ASSESSMENT OF CHILDREN WITH SPECIAL NEEDS

Education is intended to provide all students with the skills and competencies they need to enhance their lives and the lives of their fellow citizens. This function would be extremely difficult even if all students entered school with the same abilities and competencies and even if students learned in the same way and at the same rate. However, they do not. Some are very smart, and some are not; some have mastered much of the first-grade curriculum before they enter school, whereas others need unusual amounts of help to learn the same material; some are fluent in English, and others are not; many have appropriate school behavior, and some do not.

In addition, the students attending schools today are a much more diverse group than in the past. Today's classrooms are **multicultural, multiethnic, and multilingual**. Students demonstrate a significant range of academic skills; in some large urban environments, for example, sixth graders are reading more than 2 years below grade level, and there is as much as a 10-year range in skill level in Math in a sixth-grade classroom.

Most of children and youth are attending schools in their own neighborhoods—this was not always the case in the past—and fewer students with disabilities are in separate buildings or separate classes, instead, now a days; these children are learning in classes with their peers. Thus, the focus of this course is on students who are attending their education in special and inclusive education. All students are entitled to a free and appropriate public education. The job of schools and the personnel who work in them are to enhance the competence of all students, and we are to build the capacity of systems (broadly conceived as communities, schools, parents and caregivers, and service agencies) to meet the needs of individual students. School personnel are confronted with the significant challenge of meeting the needs of a very diverse group of students. This is why assessment is such an important activity in all educational and Psychological arenas.

SECTION I: ASSESSMENT IN CLASSROOMS

1.1.Definition of Basic Concepts in Assessment

➕ **Test,** most persons think of test as; school examinations, college entrance examinations, or employment tests involving writing or marking answers. This constitutes a narrow view of the world of testing. However, the broader and more inclusive definition of an educational and psychological test is defined as **a systematic** procedure for observing (i.e., getting information) and describing one or more characteristics of a person with the aid of either a numerical scale (measurement such as test scores) or a category (qualitative means).

A test is a task or a serious of tasks or questions that students must answer or perform. It is used to get information regarding the extent to which the students have mastered the subject matter taught and the attainment of instructional objectives. A test is a particular form of measurement. It is a formal, systematic procedure to gather information about student's behavior or performance.

🔸 **Measurement is** concerned with systematic collection, quantification, and ordering of information gathered through different mechanisms**. It is a process of quantifying or assigning a number to performance according to explicit rules.** In other words, we assign numbers to any behaviors, characteristics, properties or attributes (for example, academic performance of students) based on agreed upon rules. For instance, if a person tells you the size of a table he measure is 45, what do you understand by this number? What could be your next question to the person? If a student said, "I scored 25," what further information would you need to understand the meaning of the score? In both case you need the units. In the first instance, you may ask, "45 what?" Centimeters, meters, millimeters, inches… What is the unit? In the latter case, 25 out of what maximum score? Out of 100, 50, 25? Thus assigning number to the length of the table (i.e., the attribute of the table) and to the behavior (i.e., performance) of the student is not enough. We need a rule (for example, the units).

Measurement can take many forms, ranging from the application very elaborate and complex electronic devices, to paper – and- pencil exams, to rating scales or checklists.

🔸 **Assessment** is a general term that includes all the different ways teachers gather information in their classrooms from their students. Assessment includes observations, oral questions, paper-and-pencil tests, homework, laboratory work, research paper, project works, field reports, and the like. It is a process of collecting, synthesizing, and interpreting information to aid in decision-making (Nitko, 1996; and Airasian, 1996). Assessment is concerned with the totality of the educational setting and is an inclusive term, that is, it subsumes measurement and others that cannot be directly measured. Assessment focuses not only on the nature of the learner, but also on what is to be learned and how (Payne, 1997).

🔸 **Evaluation is** the process of making judgment about student's performance, instruction, or classroom climate based on the data we gathered through various techniques. It occurs after assessment. It depends on the information that has been collected, synthesized and thoroughly interpreted. This is an important aspect of the teaching-learning process. It is the point where the instructor is in a position to make informed judgments about students' learning progress and the effectiveness of teaching.

Evaluation includes both quantitative and qualitative descriptions of student behavior plus value judgment concerning the desirability of that behavior. The following simple mathematical arrangement shows the relationship between measurement and evaluation.

Evaluation = Quantitative description of students' behavior (**measurement)** + qualitative description of students' behavior (**non-measurement**) + **value judgment**

Thus, evaluation may or may not be based on measurement (or tests) but when it is, it goes beyond the simple quantitative description of students' behavior. Evaluation involves judgment.

From instructional standpoint, **evaluation** may be defined as a systematic process of determining the extent to which instructional objectives are achieved by the learners. There are two important aspects of this definition. First, evaluation implies a systematic process, which omits causal, uncontrolled observation of students. Second, evaluation always assumes that instructional objectives have been previously identified. Without previously determined objectives, it is almost impossible to judge the nature and extent of students' learning progress.
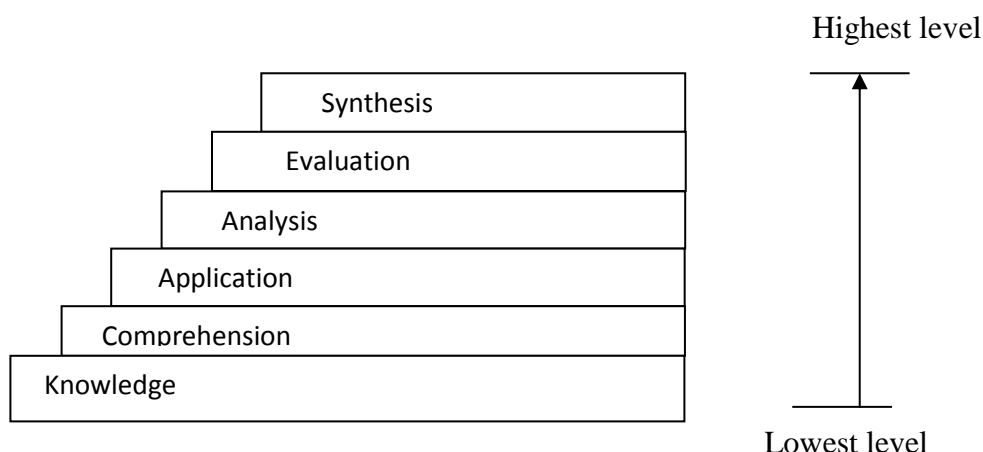
1.2 Teacher Made Tests or Classroom Tests

1.2.1 Educational Objectives

Educational Objectives are important for preparing tests. These educational **objectives can be classified into three major domains of objectives** based on behavioral constructs. **(Bloom's Taxonomy of Educational Objectives**): Bloom and his associates have developed systems for classifying learning outcomes and identifying behavior that can be expected of learners. They refer to their respective classification schemes as taxonomies of behavioral objectives, since they are intended as targets or goals of instruction.

Accordingly, they have categorized objectives into three domains:

1. **The cognitive domain**; which represents intellectual abilities and skills.

    Example: - Remembering of facts

                        - Solving problems mentally

2. **The Affective domain**; which represents attitudes, feeling, interest, beliefs and values.

    Example: - Desire to read

                        - Willingness to work at something

3. **Psychomotor domain**; which represents bodily movements and physical performance.

    Example: Muscle flexibility

1. The Cognitive Domain: This domain contains six hierarchical arranged levels. It ranges from the knowledge level (simplest) to the evaluation level (most complex). The higher level objectives are assumed to include, and be dependent on, lower level cognitive skills. The hierarchy is indicated below by the diagram.

Highest level

| Synthesis |
| Evaluation |
| Analysis |
| Application |
| Comprehension |
| Knowledge |

Lowest level

The most commonly assessed school behaviors are in the cognitive domain. Cognitive behaviors include a range of intellectual activities such as memorizing, interpreting, applying, problem solving, analyzing, and thinking critically. Bloom's cognitive taxonomy (classification) is organized with six levels, with each level representing a more complex type of cognitive thinking or behavior. Starting with the simplest and moving to the most complex, the six levels are: knowledge, comprehension, application, analysis, synthesis and evaluation.

Knowledge (Remembering previously learned material): Objectives at the knowledge level require the learner to remember or recall information such as facts, terminology, problem solving strategies, or rules.

Some action verbs that are used to describe learning outcomes at knowledge level are:

Define	Identify	Outline	Recall	List
Select	State	Recite	Match	Name

Example

1. Students will list the six levels of cognitive domain.
2. State the formula for the area of a circle.

Comprehension (Grasping the meaning of material): Objectives at this level require some degree of understanding. Learners are expected to be able to change the form of a communication, translate, restate what has been read, see connections or relations between parts of a communication, or draw conclusions from information or see the consequences of it.

Some action verbs that are used to describe learning objectives at this level are:

defend	*predict*	*estimate*	*convert*	*explain*	*Infer*
summarize	*distinguish*	*discriminate*	*extend*	*paraphrase*	

Example : Given the mathematical formula for the area of a circle, paraphrase it using your own words.

Application (Using information in concrete situations): Objectives written at this level require the students to use previously acquired information in a setting other than the one in which it was learned. Application objectives differ from comprehension objectives in that application questions require the presentation of a problem in a different and often applied context.

Some action verbs that are used to describe learning outcomes are:

change	*modify*	*produce*	*develop*	*employ*
compute	*solve*	*relate*	*organize*	*prepare*
demonstrate	*operate*	*transfer*	*use*	

Example: Compute the area of actual circles.

Analysis (Breaking down material into parts): Objectives written at the analysis level require the learner to identify logical errors (e.g. point out a contradiction or an erroneous inference) or to differentiate among facts, opinions, assumptions, hypotheses, or conclusions. Questions at the analysis level often require the student to draw relationships among ideas or to compare and contrast. Some action verbs that are used to describe learning outcomes at this level are:

break down	*distinguish*	*point out*	*differentiate*	*subdivide*	*diagram*
deduce	*illustrate*	*relate*	*outline*	*separate out*	*inter*

Example

1. *Given a math word problem, determine the strategies that would be necessary to solve it.*
2. *Given a presidential speech, the student will be able to point out the positions that attack an individual rather than the individual's program.*

Synthesis (Putting parts together into a whole)**:** Objectives written at the synthesis level require the student to produce something unique or original. At this level, students are expected to solve an unfamiliar problem in a unique way or to combine parts to form a unique or novel whole.

Some action verbs that are used to describe learning outcomes at this level are:

compile	*categorize*	*rewrite*	*design*	*devise*
create	*compose*	*summarize*	*formulate*	

Example

1. *Integrate several different strategies to solve a mathematical problem.*

2. *Given a short story, the student will write a different but plausible ending.*

Evaluation (Judging the value of a product for a given purpose, using definite criteria): Instructional objectives written at this level require the learner to form judgment and make decisions about the value or worth of methods ideas, people, or products that have a specific purpose. Students are expected to state the bases for their judgments (e.g. the criterion or principles they used to reach their conclusions).
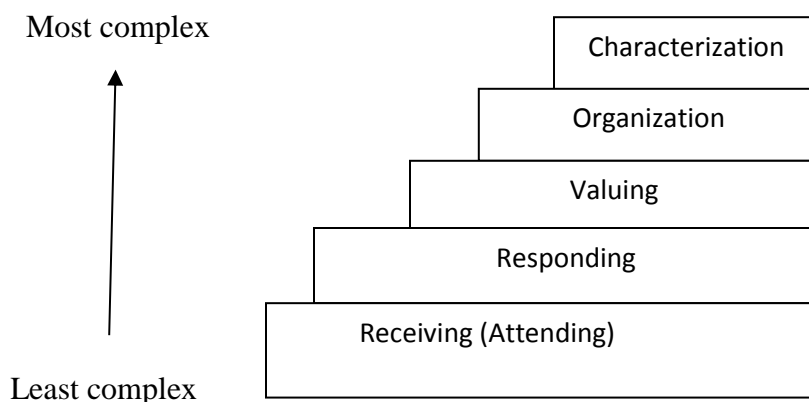
Some action verbs that describe learning outcomes at the evaluation level are:

appraise	*criticize*	*support*	*conclude*	*contrast*
compare	*defend*	*validate*	*interpret*	*justify*

Example

1. *When you have finished solving a problem determine the degree to which that problem was solved as efficiently as possible.*
2. *Given a previously unread short story, the student will criticize the content and form of the story.*

2. The Affective Domain: The second domain of behavior is affective domain. Affective behavior involves feelings, attitudes, interests, preferences, values, and emotions. Emotional stability, motivation, trustworthiness, self-control and personality are all examples of affective characteristics. Although affective behaviors are rarely assessed formally in schools and classrooms, teachers constantly assesses affective behavior informally. They describe their students affective characteristics, based on their informal observations and interactions with the students. Krathwohl, Bloom and Masia (1964) classified affective domain into five levels of hierarchy. The levels of affective taxonomies contain low involvement behaviors such as simply paying attention while the higher level contains high involvement behavior characterized by strong interest commitment, and valuing.

