fx'^2}{N} - c^2}$$

where f = frequency

x' = deviation of the midpoint of the class interval from assumed mean

N = Total number of scores

i = length of the C.I.

c = correction in terms of the class interval

In the present case

$\Sigma fx'^2 = 146$

$N = 33$

$i = 3$ and $\Sigma fx' = 28$

While learning about Mean it was noted that

$$c = \frac{\Sigma fx'}{N}$$

Therefore, in the formula for S.D.

instead of c^2 one may write $\left(\frac{\Sigma fx'}{N}\right)^2$

Now, substituting the above values in the formula

$$\text{S.D. or } \sigma = i\sqrt{\frac{\Sigma fx'^2}{N} - c^2}$$

$$= 3\sqrt{\frac{\Sigma fx'^2}{N} - \left(\frac{28}{33}\right)^2}$$

$$= 3\sqrt{\frac{146}{33} - \left(\frac{28}{33}\right)^2}$$

$$= 3\sqrt{4.424 - 0.719}$$

$$= 3\sqrt{3.705}$$

$$= 3 \times 1.925$$

$$= 5.77$$

Standard Score

It is known that to compare the performance of a student on different subjects or different characteristics, one cannot directly compare the corresponding scores on any test.

In view of this, one may convert these scores on different tests, which are to be understood with respect to different Means and Standard Deviations of the complete set of scores, into standard scores. This is done by applying the formula

$$Z = \frac{X - M}{S.D.}$$

Here, Z is the converted score
X is the original score
M is the Mean
and S.D. is the Standard Deviation

These scores are, accordingly, called standard scores.

These are also called Z - scores

This derived score distributions will always have $M = 0$ and S.D. = 1. Thus, they can be directly compared.

With the help of these converted scores teacher can directly compare the scores of students on tests of different subjects or characteristics.