Most complex

Characterization

Organization

Valuing

Responding

Receiving (Attending)

Least complex

Receiving (Attending): The student is aware of or passively attending to certain phenomena and stimuli (that is, listening, observing). Some action verbs that describe learning outcomes at receiving level are:

listen	*attend*	*share*	*look*
notice	*be aware*	*control*	*hear etc…*

Example: The student will be able to listen to all music concerts by Teddy Afro without leaving his/her seat.

Responding: The student's response at these levels indicates more than passive listening and attending; they require active participation. In the most basic form of responding, a student might choose some other activity. More complete responding would be indicated by a student's willingness to engage in an activity, even when allowed a choice. The highest level with in this category is indicated by satisfaction after engaging in a response.   Some action verbs that describe learning outcomes at responding level are:

follow	*play*	*Practice*	*participate*
discuss	*applaud*	*comply*	*obey*

Example: The student will follow the directions given in the book without argument when asked to do so.

Valuing: The student displays a behavior consistent with a single belief or attitudes in situations where he/she is not forced to comply or obey (i.e. demonstrates a definite preference, displays a high degree of certainty and convocations). Some action verbs that describe learning outcomes at valuing level are:

help	*express*	*act*	*argue*	*display*
debate	*organize*	*convince*	*prefer*	

Example: The student will express an opinion about unclear disarmament whenever national events raise the issue.

Organization: The students is committed to a set of values and displays or communicates his or her beliefs or values (i.e. develops a rationale for a set of values; makes judgments about sets of values). Some action verbs that describe learning outcomes at this level are:

select	*formulate*	*balance*	*abstract*
decide	*systematize*	*compare*	*define*

Example: The students will be able to formulate the reasons why she or he supports civil rights legislation that does not support his or her belief.

Characterization: Objectives at the characterization level require that all the behavior displayed by the student be consistent with his or her value. At this level, students not only acquired the behaviors at all

previous levels but also have integrated their values into a system representing a complete and persuasive philosophy. Evaluation of this level of behavior involves the extent to which the student has developed a consistent philosophy of life (example, exhibits respect for the worth and dignity of human beings in any and all situations). Some action verbs that describe learning outcomes at this level are:

display	*manage*	*exhibit*	*require*	*internalize*
avoid	*resolve*	*revise*	*resist*	

 Example

The student will exhibit a helping attitude toward handicapped students by assisting with their mobility both in and outside the classroom.

3. **The Psychomotor Domain:** This level deals with the skill of the students. This domain includes behaviors that rely heavily on muscular system. For example, running, jumping, speaking, typing, riding bicycle, and driving cars, etc. It has five hierarchically arranged levels. The following diagram shows the levels from least complex (imitation) to most complex (naturalization).
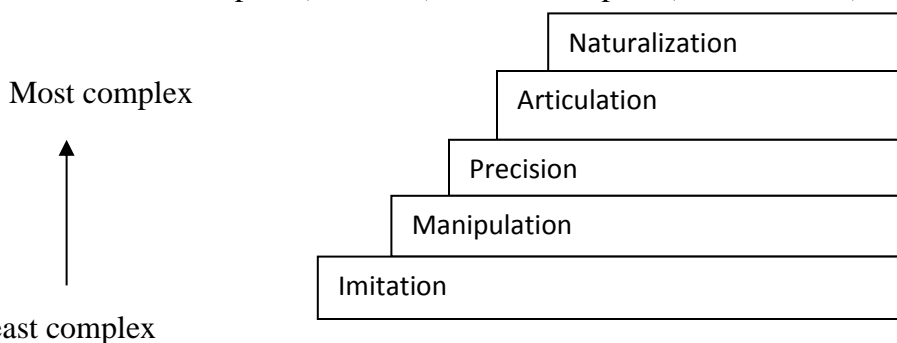
Imitation: Objectives at this level require the student to be exposed to an observable action and then overtly imitate that action, such as when a teacher demonstrates use of the microscope by placing a slide on the specimen tray. The performance is generally crude and imperfect. At this level students are expected to observe and be able to repeat the action being visually demonstrated.

   Some action verbs that describe outcomes at this level are:

repeat	*hold*	*follow*
place	*grasp*	*balance*

Example:

After being shown a free-hand drawing of a parallelogram, the student will be able to reproduce the drawing.

Manipulation: Objectives at this level require the student to perform, or practice selected actions from written or verbal directions without the aid of visual modern or direct observation. Action verbs at this

level are the same as that of the imitation level except that they are performed from spoken or written instructions.

Example: With the instruction on the handout in front of you, practice focusing your microscope until the outline of the specimen can be seen.

Precision: Objectives at this level require the student to perform an action independent of either a visual model or a written set of directions.

Students are expected to reproduce the action with control and to reduce errors to a minimum. Some action verbs that describe outcomes at this level include performing the behavior:

Accurately	*with control*	*proficiently*
Independently	*without error*	*with balance*

Example: The student will be able to accurately place the specimen on the microscope tray.

Articulation: Objectives at this level require the student to display the coordination of a series of related acts by establishing the appropriate sequence and by performing the acts accurately, with control as well as with the speed and timing. Some action verbs that describe outcomes at this level include performing the behavior.

harmony	*speed*	*confidence*	*coordination*	*proportion*
integration	*timing*	*stability*	*smooth mass*	

Example

The student will be able to accurately complete ten simple arithmetic problems on an electronic calculator quickly and smoothly with 90 second.

Naturalization: Objectives at this level require a high level of proficiency in the skill or performance being taught. At this level, the behavior is performed with the least expenditure of energy and becomes routine, automatic, and spontaneous. Students are expected to repeat the behavior naturally and effortlessly time and again. Some action verbs that describe outcomes at this level include performing the behavior:

naturally	*effortlessly*	*professionally*	*routinely*
with ease	*automatically*	*with perfection*	*spontaneously*

Example

At the end of the lesson, students will be able to automatically draw correct right triangles without the aid of a template.

1.2.2 Classification of Achievement Test

Achievement test may be classified in the following ways:

1. By mode of Response.
   - Oral test
   - Written test
   - Practical test

2. By Purpose of Testing
   - Placement Test
   - Formative Test
   - Diagnostic Test
   - Summative Test

3. By Desired Speed of Response
   - Power test
   - Speed test

4. By Degree of Accuracy Employed in Preparation an Applicability
   - Teacher Made Tests
   - Standardized Tests

5. By Mode of Interpreting Results
   - Norm –referenced Testing
   - Criterion-referenced Testing
   - Self-referenced Testing

6. By Format of Test Items
   - Objective Test Items
   - Essay Test Items

The classroom test, which is otherwise called teacher-made test, is an instrument of measurement and evaluation. It is a major technique used for the assessment of students learning outcomes.

The classroom tests can be achievement or performance test and/or any other type of test like practical test, etc. prepared by the teacher for his specific class and purpose based on what he has taught.

The development, of good questions or items writing for the purpose of classroom test, cannot be taken for granted. An inexperienced teacher may write good items by chance. However, this is not always possible. Development of good questions or items must follow a **number of principles** without which no one can guarantee that the responses given to the tests will be relevant and consistent. In this part, we shall examine the various aspects of the teacher's own test.

As a teacher, you will be faced with several problems when it comes to your most important functions – evaluating of learning outcomes. You are expected to observe your students in the class, workshop, laboratory, field of play etc and rate their activities under these varied conditions. You are required to correct and grade assignments and home works. You are required to give weekly tests and end of term examinations.

Most of the times, you are expected to decide on the fitness of your students for promotion on the basis of continuous assessment exercises, end of term examinations' cumulative results and promotion examination given towards the end of the school year. Given these conditions it becomes very important that you become familiar with the planning, construction and administration of good quality tests. This is because in the next few years when you graduate as a teacher your tests will be used to play a very important role in the growth and progress of your country's Youth.

Some Pit Falls or Short Comings in Teacher Made Tests: Educational objectives are classified by Bloom et al (1956) into three domains and different levels such as recall or memory or knowledge comprehension, application, analysis, evaluation, and synthesis. It means that you do not only **set objectives** along these levels but also **test them along** the levels. The following observations have been made about teacher-made tests.

They are listed below in order to make you avoid them when you construct your questions for your class tests.

i. *Most teacher-test are **not appropriate** to the different levels of learning outcomes. The teachers specify their instructional objectives covering the whole range simple recall to evaluation. Yet the teachers' items fall within the recall of specific facts only*

ii. *Many of the test **exercises fail to measure what they are supposed to measure**. In other words most of the teacher-made tests are not valid. You may wonder what is validity. It is a very important quality of a good test, which implies that a test is valid if it measures what it is supposed to measure.*

iii. *Some classroom tests **do not cover comprehensively** the topics taught. One of the qualities of a good test is that it should represent the entire topic taught. However, these tests cannot be said to be a representative sample of the whole topic taught.*

iv. *Most tests prepared by teacher **lack clarify in the wordings**. The questions of the tests are ambiguous, not precise, not clear and most of the times carelessly worded. Most of the questions are general or global questions.*

v. *Most teacher-made tests **fail item analysis** test. They fail to discriminate properly and not designed according to difficulty levels. These are not the only pit falls. Nevertheless, you should try to avoid both the ones mentioned here and those not mentioned here. Now let us look at the various types of teacher-made tests items.*

1.2.3 Types of Test Forms used in The Classroom: There are different types of test forms used in the classroom. These can be **essay test items** and **objective test items**. Therefore, we are going to describe

the essay test items and objectives test items. These are the most common tests, which you can easily construct for your purpose in the class.

1. **Essay Test Items:** Essay tests are tests consisting of questions (items) designed to elicit from the learners through freedom of response the extent to which they have acquired the behaviour called for in the course objectives. The answers to such questions, which the learners are confronted vary in quality and degree of correctness. Most times, these answers are not complete and thorough.

They also have **poor psychometric quality** or **measurement qualities** although popular among classroom teachers especially those who are deficient in the skill required for item construction. For this reason, we should examine how to construct essay items in this part and subsequently examine how to administer and score such items in order to improve their reliability and validity.

When to use Essay Questions

✓ *You should use essay questions in the measurement of complex achievement when its distinctive feature of freedom of response is required. Learners are free to select, relate and present ideas in their own words. This freedom enhances the value of essay questions as a measure of complex achievement but it introduces scoring difficulties that make them insufficient as measure of factual knowledge.*

✓ *Essay questions should also be used to measure those learning outcomes that cannot be measured by objective test items. The specific features of essay questions can be utilized most fully when their shortcomings are offset by the need for such measurement.*

✓ *They should be used when learning outcomes concerned with the abilities to select, organize, integrate, relate, and evaluate ideas require the freedom of response and the originality provided by essay questions. More so, when these outcomes are of such great educational significance that the expenditure of energy in the difficulty and time-consuming task of evaluating the answers can be easily justified.*

Advantages of Essay Test

✓ *It measures complex learning outcomes that cannot be measured by other means. For instance, it has the ability to measure learner's communication skills. That is, the learner's ability to produce an answer, synthesize and organize ideas and present them readably in a logical and coherent form. This is the major advantage.*

✓ *It also enables the measurement of organizational and divergent thinking skills by laying emphasis on the integration and application of thinking and problem solving skills, creativity and originality.*

✓ *It is very applicable for measuring learning outcomes at the higher levels of educational objectives such as application, analysis, synthesis and evaluation of levels of the cognitive domain.*

✓ *It is easy and economical to administer. It can be easily and conveniently written on the chalkboard because of the few items involved. This saves materials and time for production.*

✓ *Essay item is easy to construct and does not take much time. This fact has to be guarded seriously to avoid constructing questions that can be very misleading by not asking for specific behaviours emphasized in a particular set of learning outcomes.*

✓ *It can be used to measure in-depth knowledge especially in a restricted subject matter area.*

✓ *It does not encourage guessing and cheating during testing.*

Disadvantages of Essay Test: Despite the advantages already proffered/proposed for essay test, it does not satisfy the two most important qualities of a good measuring instrument. Its disadvantages include that:

✓ *It is inadequate in sampling subject matter content and course objectives since it provides limited sampling. The provision of few questions results in the invalid and narrow coverage of subject matter and instructional objectives. Also as Nenty (1985), rightly pointed out, fewness of the number of questions often asked encourages permutation of some content areas and creaming of ideal responses to suspected questions. In this regard, essay questions discourage the development of good study habit.*

✓ *In addition to the invalidity of the measurement, evaluating the answers to carelessly developed questions tends to be confusing and time-consuming task. This results in poor reliability in scoring. Studies have shown that answers to essay questions are scored differently by different teachers and that even same teachers score the answers differently at different times. A variation, which ranges from near perfect scores to those that are representing dismal failure. This may be attributed to the inability of scorers to clearly identify the learning outcomes being measured. When the evaluation of answers is not guided by clearly defined outcomes, it tends to be based on less stable, initiative judgments.*

✓ *Sometimes an essay question implies many skills other than that which the item was intended to measure. The testee therefore perceives and reacts to the same questions differently. The differences in the perception of the questions encourage bluffing and hides differences in the knowledge of basic factual material and the learners ability to use and organize such facts.*

✓ *The essay test item does not readily lend itself to empirical study of item qualities like difficulty and discrimination based on which improvements on the item could be made.*

Classification of Essay Test Items: Essay questions are classified into two types namely the restricted response type and the extended response type. The classification is based on the degree of freedom of response associated with the question. For instance, an essay question may require just a few short

sentences as answer as in the short-answers objective item where a sentence or two could be all that is required. Whereas, another essay question may give the examinees complete freedom in making their responses and their answers may require several pages.

However, there are variations in freedom of response that fall within these extreme conditions. But for convenience, essay questions are presently classified as unrestricted response type in which examinees are given almost complete freedom in making their responses and the restricted response type in which the nature, length or organization of the response is limited.

Extended Responses Essay Questions: These are responses to essay questions in which the examinee is only restricted by time as no bound is placed as regards the depth, breadth and the organization of the response.

An example of questions in this category include:

i. Describe the sampling technique used in research studies.

ii. Discuss the social, religious and economical implication of religious riots/uprisings in Ethiopia.

iii. Explain the various ways of preventing accident in a school workshop or laboratory.

In response to such a question the examinee demonstrates his ability to select and recall the facts, which he thinks are pertinent, organize and present his ideas in a logical and coherent form. This freedom to decide which facts the thinks is most pertinent to select his own method of organization and to write as much as seems necessary for a comprehensive answer tends to reveal the ability to evaluate ideas, relate them, coherently and to express them succinctly/briefly. In addition, they expose the individual differences in attitudes, values and creative ability. This type of essay item is mostly useful in measuring learning outcomes at the higher cognitive levels of educational objectives such as analysis, synthesis and evaluation levels. Although, the extended response essay type are also limited by two weaknesses, which are:

✓ *They are insufficient for measuring knowledge of factual materials because they call for extensive details in selected content area at a time.*

✓ *Scoring such responses is usually difficult and unreliable since the examinees have free will in the array of factual information of varying degree of correctness, coherence and expression. These limitations are minimized in the Restricted Response Type.*

Restricted Response Essay Questions: In this type the examinee is limited to the nature, length or organization of response to be made. The items are directional questions and are aimed at the desired responses. This limits the examinee freedom to select, recall, and synthesize all that he knows and to

present them logically as he may wish. This type of essay item is most useful in measuring learning outcomes at the lower cognitive levels of educational objectives, that is, knowledge, comprehension and application levels.

Examples of restricted response essay question include:

- ✓ *Give three advantages and two disadvantages of essay tests.*
- ✓ *State four uses of tests in education.*
- ✓ *Explain five factors which influence the choice of building site.*
- ✓ *Mention five rules for preventing accident in a workshop.*

The restricted nature of the expected response in this type of items makes it more efficient for measuring knowledge of factual material. It reduces to a reasonable extent the difficulty of scoring and encourages more reliability in scoring. However, the restriction makes it less effective as a measure of ability to select, organize and integrate ideas and present them in an original and coherent form, which is one of the major advantages of essay test.

Constructing the Essay Questions: You are now aware of the handicaps of essay questions as a measuring instrument. Therefore, an essay test is a useful measurement instrument only to the extent that it is constructed, administered and scored to ensure a high level of objectivity. For this reason, essay test items should consist of items that will ensure the same understanding and elicit only the skill or ability one is interested in measuring from every examinee. Also, the responses are to be such to which two or more examiners would assign the same score and should attract consistent interpretation from everybody. You know that this is difficult to achieve and needs a lot of effort. Hence the following points are suggested as guide for construction of good essay test item that call for the desired behaviour.

i. *Restriction of the use of essay questions to only those learning outcomes that cannot be satisfactorily measured by objective items. That is, essay questions are to be used only when it's desirable and very adequate for measuring the learning outcomes for full realization of learner's achievement. In other words, they are to be used for questions that call for complex learning outcomes that pertain to the organization, integration and expression of ideas which would not have been possible without the use of essay test items.*

ii. *Formulation of questions that call forth the behaviour specified in the learning outcomes. Essay questions should be designed to elicit only the skill which the item was intended to measure. This can be achieved by expressing clearly and precisely the question in line with clearly defined instructional objective.*

In addition, an action verb like compare, contrast, illustrates, differentiates, criticized and so on could be used to give the test items more focus.

iii. Phrase each question to clearly indicate the examinees task. An essay question has to specify precisely what is required of the examinee. Ensure that the testee's task is clearly indicated by delimiting the area covered by the item, using descriptive words to give specific direction towards the desired response. Indicate the score allotted to the test. This suggestion easily lend itself to restricted response type and care should be taken not to narrow the questions when constructing the extended response type in order not to reduce it's effectiveness as a measuring of the ability to select, organize and integrate ideas. Also, adapt the length and complexity of the answer to the testees' level of maturity.

iv. Indication of approximate time limit for each question. It is necessary to indicate time allotted to each question to enable the testees to pace their writing on each question and to allay any anxiety that might arise. The timing should take care of slower testees writing speed so as not to put them at disadvantage for a satisfactory response.

v. Avoidance of the use of optional questions. The provision of optional questions although generally favoured by testees obviously means that they are taking different tests and therefore the common basis for evaluating their achievement is lost.

Moreover, optional questions might also influence the validity of test results since some examinees may be favoured in their advanced preparation of selected areas of study. It is also not easy to construct essay questions of the same difficulty level. Hence, making valid comparisons of performance among them especially for norm reference setting will not be possible.

2.**Objective Test Items:** Objective tests are those test items that are set in such a way that one and only one correct answer is available to a given item. In this case, every scorer would arrive at the same score for each item for each examination, even on repeated scoring occasions. This type of items sometimes calls on examinees to recall and write down or to supply a word or phrase as an answer (free – response type). It could also require the examinees to recognize and select from a given set of possible answers or options the one that is correct or most correct (fixed-response type). This implies that the objective test consists of items measuring specific skills with specific correct response to each of the items irrespective of the scorer's personal opinion, bias, mood or health at the time of scoring.

When to use objective test

- ✓ It is used when highly structured task are needed to limit the type of response the examinees can make and to obtain correct answers from learners by demonstrating the specific knowledge or skill called for in the item.

- ✓ It is used to appraise more effectively the achievement of any of the educational objectives of simple learning outcomes as well as the complex outcomes in the knowledge, understanding, and application and even in higher levels covering large content areas if skillfully constructed. It is possible to set as many as 120 objective tests spread over many lesson units and several cognitive levels of educational objective for one hour or two hours.

- ✓ It is used when objective, quick, easy and accurate scoring is desired especially when the number of examinees is large.

- ✓ It is used to measure understanding, thinking skills and other complex learning outcomes of the learners.

- ✓ It can also be used for diagnosis of learning deficiency and the result used for remediation process.

Advantages of Objective Test

- ✓ Objective test enhances the assessment of learner's response to test items because the scoring is not influenced by the scorer's bias or disposition at the time of scoring but by the correctness of the answer. By overcoming the subjectivity of the essay test, the reliability of the test as measuring instrument is enhanced.

- ✓ Scoring of objective test is easy and takes little time. It is also scored by a machine and facilitates high efficiency in testing a large number of examinees.

- ✓ The result of objective test especially the multiple choice items can be used for diagnostic purposes since they provide clues for factual errors and misunderstanding that need remediation.

- ✓ It is adequate for sampling the subject matter and instructional objectives of the course because the relatively large number of items set enhances effective coverage of the content areas on which the test is based. The result provides a more valid and reliable ability of the examinees performance.

- ✓ It is efficient for measuring knowledge of facts. It can also be designed to measure understanding, thinking skills and other complex outcomes.

- ✓ Objective test items can be pre-test, refined through item analysis, standardized and reused a number of times if properly handled.

✓ *It is fair to all examinees since it does not call on other skills outside the skill it is intended to measure. That is, its validity is not affected by good handwriting, bluffing or the verbiage.*

Disadvantages of Objective Test

✓ *It does not encourage the development of examinees originality in desirable skills such as the ability to select, organize or synthesize ideas and to present them correctly in a logical and coherent form. The complete structuring of task is not suitable for assessing learning abilities in this form.*

✓ *It tends to measure only factual knowledge. This disadvantage can be overcome by developing items for the objective items rigorously following the steps involved in item development process.*

✓ *Development of good objective test items requires training of test developers in the skills necessary for constructing effective, valid and reliable items.*

✓ *It needs time, commitment and adequate planning.*

✓ *Objective test items lend themselves to guessing especially when the test items are not skillfully developed.* An examinee can guess correctly on few items and earn some undeserved points even in a well-constructed objective test. It is also easier to cheat in an objective test than in essay test if the test is poorly administered.

Types of Objective Test: The objective test can be classified into those that require the examinee to supply the answer to the test items (free-response type) and those that require the examinee to select the answer from a given number of alternatives (fixed response type). The free-response type consists of the short answer and completion items while the fixed response type is commonly further divided into true-false or alternative response matching items and multiple-choice items.

The Free Response Test Items: The free response type of objective test tends to represent a compromise between the essay and the objective items. The free response type namely short-answer item and the completion item both are supply-type test items consisting of direct questions which require a short answer (short-answer type) or an incomplete statement or question to which a response must be supplied by an examinees (completion type). The answers to such questions could be a word, phrase, number or symbol. It is easy to develop and if well developed, the answers are definite and specific and can be scored quickly and accurately. An example of question in the class is:

Short Answer: Who was the first president of Debre Markos University? (Tewodros Bekafa, PhD)

Completion: The name of the first president of Debre Markos University is---------------(Tewodros Bekafa, PhD).

The free – response type is very adaptable for item construction in mathematics, physical sciences and other areas where questions are computational problems requiring examinees to supply the solutions.

Uses

✓ *It is suitable for measuring a wide variety of relatively simple learning outcomes such as recall of memorized information and problem solving outcomes measured in mathematics and sciences.*

✓ *It can be used to measure the ability to interpret diagrams, charts, graphs and pictorial data.*

✓ *It is used when it is most effective for measuring a specific learning outcome such as computational learning outcomes in mathematics and sciences.*

Advantages

✓ *It measures simple learning outcomes, which makes it easier to construct.*

✓ *It minimizes guessing because the examinees must supply the answer by either think and recall the information requested or make the necessary computations to solve the problem presented. It is unlike the selection item where partial knowledge might enable the examinee to choose the correct answer.*

Disadvantages

✓ *It is not suitable for measuring complex learning outcomes. It tends to measure only factual knowledge and not the ability to apply such knowledge and it encourages memorization if excessively used.*

✓ *It cannot be scored by a machine because the test item can, if not properly worded, elicit more than one correct answer. Hence the scorer must make decision about the corrections of various responses. For example, a question such as "Where was scientist Aklilu Lema born?" Could be answered by name of the town, state, country or even continent. Apart from the multiple correct answers to this question, there is also the possibility of spelling mistakes associated with free-response questions that the scorer has to contend with.*

The Alternative Response Test Item

The alternative response test item commonly called the true-false test item because the true-false option is commonly used consists of item with declarative statement to which the examinee is asked to give either of two options concerning the item. The two options could be true or false, right or wrong, correct or incorrect, yes or no, fact or opinion, agree or disagree and so on.

Most times the alternative response item includes opinion statement and the examinee is also required to response to them as merely true or false. The opinion item is not desirable from the standpoint of testing,

teaching and learning. If opinion statement is to be used, it has to be attributed to some source thereby making it possible to assign the option of true or false to the statement based on knowledge concerning the belief held by an individual or the values supported by an organization or institution. An example of alternative response item is as follows:

Read the following statement if the statement is true circle the **T** if it is false circle the **F.**

T. F Solar Energy is the energy radiated from the sun.

The correct answer to the example above is true and is always true.

Uses

✓ It is commonly used to measure the ability to identify the correctness of statements of fact, definitions of terms, statements of principles and other relatively simple learning outcomes to which a declarative statement might be used with any of the several methods of responding.

✓ It is also used to measure examinee ability to distinguish fact from opinion; superstition from scientific belief.

✓ It is used to measure the ability to recognize cause – and – effect relationships.

✓ It is best used in situations in which there are only two possible alternatives such as right or wrong, more or less, and so on.

Advantages

✓ It is easy to construct alternative response item but the validity and reliability of such item depend on the skill of the item constructor. To construct unambiguous alternative response item, which measures significant learning outcomes, requires much skill.

✓ A large number of alternative response items covering a wide area of sampled course material can be obtained and the examinees can respond to them in a short period of time.

Disadvantages

✓ It requires course material that can be phrased so that the statement is true or false without qualification or exception as in the Social Sciences.

✓ It is limited to learning outcomes in the knowledge area except for distinguishing between facts and opinion or identifying cause – and – effect relationships.

✓ It is susceptible to guessing with a fifty-fifty chance of the examinee selecting the correct answer on chance alone.

The chance selection of correct answer has the following effects.

i. It reduces the reliability of each item thereby making it necessary to include many items in order to obtain a reliable measure of achievement.

ii. The diagnostic value of answers to guess test items is practically nil because analysis based on such response is meaningless.

iii. The validity of examinees response is also questionable because of response set.

Response set is a consistent tendency to follow a certain pattern in responding to test items. For instance some examinees will consistently mark "true" those items they do not know while others will consistently mark them "false. Any given test will therefore favour one response set over another thereby introducing an element into the test score that is irrelevant to the purpose of the test.

The Matching Test Items: The matching test items usually consist of two parallel columns. One column contain a list of word, number, symbol or other stimuli (premises) to be matched to a word, sentence, phrase or other possible answer from the other column (responses) lists. The examinee is directed to match the responses to the appropriate premises. Usually, the two lists have some sort of relationship. Although the basis for matching responses to premises is sometimes self-evident but more often it must be explained in the directions.

The examinees task then is to identify the pairs of items that are to be associated on the basis indicated. Sometimes the premises and responses list is an imperfect match with more list in either of the two columns and the direction indicating what to be done. For instance, the examinee may be required to use an item more than once or not at all, or once. This deliberate procedure is used to prevent examinees from matching the final pair of items on the basis of elimination.

An example of matching item is given below.

Direction: Match the measure of angles in column "A" with the names under column
"B" and write the letter of the correct answer in the space provided. A
response may be used more than once or not at all.

	<u>Column A</u>	<u>Column B</u>
_____1.	An angle that measures between 180^0 - 360^0	A. Acute angles B. Complementary angle
_____2.	An angle that measures between 90^0 - 180^0	C. Obtuse angle D. Reflex angle
_____3.	An angle that measures 90^0	E. Right angles
_____4.	An angle that measures between 0^0 - 90^0	F. supplementary angle

Uses

✓ *It is used whenever learning outcomes emphasize the ability to identify the relationship between things and a sufficient number of homogenous premises and responses can be obtained.*

✓ *Essentially used to relate two things that have some logical basis for association.*

✓ *It is adequate for measuring factual knowledge like testing the knowledge of terms, definitions, dates, events, references to maps and diagrams.*

Advantages

✓ *The major advantage of matching exercise is that one matching item consists of many problems. This compact form makes it possible to measure a large amount of related factual material in a relatively short time.*

✓ *It enables the sampling of larger content, which results in relatively higher content validity.*

✓ *The guess factor can be controlled by skillfully constructing the items such that the correct response for each premise must also serve as a plausible response for the other premises.*

✓ *The scoring is simple and objective and can be done by machine.*

Disadvantages

✓ *It is restricted to the measurement of factual information based on rote learning because the material tested lend themselves to the listing of a number of important and related concepts.*

✓ *Many topics are unique and cannot be conveniently grouped in homogenous matching clusters and it is sometimes difficult to get homogenous materials clusters of premises and responses that can sufficiently match even for contents that are adaptable for clustering.*

✓ *It requires extreme care during construction in order to avoid encouraging serial memorization rather than association and to avoid irrelevant clues to the correct answer.*

The Multiple Choice Test Items: The multiple choice item consists of two parts – a problem and a list of suggested solutions. The problem generally referred to as the stem may be stated as a direct question or an incomplete statement while the suggested solutions generally referred to as the alternatives, choices or options may include words, numbers, symbols or phrases. In its standard form, one of the options of the multiple choice item is the correct or best answer and the others are intended to mislead, foil, or distract examinees from the correct option and are therefore called distracters, foils or decoys. These incorrect alternatives receive their name from their intended function – to distract the examinees that are in doubt about the correct answer.

An example of multiple-choice item is given below.

Which one of the following factors contributed least to the selection of Addis Ababa as the Capital City of Ethiopia?

(A) Central location

 (B) Good climate

(C) Good highways

(D) Low population

(E) Good strategy for security

The best-answer form of Multiple Choice Item is usually more difficult than the correct answer form. This is because such items are used to measure more complex learning outcomes. It is especially useful for measuring learning outcomes that require the understanding, application or interpretation of factual information.

 An example is given below.

Which one of the following does best describe the property of speed?

 (A) It has magnitude

 (B) It has direction

(C) It is a scalar quantity

(D) It is a vector quantity

(E) It has magnitude and direction.

Uses

✓ *The multiple-choice item is the most widely used of the types of test available. It can be used to measure a variety of learning outcomes from simple to complex.*

✓ *It is adaptable to any subject matter content and educational objective at the knowledge and understanding levels.*

✓ *It can be used to measure knowledge outcomes concerned with vocabulary, facts, principles, method and procedures and also aspects of understanding relating to the application and interpretation of facts, principles and methods.*

✓ *Most commercially developed and standardized achievement and aptitude tests make use of multiple-choice items.*

Advantages

✓ *The main advantage of multiple-choice test is its wide applicability in the measurement of various phases of achievement.*

✓ *It is the desirable of all the test formats being free of many of the disadvantages of other forms of objective items. For instance, it present a more well-defined problem than the short-answer item,*

avoids the need for homogenous material necessary for the matching item, reduces the clues and susceptibility to guessing characteristics of the true-false item and is relatively free from response sets.

✓ *It is useful in diagnosis and it enables fine discrimination among the examinees on the basis of the amount of what is being measured possessed by them.*

✓ *It can be scored with a machine.*

Disadvantages

✓ *It measures problem-solving behaviour at the verbal level only.*

✓ *It is inappropriate for measuring learning outcomes requiring the ability to recall, organize or present ideas because it requires selection of correct answer.*

✓ *It is very difficult and time consuming to construct.*

✓ *It requires more response time than any other type of objective item and may favour the test-wise examinees if not adequately and skillful constructed.*

Constructing the Objective Test Items

You have seen that simple put a test item is a statement sometimes in question form that tries to elicit a testee's level of knowledge, ability or understanding of a specific subject matter. Therefore, writing a good test item is an art that requires some skill, time, perseverance, and creativity.

The following are some general guidelines for the construction of any type of objective test item.

✓ *The wording of the item should be clear and as explicit as possible.*

✓ *Avoid setting interrelated items*

✓ *Items should be designed to test important and not trivial facts or knowledge.*

✓ *Write an item to elicit discriminately the extent of examinees possession of only the desired behaviour as stipulated in the course instructional objectives answers.*

✓ *Ensure that there is one and only one correct or best answer to each item.*

✓ *Avoid unintentionally giving away the answer through providing irrelevant clues.*

✓ *Use language appropriate to the level of the examinees.*

✓ *Items in an achievement test should be constructed to elicit specific course content and not measure general intelligence.*

✓ *Have an independent reviewer to vet/check up your test items.*

1.3. **Using Test Adaptations and Accommodations**

Accommodations and Modifications

Accommodations do not reduce learning expectations. They meet specific instruction and assessment needs of students with disabilities and allow for educators to know that measures of a student's work are valid. Modifications refer to practices that change, lower, or reduce learning expectations. Modifications may change the underlying construct of an assessment.

Examples of modifications include the following:

✓ *Requiring a student to learn less material (e.g., fewer objectives, shorter units or lessons, fewer pages or problems)*

✓ *Reducing assignments and assessments so a student only needs to complete the easiest problems or items*

✓ *Using an accommodation that invalidates the intended construct*

✓ *Revising assignments or assessments to make them easier (e.g., crossing out half of the response choices on a multiple-choice test so that a student only has to pick from two options instead of four)*

✓ *Giving a student hints or clues to correct responses on assignments and tests providing modifications to* students during classroom instruction and classroom assessments may have the unintended consequence of reducing their opportunities to learn critical content. If students have not had access to critical, assessed content, they may be at risk of not meeting graduation requirements. Providing a student with a modification during a state accountability assessment may constitute a test irregularity and may result in an investigation of the school or district's testing practices. It could also affect a student's score.

IEP Team Considerations for Instructional Accommodations

To ensure that students with disabilities are engaged in standards-based instruction, every IEP team member needs to be familiar with state policies.

The team should consider:

✓ *student characteristics and needs (see tool 1 in appendix);*

✓ *instructional tasks expected of students to demonstrate proficiency in grade-level content in state standards; and*

✓ *consistency between accommodations documented in the standards-based IEP that is used for classroom instruction and those used on assessments.*

IEP team members should ask, Does the student really need any accommodation? A student may not be receiving an accommodation he or she really needs or may be receiving too many. Research indicates that more is not necessarily better, and that providing students with accommodations that are not truly needed may have a negative impact on performance. The better approach is to focus on a student's identified needs within the general education curriculum.

1.4 **Assembling and Administering Classroom Test**

Assembling the Classroom Test

The preparation of test items for use in a test is greatly facilitated if the items are properly recorded, if they are written at least several days before they are to be used, and if extra items are developed. This simplifies the task of reviewing them for defects; extra items make it easy to eliminate those items found to be defective. It also provides some latitude in fitting the final draft of the test to the test plan.

i. Recording of Test Items: As test items are being constructed, it is desirable to write each item on a separate index card. In addition to the test item, the card should contain information concerning the instructional objectives, the specific learning outcome, and the content measured by the item. A space should also be reserved on the card for item analysis information. This is usually placed on the back of the data that each time is used. Placing each item on a separate card provides the flexibility needed in preparing a test for use. As the test items are being reviewed and edited, items can be eliminated, added, or revised with very little difficulty. The same holds true when arranging the items for the test. They can be arranged and rearranged merely by sorting the cards.

ii. Review of Test Items: Whatever efforts were made during test development in one way or other, defects inadvertently creep in during construction. Defects of test items might be encountered as we concentrate on the clarity and conciseness of question, a verbal clue slips in unobserved. As we attempt to increase the difficulty of an item, we unwittingly introduce some ambiguity or as we rework an item to make the incorrect choices more plausible, the behavior called forth by the item is unintentionally modified. In short, we focus our attention so closely on some aspects of item construction that we overlook others.

However, such problem can most easily be detected by:

    a. *reviewing the items after they have been set aside for a few days, and*

    b. *asking a fellow teacher to review and criticize the items*

iii. Arranging of Items in the Test: Once the test items are written, edited, and revised for errors, the next task is to arrange them in some order. During test items arrangement, there are factors to be considered:

    a. *the types of items used*

    b. *the learning outcomes measured*

    c. *the difficulty of the items and*

    d. *the subject matter measured*

First and foremost, the items should be arranged in sections by *item type*. That is, all true-false items should be grouped together, then all matching items, then all multiple-choice items and so on. This arrangement provides for the fewest sets of directions, it is easier for the pupils since they can retain the same mental set throughout each section, and it greatly facilitates scoring.

Where two or more items types are included in a test there is also some advantage in keeping the simpler item types together and placing the more complex items types in the test as follows:

    1. *True-false*

    2. *Matching items*

    3. *Supply type (short answer and completion)*

    4. *Multiple-choice*

    5. *Interpretive exercises*

    6. *Essay questions*

Arranging the sections of the test in the above order, provides a sequence that roughly approximates the complexity of the learning outcomes measured, ranging from the simple to the complex. Items that measure similar outcomes should be placed together and then arranged in order of ascending difficulty i.e., beginning with the easiest items and proceeding gradually to the most difficult one. Beginning with easy items assist pupils by raising their motivation and confidence keep up working on the remaining items that follow. For example, the items in the multiple-choice section might be arranged in the following order:

    a. *knowledge of terms*

    b. *knowledge of specific facts*

    c. *knowledge of principles*

    d. *application of principles*

Keeping the items that measure similar outcomes together is especially helpful in determining the types of learning outcomes causing pupils the greatest difficulty.

iv. Preparation of Directions for the Test: Sometimes due to problems with clarity and incompleteness of direction, pupils would be confused. Whether written, oral or both, the directions are a vital part of the test and should include at least the following points

    A. *purpose of the test*

    B. *time allowed for completing the test*

    C. *basis for answering*

    D. *procedure for recording the answers*

The amount of detail devoted to each of these points depends mainly on the age level of the pupils, comprehensiveness of the test, complexity of the test items, and the experience of the pupils with the testing procedure used. Use of new item types and separate answer sheets, for example, requires much more detailed direction than familiar items requiring pupils merely to circle or underline the answers.

A. Purpose of the test: One element that could be included in the direction is the purpose of the test. This reflects the exam designed for diagnostic, formative or summative evaluation. It is usually indicated at the time the test is announced or at the beginning of the semester when the evaluation procedures are described as a part of the general orientation of the course. Though this can be done orally, it is preferable if the heading of the exam says something about the purpose, for instance. When a final examination is prepared for several sections the examination paper will have the following information on the top of the first page.

B. Time allotted for completing the test: It is desirable to inform to the pupils how much time they will have to answer the test items and how to distribute their time on each part or item. This enables students effectively distribute and use their time and minimize unnecessary wastage of time on few test items.

Determining the amount of time required to complete a test is not a simple matter. One has to take into consideration the type of test items used, the age and ability of the pupils, and the complexity of the learning outcomes to be measured. As a rough guide,

♦ *The average high school pupil should be able to answer two true-false items, one multiple-choice item or one short answer item per minute of testing time.*

♦ *However, elementary school pupils generally require much more time per item than high school pupils. This happen due to the reading skill level of the pupils (which is at lower level for elementary school pupils) which is an important determiner of the amount of time needed by a specific group.*

Sample of test direction

	DEBRE MARKOS UNIVERSITY	
Purpose of the test	COLLEGE OF SOCIAL SCIENCES AND HUMANITIES	**Time allotted for the**
	DEPARTMENT OF PSYCHOLOGY	
	ASSESSMENT AND EVALUATION OF LEARNING (PGDT 423)	
	FINAL EXAMINATION (50%)	**DATE: AUGUST 20, 2013**
How to record		**TIME ATTOWED: 2.00 HRS**

NAME _____ID NO._____ DEP_____ SECTION _____

General direction: This test paper contains four parts: True-False, matching, short answer, and multiple-choice type items. Give your answers on the blank space provided and for selection type encircle the answer of your choice.

Directions for responding

C. Basis for answering: The directions for each section of the test should indicate the basis for selecting or supplying answers. The test writer should clearly indicate how students should respond. Generally, writing a direction for true-false or multiple choice items can be relatively simple. For instance, a statement like "select the choice which best completes the statement or answers the questions" might be sufficient for multiple choice items.

If we are preparing direction for elementary school children, sometimes desirable to include sample test items correctly marked so that pupils can check their understanding of the basis of answering. However, essay questions frequently require special directions concerning the expected answer. For instance, if the test items demand students to select and organize ideas, this should be indicated to the students so that they have more adequate basis for responding.

D. Procedure for recording the answers: Pupils either requested to provide their answer on the test form itself or on separate answer sheets. Using the test form or a separate answer sheets for recording the answers depends on the following factors:

→ *Length of the test*

→ *Number of examinees*

→ *Age of the pupils*

If the test is short, small group of pupils and young examinees, the answers are generally recorded directly on the test paper. However, for most other situations, separate answer sheets preferred because they reduce the time needed for scoring and they make it possible to use the test papers over again.

v. Reproducing the Test

♦ *The test items should be spaced and arranged so that they can be read, answered and scored with the least amount of difficulty before they are ready for reproduction*

♦ *Most test reproduction in the schools is done on photocopying machines. As you well know the quality of such copies can vary tremendously. Regardless of how valid and reliable your test might be poor copies will make it less so.*

♦ *Regardless of the reproduction process selected, it is desirable to proofread the entire test before it is administered. Charts, graphs and other pictorial must be checked especially carefully to be certain that the reproduction has been accurate and the details are clear.*

1.4.2 Administering of the Tests: The guiding principle in administering any classroom test is that all pupils must be given a fair chance to demonstrate their achievement of the learning outcomes being measured. This means the physical and psychological environment should be suitable to their best efforts and the control of factors that might interfere with valid measurement. Physical factors that interfere with students' performance are:

- *Adequate working space*
- *Quiet room*
- *Appropriate light*
- *Ventilated room*
- *Comfortable seat and so on.*

Similarly, pupil will not perform at their best if they are overly tense and anxious during testing. Some of the excessive test anxieties caused by:

- *threatening pupils with tests, if they don't behave*
- *warning pupils to do their best " because this test is important"*
- *telling pupils they must work fast to complete the test on time*
- *threatening terrible consequences if they fail the test.*

A teacher may induce anxiety or intense fear to the pupils both by word and deed. Therefore, students should be assured and reassured that the test results are to be used to help them improve their learning and the time assigned for the test is enough to complete the test.

Other psychological factors to be considered by the teacher are:

- *time of testing, if tests are administered just before "The big game" or "the big holiday", the results may not be representative.*
- *individual pupil fatigue, the onset of illness, or worry about a particular problem may prevent maximum performance.*

When teachers observe such problems they would take some measures like arranging the time of testing in terms of such factors and permitting the postpone, when appropriate, can enhance the validity of the results. In addition to the above mentioned psychical and psychological factors that interfere with the performance of students there are some practices we need to avoid during test administration.

Things to be avoided are as follows:

i) **Do not talk unnecessarily before letting students start working:** There may be a tendency among examinees to overlook instructions unrelated to the test. For example, many pupils may fail to pay attention and remember arrangements such as make-up classes; if the announcement is delivered during test administration, even this situation, it may cause frustration for some students.

ii) **Keep interruptions during the test to minimum:** Sometimes it happens that there are some corrections to be given on the test items. If we have some corrections or related instructions, it has to be informed to the examinees at the beginning of the test. At times a pupil will ask to have an ambiguous items clarified and it will be desirable to explain the items to the entire group at the same time; such interruptions are necessary but should be kept to a minimum. All other distractors both from within and out the classroom should, of course, also be eliminated.

iii) **Avoid giving hints to pupils who ask about individual items :** During test administration, students may raise or ask you several things related to the test such as lack of clarity, difficult words, ambiguity, and difficulty of reading spelling and grammatical errors. It is advisable to respond as supervisor of the test. But, if you believe that the item needs clarification, the clarification of the item should be given to the entire class by demanding everybody's attention. Whereas in a situation where the question raised by the examinee is not real (e.g. not ambiguous) the pupil should be told to answer it as best he can; unless while you are making clarifications on individual basis, you may unknowingly provide unintended hints.

iv) **Discourage cheating:** Any activity of a student or group of students whose purpose is to give any of them higher grades than they would be likely to receive on the basis of their own achievements is considered to be cheating. Thus, the term cheating covers a wide variety of activities, such as:

1. *The sideways glance at a fellow student's answers*
2. *The preparation and use of a crip sheet , etc*

v) Activities that do not match with test administrations: Supervisors or invigilators frequently may involve in other activities such as, reading books, reading the exam itself, or other personal activities. These conditions undoubtedly give opportunities for the students to cheat. Under other conditions, however, it might be necessary to discourage cheating by special seating arrangements and careful supervisor.

1.5 Evaluating a test Or Item analysis

One can go through the following steps to conduct item analysis

1. Arrange the scored test paper in order from the highest to the lowest or in reverse order

2. From the arranged test papers, form two groups. That is upper and lower groups. If the number of students (test papers) is less than or equal to 40 divide in to two equal parts. But if the number of students is more than 40 take the upper 27% the test paper from both the upper group (high achiever) and lower group (low achiever) and leave the middle. For instance, if the number of students in the class is 70 and as a teacher you want to conduct item analysis you have to compute 27% of the upper and lower group that are going to involve in the analysis process. The remaining paper in the middle will be placed aside.

   That is, they will not be used in the analysis.

$$70 \text{ X } \frac{27}{100} = 18.9$$

   So, by using the formula we are going to use 19 test papers from the upper group and 19 test papers from the lower group, together it accounts 38 papers. The remaining 32 test papers that are found in the middle not needed for analysis.

3. For each test item tabulate the number of pupils in the upper and lower groups who selected each alternative. This tabulation can be done directly on the test paper or the test item card.

Example for item X (class size = 30)

Group	Alternatives			
	A	B	C*	D
Upper (15)	1	0	11	3
Lower (15)	2	0	7	6

(*) indicates the keyed (correct) answer

4. Add the number of students who choose the *correct alternative* from the upper and lower groups and divide the result by the total number of upper and lower group students and finally multiply the result by 100. This will give the difficulty level (index) of the item. The difficulty level is denoted by P and the formula is given as follows:

$$P = \frac{Number\ of\ studens\ who\ answered\ the\ item\ correctly\ from\ both\ groups}{Total\ number\ of\ students\ of\ bothgroups} \ X\ 100$$

$$P = \frac{R_U + R_L}{T} \ X\ 100$$

Where, R_U = Number of students who answered the item correctly from the upper group

R_L = Number of students who answered the item correctly from the lower group

T = Total number of upper and lower group students

For example, for the above data, the difficulty level for item X becomes;

$$P = \frac{R_U + R_L}{T} \ X\ 100 = \frac{11 + 7}{30} \ X\ 100 = \frac{18}{30} \ X\ 100 \ = 60\%$$

The difficulty of a test item is indicated by the percentage of pupils/students who got the item right. When a P levels are less than about 25%, the item is considered relatively difficult. When P levels are above 70%, the item is considered relatively easy. Generally, test construction experts try to build test items that have most items between P levels of 20% to 80% with an average P level of about 50%.

5. Subtract the number of students in the lower group who choose the *correct alternative* from the upper group and divide the result by half of the total number of upper and lower group students. This will give the discrimination power (index) of the item. It indicates to what extent an item discriminates the upper and lower group students. The discrimination power is denoted by D and the formula is given as follows:

$$D = \frac{\begin{array}{c}(Number\ of\ students\ who\ got\ an \\ item\ correct\ from\ the\ upper\ group)\end{array} - \begin{array}{c}(Number\ of\ students\ who\ got\ an \\ item\ correct\ from\ the\ upper\ group)\end{array}}{Half\ of\ the\ total\ number\ of\ upper\ and\ lower\ groups}$$

$$D = \frac{R_U - R_L}{1/2 T}$$

For example, for the above data, discrimination index can be computed as follows

$$D = \frac{11 - 7}{1/2 (30)} = \frac{4}{15} = 0.267$$

The interpretation of an item in relation to item discrimination power can be seen as follows:

- An item has maximum positive discriminating power if all pupils/students from the upper group got the item right and all from the lower group miss it.

$$D = \frac{15 - 0}{15} = 1$$

- An item has zero discriminating power if all pupils both from the lower and the upper group got the item right or miss it or equal number of students answered the item correctly from both groups.

For instance,

$$D = \frac{15 - 15}{15} = 0 \text{ or } \frac{0 - 0}{15} = 0$$

- The item has negative discriminating power if more pupils/students from the lower group than the upper group got the items right.

$$D = \frac{0 - 15}{15} = -1$$

Generally, an item is considered as having average discriminating power of its D index is closer to 0.5. An item with a maximum positive discriminating power would be one where all pupils in the upper group got the item right and all the pupils in the lower group got the item wrong.

In general, the following suggestion helps you to interpret the result of item discrimination.

If the value of D

is ≥ 0.40, the item is very good

$0.30 - 0.39$, the item is reasonably good but subjected to improvement

$0.20 - 0.29$, the item is moderately good but needs revision

is < 0.20, the item is poor, needs serious revision or rejected

6. Evaluating the effectiveness of distracters: Distracters effectiveness is determined by inspection or observation. No need of calculation of an index. A good or effective distracter is the one that attracts more students from the lower group than the upper group. Thus, it should discriminate the well to do students than the poor. The main purpose of destructor is to distract or attract uninformed or unprepared students from getting the correct answer.

Examination of the following item analysis data will illustrate the ease with which the effectiveness of distracters can be determined by observation.

	Alternatives			
Group	*A	B	C	D
Upper (10)	5	4	0	1
Lower (10)	3	2	0	5

(*) indicates, A is the correct answer

The following are the comments given:

- *The item has positive discrimination index-since 5 in the upper group and 3 in the lower group got the item right. However, the P value is 0.20 that is fairly low. This may happen due to the ineffectiveness of some of the distracters.*

- *Option "B" is a poor distracter because it attracts more pupils from the upper group than from the lower group. This may be due to some ambiguity incurred in the statement of the item.*

- *Option "C" is completely ineffective as a distracter because it attracted no one, on the other hand alternative "D" is functioning as intended. It attracted a large proportion of pupils from the lower group.*

- *One may improve the low discriminating power of the item by removing any ambiguity in the statement of the item and revising or replacing alternatives "B" and "C".*

Generally, item analysis data merely indicate poorly functioning items not the cause of the weakness.

Activity 2: Assume that the following item analysis data are obtained from one of the course you have taken. This multiple –choice items were followed by four alternatives and the asterisk (*) represents the choice of the correct answer. Based on the data of 5 items given below, determine P and D value and interpret the results for each. In addition, comment on the effectiveness of the distracters.

Item Number	Group	Alternatives			
		A	B	C	D
1	Upper 27 %	8*	5	7	7
	Lower 27 %	2	9	8	8
2	Upper 27 %	10	5	10*	2
	Lower 27 %	5	1	8	13
3	Upper 27 %	0	22*	3	2
	Lower 27 %	12	1	8	6
4	Upper 27 %	0	4	3	20*
	Lower 27 %	0	7	12	8
5	Upper 27 %	0	3	21	3
	Lower 27 %	4	3	18	2

(*)indicates the correct answer

Part II: Completion type

Direction: Based on item analysis data done for an item of Mid-term examination of assessment and evaluation of learning, answer those questions that follow the table. Select and complete your answer on the space provided from the list given in the alternatives section.

Item no	Group	N	Alternatives			
			A	B *	C	D
5	Upper	15	6	6	0	3
	Lower	15	6	2	0	7

(*) indicates the correct answer

a. The difficulty level of the item is --------------

b. The discriminating power of the item is ------------

c. Evaluation of the effectiveness of each destructor-------------

Part III: Supply Type

Direction: Using the item analysis data given below answer the questions that follow it accordingly

Item no	Group	N	Alternatives			
			A	B	C *	D
3	Upper	10	3	4	3	0
	Lower	10	0	3	4	3

(*) indicates the correct answer

    a. Calculate item difficulty (show the step) and interpret the result.

    b. Calculate item discriminating power and interpret the result.

    c. Comment on the distracters effectiveness.

SECTION II: Assessment of Children with Special Needs

*2. Methods of Gathering Information: I*nformation-gathering process involves:

    ✓ *observing the student's interactions with parents, teachers, and peers;*

    ✓ *interviewing the student and significant others in his or her life;*

    ✓ *examining school records and past evaluation results;*

    ✓ *evaluating developmental and medical histories;*

    ✓ *using information from checklists completed by parents, teachers, or the student;*

    ✓ *evaluating curriculum requirements and options;*

    ✓ *evaluating the student's type and rate of learning during trial teaching periods;*

    ✓ *using task analysis to identify which task components already have been mastered and in what order un-mastered skills need to be taught; and*

    ✓ *collecting ratings on teacher attitude towards students with disabilities, peer acceptance, and classroom climate.*

(Roth-Smith, 1991) Clearly, gathering information about the student using such a variety of techniques and information sources can be expected to shed considerable light upon the student's strengths and needs, the nature of his or her disability and how it affects educational performance, and what type of instructional goals and objectives should be established for the student. More detail about many of these methods of collecting information about the student will be presented below**.**

2.1. **Reviewing school Records:** School records can be a rich source of information about the student and his or her background. The number of times the student has changed schools may be of interest;

frequent *school changes* can be disruptive emotionally as well as academically and may be a factor in the problems that have resulted in the student's being referred for assessment. *Attendance* is another area to note; are there patterns in absences (e.g., during a specific part of the year, as is the case with some students whom have respiratory problems or allergies), or is there a noticeable pattern of declining attendance, which may be linked to a decline in motivation, an undiagnosed health problem, or a change within the family?

The student's past history of *grades* is usually of interest to the assessment team as well. Is the student's current performance in a particular subject typical of the student, or is the problem being observed something new? Are patterns noticeable in the student's grades? For example, many students begin the year with poor grades and then show gradual improvement as they get back into the swing of school. For others, the reverse may be true: During the early part of the year, when prior school material is being reviewed, they may do well, with declines in their grades coming as new material is introduced. Also, transition points such as beginning the fourth grade or middle school may cause students problems; the nature and purpose of reading, for example, tends to change when students enter the fourth grade, where reading to learn content becomes more central. Similarly, middle school requires students to assume more responsibility for long-term projects (Hoy & Gregg, 1994).

These shifts may bring about a noticeable decline in grades for some students.

2.2. **Interviews:** Interviewing the student in question, his or her parents, teachers, and other adults or peers can provide a great deal of useful information about the student. Ultimately, "an interview should be a conversation with a purpose" (Wallace, Larsen, & Elksnin, 1992,), with questions designed to collect information that "relates to the observed or suspected disability of the child". Preparing for the interview may involve a careful review of the student's school records or work samples, for these may help the assessment team identify patterns or areas of specific concern that can help determine who should be interviewed and some of the questions to be asked. *Parents*, for example, may be able to provide detailed information about the child's academic or medical background. It is especially important that they contribute their unique, "insider" perspective on their child's functioning, interests, motivation, difficulties, and behavior in the home or community. They may have valuable information to share about possible solutions to the problems being noted.

Teachers can provide insight into the types of situations or tasks that the child finds demanding or easy, what factors appear to contribute to the child's difficulties, and what has produced positive results (e.g., specific activities, types of rewards) (Wodrich & Joy, 1986).

The *student*, too, may have much to say to illuminate the problem. "All persons interviewed should be asked if they know of information important to the solution of the academic or behavior problem that was not covered during the interview" (Hoy & Gregg, 1994,).

Organizing interview results is essential. Hoy and Gregg (1994) suggest that the interviewer might summarize the "perceptions of each person interviewed in a way that conveys similarities and differences in viewpoints" including:

- ✓ perceptions of the primary problem and its cause,
- ✓ what attempts have been made to solve or address the problem,
- ✓ any recent changes in the problem's severity, and
- ✓ student strengths and weaknesses.

2.3. Testing and Looking at Student's Work

Test scores are also important to review. Comparing these scores to a student's current classroom performance can indicate that the student's difficulties are new ones, perhaps resulting from some environmental change that needs to be investigated more fully, or the comparison may show that the student has always found a particular skill area to be problematic. "In this situation, the current problems the student is experiencing indicate that the classroom demands have reached a point that the student requires more support to be successful"

Looking at Student's Work: often, an initial part of the assessment process includes examining a student's work, either by selecting *work samples* that can be analyzed to identify academic skills and deficits, or by conducting a *portfolio assessment,* where folders of the student's work are examined.

When collecting work samples, the teacher selects work from the areas where the student is experiencing difficulty and systematically examines them. The teacher might identify such elements as how the student was directed to do the activity (e.g., orally, in writing), how long it took the student to complete the activity, the pattern of errors (e.g., reversals when writing, etc.), and the pattern of correct answers. Analyzing the student's work in this way can yield valuable insight into the nature of his or her difficulties and suggest possible solutions.

Maintaining *portfolios* of student work has become a popular way for teachers to track student progress. By assembling in one place the body of a student's work, teachers can see how a student is progressing over time, what problems seem to be re-occurring, what concepts are being grasped or not grasped, and what skills are being developed.

The portfolio can be analyzed in much the same way as selective work samples, and can form the basis for discussions with the student or other teachers about difficulties and successes and for determining what modifications teachers might make in their instruction.

2.3. **Observation:** Observing the student and his or her environment is an important part of any assessment process. Observations in the classroom and in other settings where the student operates can provide valuable information about his or her academic, motor, communication, or social skills; behaviors that contribute to or detract from learning; and overall attitude or demeanor. Observing the student's environment(s) and his or her behavior within those environments can identify the factors that are influencing the student. For the information from observations to be useful, the team must first define the purpose for the observation and specify:

- ✓ Who will make the observation;
- ✓ Who or what will be observed;
- ✓ Where the observation will take place (observing a range of situations where the student operates is recommended);
- ✓ When the observation will take place (a number of observations at different times is also important); and
- ✓ How the observations will be recorded.

Observations are a key part of some of the assessment methods that will be discussed later in this section, including curriculum-based assessment, ecological assessment, and task analysis. There are many ways in which to record what is observed; the box below lists and briefly describes the more common observational methods.

While observations can yield useful information about the student and his or her environments, there are a number of errors that can occur during observations and distort or invalidate the information collected. One source of error may come from the observer—he or she must record accurately, systematically, and without bias. If his or her general impression of the student influences how he or she rates that student in regards to specific characteristics, the data will be misleading and inaccurate.

This can be especially true if the student comes from a background that is different from the majority culture.

In such cases, it is important that the observer have an understanding of, and a lack of bias regarding, the student's cultural or language group. Often, multiple observers are used to increase the reliability of the observational information collected. All observers should be fully trained in how to collect information

using the specific method chosen (e.g., time sampling using a checklist) and how to remain unobtrusive while observing and recording, so as not to influence the student's behavior. It is also important to observe more than once, in a number of situations or locations, and at various times, and to integrate these data with information gathered through other assessment procedures. Decisions should not be made based upon a narrow range of observational samples.

Common Observational Techniques

Anecdotal Records: The observer describes incidents or behaviors observed in a particular setting in concrete, narrative terms (as opposed to drawing inferences about feelings or motives). This type of record allows insight into cause and effect by detailing what occurred before a behavior took place, the behavior itself, and consequences or events that occurred after the behavior.

Event Recording: The observer is interested in recording specific behavioral events (such as how many times the student hits or gets out of his or her seat). A tally sheet listing the behaviors to be observed and counted is useful; when the observer sees the behavior of interest, he or she can simply make a tick mark on the sheet.

Duration Recording: This method usually requires a watch or clock, so that a precise measurement of how much time a student spends doing something of concern to the teacher or assessment team (e.g., talking to others, tapping, rocking) can be recorded.

Time-sampling Recording: With this technique observers count the number of times a behavior occurs during a specific time interval. Rather than observe for long periods of time and tally all incidences of the behavior causing concern, the observer divides the observation period into equal time units and observes and tallies behavior only during short periods of time. Based upon the time sampling, predictions can then be made about the student's total behavior.

Checklists and Rating Scales: A checklist usually requires the observer to note whether a particular characteristic is present or absent, while a rating scale typically asks the observer to note the degree, which a characteristic is present or how often a behavior occurs. There are many commercially available checklists and rating scales, but they may be developed locally as well.

2.5. Ecological Assessment: Ecological assessment basically involves directly observing and assessing the child in the many environments in which he or she routinely operates. The purpose of conducting such an assessment is to probe how the different environments influence the student and his or her school performance. Where does the student manifest difficulties? Are there places where he or she appears to function appropriately? What is expected of the student academically and behaviorally in

each type of environment? What differences exist in the environments where the student manifests the greatest and the least difficulty? What implications do these differences have for instructional planning? As Wallace, Larsen, and Elksnin (1992) remark: "An evaluation that fails to consider a student's ecology as a potential causative factor in reported academic or behavioral disorders may be ignoring the very elements that require modification before we can realistically expect changes in that student's behavior.

2.6.**Prereferral Procedures:** Many school systems recommend or require that, before an individualized evaluation of a student is conducted, his or her teacher meet with an assistance team to discuss the nature of the problem and what possible modifications to instruction or the classroom might be made. These procedures are known as *prereferral.* Prereferral procedures have arisen out of a number of research studies documenting faulty referral practices, including, among other practices, the over referral of students who come from backgrounds that are culturally or linguistically different from the majority culture, those who are hard to teach, or those who are felt to have behavioral problems.

According to Overton (1992), "the more frequent use of better prereferral intervention strategies is a step forward in the prevention of unnecessary evaluation and the possibility of misdiagnosis and over identification of special education students". This process recognizes that many variables affect learning; rather than first assuming that the difficulty lies within the student, the assistance team and the teacher will look specifically at what variables (e.g., classroom, teacher, student, or an interaction of these) might be affecting this particular student. Examining student records and work samples and conducting interviews and observations are part of the assistance team's efforts. These data gathering approaches are intended to specify the problem more precisely and to document its severity. Modifications to the teacher's approach, to the classroom, or to student activities may then, be suggested, attempted, and documented; if no progress is made within a specific amount of time, then the student is referred for an individualized evaluation. It is important for teachers to keep track of the specific modifications they attempt with a student who is having trouble learning or behaving, because these can provide valuable information to the assessment team at the point the student is referred for evaluation.

3. Major Steps in the Process of Assessment

    3.1 **Assessment Steps and Purpose**

Assessment in educational settings serves five primary purposes:

- ✓ screening and identification: to screen children and identify those who may be experiencing delays or learning problems;

- ✓ eligibility and diagnosis: to determine whether a child has a disability and is eligible for special education services, and to diagnose the specific nature of the student's problems or disability;

- ✓ IEP development and placement: to provide detailed information so that an Individualized Education Program (IEP) may be developed and appropriate decisions may be made about the child's educational placement;

- ✓ instructional planning: to develop and plan instruction appropriate to the child's special needs; and

- ✓ evaluation: to evaluate student progress.

In other words , Steps in Assessment include the following:

A. Screening	D. Program planning
B. Making a referral	E. Monitoring individual progress
C. Determining Eligibility	F. Evaluating the program

Assessment Steps and Purposes

Assessment Questions	Assessment Steps and Purposes	Approaches
Is there a possibility of a disability?	**Screening:** To determine whether students may have a disability and should be referred for further assessment	Norm referenced instruments Curriculum-based assessment Criterion-referenced assessments Observations Checklists
Who are the teachers, therapists, or other professionals who should be involved? What approaches are indicated?	**Referral:** To request further assessment after completing a referral form. To determine the professionals who should be involved and the assessment approaches that are indicated	
Does the student have a disability? What disability does the student have? Does the student meet the criteria for services? What are the strengths and weaknesses? What is the student having trouble doing? What does the student	**Eligibility:** To determine if there is a disability To determine the need for special education and related services To compare the student's performance with the performance of the peer group To determine specific strengths and weaknesses To understand why the student is having difficulty	Norm referenced instruments Curriculum-based assessments Criterion-referenced assessments Observations Error analysis Interviews Checklists Student, parent, and /or teacher conferences Performance assessments

understand?		
What types of special education and related services should be provided? What classroom modifications and adaptations should be implemented? What does the student not understand? Where should instruction begin?	**Program Planning**: To determine the locations and services to be received To assess the physical, learning, and social classroom environments To understand what the student knows and does not know To determine where instruction begin To plan the student's program To determine instructional approaches	Norm referenced instruments Curriculum-based assessments Criterion-referenced assessments Observations Error analysis Interviews Checklists Student, parent, and /or teacher conferences Performance assessments
Once instruction begins, is the student making progress? Should the instruction be modified?	**Program Monitoring:** To understand the pace of instruction To understand what the student knows prior to and after instruction To understand the strategies and concepts the student uses To monitor the student's program	Norm referenced instruments Curriculum-based assessments Criterion-referenced assessments Observations Error analysis Interviews Checklists Student, parent, and /or teacher conferences Portfolios, Oral descriptions ,etc
Has the student met the goals of the IEP? Has the instructional program been successful for the student? Has the student made progress? Has the instructional program achieved its goals?	**Program Evaluation:** To determine whether the IEP goals have been met To determine whether the goals of the program have been met To evaluate program effectiveness	Norm referenced instruments Curriculum-based assessments Criterion-referenced assessments Observations Error analysis Interviews Checklists Student, parent, and /or teacher conferences Portfolios, Oral descriptions ,etc

Preparing To Administer Assessment Instruments
Before the Testing Begins
- ✓ *Understand the purpose of the assessment as stated in the manual*
- ✓ *Read the test manual thoroughly*
- ✓ *Carefully review the test items*
- ✓ *Know the administrative procedures*
- ✓ *Organize the necessary materials and check to see that none are missing*
- ✓ *Reexamine the scoring procedures to verify that answers can be recorded correctly.*

When the student arrives for assessment
- ✓ *Establish and maintain rapport that convey a sense of confidence about the student's performance and avoid statements such as 'This is going to be a difficult test.'*
- ✓ *Be aware of changes in lighting or noise level once the testing begins.*
- ✓ *Maintain neutrality during the testing.*
- ✓ *Carefully record student responses in the appropriate spaces on the test form.*
- ✓ *Be sensitive to the needs of the student.*

After the Testing is Completed

- ✓ *Thank the student for participating*
- ✓ *Finish recording additional information; note any observation of the student*
- ✓ *Compute scores*
- ✓ *Interpret results*
- ✓ *Write reports*

3.2 How students are Identified for Assessment

There are at least two ways in which a student may be identified for assessment.

A. **The school (teachers and school personnel**) suspects the presence of a learning or behavior problem and asks the student's parents for permission to evaluate the student individually. Schools routinely give tests to all students in a particular grade; when a student scores too far below his or her peers, this alerts the school to a potential problem. Alternatively, the student's classroom teacher may identify that a problem exists—perhaps the student's work is below expectations for his or her grade or age, or the student's behavior is disrupting learning—and so the teacher refers the student for assessment.

B. **The student's parents may also call or write** to the school or to the director of special education and request that their child be evaluated. They may feel that the child is not progressing as he or she should be, or notice particular problems in how the child learns. If the school suspects that the child, indeed, may have a disability, then the school must conduct an assessment.

After assessment, the following children with various special needs, who are eligible for special needs education services, will be identified.

- ❖ **Autism:** a developmental disability significantly affecting verbal and nonverbal communication and social interaction, generally evident before age 3;
- ❖ **Deafness:** a hearing impairment that is so severe that the child is impaired in processing linguistic information, with or without amplification;
- ❖ **Deaf-blindness:** simultaneous hearing and visual impairments;
- ❖ **Hearing impairment:** an impairment in hearing, whether permanent or fluctuating;
- ❖ **Intellectual disability:** significantly subaverage general intellectual functioning existing concurrently with deficits in adaptive behavior;

- ❖ **Multiple disabilities:** the manifestation of two or more disabilities (such as Intellectual disability-blindness), the combination of which requires special accommodation for maximal learning;
- ❖ **Orthopedic impairment:** physical disabilities, including congenital impairments, impairments caused by disease, and impairments from other causes;
- ❖ **Other health impairment:** having limited strength, vitality, or alertness due to chronic or acute health problems;
- ❖ **Serious emotional disturbance:** a disability where a child of typical intelligence has difficulty, over time and to a marked degree, building satisfactory interpersonal relationships; responds inappropriately behaviorally or emotionally under normal circumstances; demonstrates a pervasive mood of unhappiness; or has a tendency to develop physical symptoms or fears;
- ❖ **Specific learning disability:** a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations;
- ❖ **Speech or language impairment:** a communication disorder such as stuttering, impaired articulation, a language impairment, or a voice impairment;
- ❖ **Traumatic brain injury:** an acquired injury to the brain caused by an external physical force, resulting in total or partial functional disability or psychosocial impairment, or both;
- ❖ **Visual impairment:** a visual difficulty (including blindness) that, even with correction, adversely affects a child's educational performance.

 3.3 Primary Areas of Assessment: In this section, we will look in detail at the primary areas in which students are assessed, which are: intelligence, language, perceptual abilities, academic achievement, behavior, and emotional/ social development. When the disability is related to a medically related condition (e.g., sensory deficit, orthopedic impairment, arthritis), assessment information from physicians or other medical practitioners needs to be included as well. More than one assessment technique should be used in any given area, and the assessment team should clearly understand that each area encompasses more than one ability.

In this part, we will look at what skills are involved in these traditional areas of assessment (e.g., intelligence, language, and so on) and how schools may collect information about how a student performs in each area. While standardized testing is often the default means of gathering information

about a student, it is highly recommended that other methods be used as well, including interviews, observations, and methodologies

such as ecological or dynamic assessment.

3.3.1 Intelligence: While a person's intelligence is typically measured by an intelligence test, there is considerable controversy over what, precisely, is meant by the term "intelligence." Binet, who was largely responsible for the development of the first intelligence test, viewed intelligence as a collection of faculties, including judgment, practical sense, initiative, and the ability to adapt to circumstances (Wallace, Larsen, & Elksnin, 1992). Thurman, in contrast, developed a multifactor theory of intelligence, which included such mental abilities as verbal, number, perceptual speed, reasoning, memory, word fluency, and spatial visualization. Wechsler, on the other hand, believed that intelligence was the ability of the person "to act purposefully, to think rationally, and to deal effectively with his environment" (Wechsler, 1958, p. 7, as cited in Wallace, Larsen, & Elksnin, 1992).It is important to know that different intelligence tests are based upon different definitions of what constitutes intelligence. As a result, different tests may measure different skills and abilities. It is critical, therefore, that administrators of such tests "be completely aware of an author's definition of intelligence when selecting and interpreting an intelligence test" and "to view the scores as highly tentative estimates of learning ability that must be verified by other evidence" (Wallace, Larsen, & Elksnin, 1992,). The theory underlying intelligence tests (e.g., how does one define intelligence or develop tests of intelligence?) is not the only controversy surrounding their use. How fairly they assess certain populations (e.g., minority children, persons with limited experience, children with severe language deficits), and whether or not such tests are reliable and valid (Elliott, 1987) are also areas of hot debate. In the past, intelligence measures have been misused, particularly with African American children, Native Americans, and non-English speaking children, who, based upon their scores, were placed in classes for those with mental retardation or with learning disabilities. However, given the many court cases involving standardized intelligence testing as a means of assessing minority children provoked the strength and volume of advocates' protests. Evaluators are now becoming more sensitive to issues of test bias. That is the importance of testing in a child's native language, the need for specialized training when administering and interpreting standardized tests, and the importance of combining any test scores with information gathered in other ways. Issues related to the definition of "intelligence" and the "fairness" of using measures of intelligence also become less concerning if one knows the purpose for which the test is being used.

Intelligence tests are most helpful (and probably most appropriate) when they are used to determine specific skills, abilities, and knowledge that the child either has or does not have and when such information is combined with other evaluation data and then directly applied to school programming. There are a number of skills that an intelligence test appears to measure social judgment, level of thinking, language skill, perceptual organization, processing speed, and spatial abilities. Questions that attempt to measure social judgment and common sense, numerical reasoning, concrete and abstract thinking, the ability to recognize similarities and differences between objects or concepts, and vocabulary and language skill (e.g., the ease with which a person can find words in memory) appear very dependent on experience, training, and intact verbal abilities. Perceptual organization, processing speed, and spatial abilities seem less dependent on experience and verbal skill.

Intelligence tests can also yield valuable information about a student's ability to process information. In order to learn, every person must take in, make sense of, store, and retrieve information from memory in an efficient and accurate way. Each of us can process certain kinds of information more easily than other kinds. The artist sees and reproduces accurate depictions of the world, while others struggle to produce stick figures.

The musician creates beautiful sounds from a mixture of separate tones. The writer crafts words to create a mood. Others of us do none of these things well. In school, children need certain skills to function effectively. They must be able to listen attentively so that other movements, sounds, or sights do not distract them.

They must be able to understand the words spoken to them. This often requires children to hold multiple pieces of information in memory (e.g., page number, questions to answer) and to act upon them. These children must be able to find the words they need to express themselves, and ultimately, write these words on the paper correctly/wrongly. This involves another whole series of processing skills holding a writing implement, coordinating visual and motor actions, holding information in memory until it can be transferred to paper, transforming sounds into written symbols, and understanding syntax, punctuation, and capitalization rules. They also must be able to interpret the nonverbal messages of others, such as a frown, a smile, a shake of the head. Equally important, they must do all of these things quickly and accurately and often in a setting with many distractions.

A thorough interpretation of an intelligence test can yield information about how effectively a child processes and retrieves information.

Most individually administered intelligence tests can determine, at least to some degree, a child's ability to attend process information quickly, distinguish relevant from less relevant details, put events in sequence, and retrieve words from memory. Kamphaus (1993) summarizes a number of research findings related to the use of intelligence tests:

1. Intelligence test scores are more stable for school-aged children than for preschoolers and more stable among individuals with disabilities than those without disabilities;

2. Intelligence test scores can change from childhood to adulthood;

3. It is likely that environmental factors, socioeconomic status, values, family structure, and genetic factors all play a role in determining intelligence test scores;

4. Factors such as low birth weight, malnutrition, anoxia (lack of oxygen), and fetal alcohol exposure have a negative impact on intelligence test scores; and

5. Intelligence and academic achievement appear to be highly related. This last finding supports the notion that intelligence and achievement tests may not be so different from each other and that "intelligence tests may be interpreted as specialized types of achievement measures" (Kamphaus, 1993). This is consistent with the suggestion that intelligence tests may be best used to determine specific skills, abilities, and knowledge.

3.3.2 Language: Language provides the foundation upon which communication, problem solving, and expanding, integrating, analyzing, and synthesizing knowledge take place. Deficits in language, therefore, can have a profound impact on the ability of an individual to learn and function competently and confidently as he or she interacts in the world. Language is complex and involves multiple domains—nonverbal language, oral language (i.e., listening and speaking), written language (i.e., reading and writing), pragmatic language (e.g., using language for a specific purpose such as asking for help), phonology, and audiology. How quickly a person can access words or ideas in memory further influences his or her use of language.

A child who must struggle to find an appropriate term is at a great disadvantage in a learning and social environment. As he or she grapples to retrieve a word, others have moved on. The student may miss critical pieces of knowledge, connect incorrect bits of information in memory, and have an ineffective means of showing others all that he or she knows. Such problems can result in lowered levels of achievement and in feelings of confusion, helplessness, and frustration. It is clear how important language processing can be to a child's successful adaptation to the school environment and, therefore, it is an important area to be considered in the assessment process.

Speech and language pathologists are specially trained professionals who, working with school psychologists and classroom teachers, are frequently the primary individuals gathering data related to a child's language functioning. Bloom and Lahey (1978) divide language processes into three general categories: form, content, and use. Phonology, morphology, and syntax are all considered to be components of *form.* The first of these processes, ***phonology***, refers to the knowledge a person has of the sounds in the language. While the number of sounds that exist are limited, a nearly endless number of words can be constructed from these sounds. Awareness of the basic sound units of language appears important to a child's ability to quickly and accurately locate words in memory when speaking, comprehend oral sentences, and learn to read (Liberman & Shankweiler, 1987). It is important to note that the ability to blend or separate sounds (i.e., phonological processing ability) is often overlooked in the assessment process. This may be an unfortunate oversight, given its apparent importance to the reading process. ***Morphology***, the second form element, refers to the smallest meaningful unit of language. Morphology involves the stringing together of sounds (phonemes) and includes such structures as prefixes, suffixes, word endings that describe number (e.g., dog vs. dogs), and tense (e.g., walk vs. walked). *Syntax* refers to the rules used in combining words to make a sentence. As with the sounds of language (phonology), the rules of language are finite. The acquisition of syntax is also developmental. While syntax determines the rules that guide how sentences are put together, such knowledge alone is not sufficient for constructing sentences. The meaning of words constrains what words may or may not be used together. For example, the sentence "I saw the house flying over the orchard" would make little sense, although it is syntactically correct. It is this aspect of language, the importance of meaning, that Bloom and Lahey (1978) refer to as ***content***. Content involves knowledge of vocabulary, the relationships between words, and "time-and-event" relationships (Swanson & Watson, 1989). The child must also be able to associate words with the correct environmental experience. It is generally expected that a child understands the meaning of more words than he or she can express at any point in time. As Swanson and Watson (1989) point out, when an individual appears able to express more information than he or she is able to receive and comprehend, it may suggest that he or she has difficulty in auditory input and processing. ***Use***, the final component in the Bloom and Lahey model, refers to "the pragmatic functions of language in varying contexts" (Swanson & Watson, 1989,. It views the child as an active "communicator" whose words and sentences are intentionally selected in relation to the effect the speaker wishes to have on a listener.

The speaker needs to be able to:

(a) change what is said in some way when it is apparent that he or she is not being understood,

(b) vary language use when talking with different groups (e.g., peers or adults), and

(c) use language in a variety of functional ways (e.g., to begin or end a conversation).

Thus, use (or pragmatics, as it is sometimes called) is a vital area to assess in language; to ignore how a student uses language is to ignore a basic element of language—that we communicate in a context, for a particular purpose or reason (Heward & Orlansky, 1992).

Assessing a Child's Language

Abilities. The IDEA's regulations provide a definition of speech-language impairment as "a communication disorder such as stuttering, impaired articulation, a language impairment, or a voice impairment that adversely affects" a child's educational performance [34 CFR 300.7(b)(11)]. In more specific terms, a child with a speech disorder may have difficulty in *producing* sounds properly, speaking in a normal flow or rhythm, or using his or her voice in an effective way. A child with a *language* disorder would have problems using or understanding the rules, sounds, or symbols that we use to communicate with each other. This relates to language form, content, and/or use, as discussed above (Heward & Orlansky,1992). A child with a speech impairment, a language impairment, or both, would be eligible for services under the IDEA.

There are many standardized measures of speech and language ability. Some "provide a *comprehensive* view of all language functioning," while others "measure *specific* components of linguistic performance (for example, phonology, linguistic structure, or semantics)" (Wallace, Larsen, & Elksnin, 1992,).

The range of tests and what they measure may be identified through consulting resource books on speech/ language assessment or more general test references such as *Tests* (Sweetland & Keyser, 1991), or by contacting organizations such as the American Speech-Language-Hearing Association (ASHA) (listed under "Organizations" at the end of this issue).

It is important to realize, however, that "standardized diagnostic tests are generally insensitive to the subtleties of ongoing functional communication" (Swanson & Watson, 1989,). Therefore, in addition to or in place of standardized tests, a typical speech/language evaluation should include obtaining a language sample that seeks to capture how the student performs in an actual communication situation. Language samples can be obtained through checklists or observational recording systems, or through informally conversing with the student.

Great care must be taken to ensure that assessment of students is culture-free and dialect-sensitive, as many children will speak nonstandard English or another language entirely. [The issue of cultural bias in language assessment is considered in *Topics in Language Disorders* (Terrell, 1983)]. Obtaining such a language sample from the student is often the responsibility of the speech-language pathologist.

Through interviews, observations, and teaching, teachers can also gather valuable information about a student's language use. By engaging in what is known as diagnostic teaching, the teacher can become an invaluable participant in the ongoing assessment and remediation of a child's language deficiencies. It is important, however, for teachers to be thoroughly familiar with the developmental milestones of normal language functioning. Obtaining a case history of the child (in most cases, from the parents) can also be valuable in the initial stages of assessment.

Knowing in detail how the child's language has developed can yield information relevant to the problem. In addition, gaining an understanding of the early stages of the disorder, any physical or emotional condition that may have been or be involved, whether the disorder occurs in other settings and, if so, how it manifests itself, and any insights the parents may have into how best to assess and work with their child (Wallace, Larsen, & Elksnin, 1992,). It is also important to realize that the ability to receive and understand language, and to use language verbally, is in part dependent upon how well the body performs physically. Before embarking upon an extensive (and expensive) battery of tests, examiners should ensure that any apparent speech or language impairment is not actually the result of a hearing impairment which, in effect, prevents the child from hearing words clearly and learning to use or understand them. Similarly, many children with physical disabilities may not be able to speak clearly enough to be understood but, when provided with assistive technology (e.g., speech synthesizers, computers), may show themselves to be competent users of language.

 i. Language development can be assessed in different ways

The expected ability of the child in language development include:

Word Reading Assess pre-reading (phonological awareness) and decoding skills

- ✓ Name the letters of the alphabet
- ✓ Identify and generate rhyming words
- ✓ Identify the beginning and ending sounds of words
- ✓ Match sounds with letters and letter blends
- ✓ Read aloud from a graded word list

In the classroom, Reading Comprehension can be assessed during reading instruction

✓ Match a written word with its representative picture

✓ Read passages and answer content questions

✓ Read short sentences aloud, and respond to comprehension questions

Pseudoword Decoding Assess the ability to apply phonetic decoding skills

✓ Read aloud a list of nonsense words designed to mimic the phonetic structure of words in the English language.

Reading Recognition. This subtest contains 100 items, ranging in difficulty from preschool level through high school level. Items assess skill development in matching letters, naming capital and lowercase letters, and recognizing words in isolation.

Reading Comprehension. This subtest contains 81 multiple-choice items assessing skill development in understanding what is read. After reading a sentence, the student must indicate comprehension by choosing the correct picture out of a group of four.

Assessment of Reading Comprehension: Diagnostic tests assess five different types of reading comprehension:

1. *Literal comprehension entails understanding the information that is explicit in the reading material.*

2. *Inferential comprehension means interpreting, synthesizing, or extending the information that is explicit in the reading material.*

3. *Critical comprehension requires analyzing, evaluating, and making judgments about the material read.*

4. *Affective comprehension involves a reader's personal and emotional responses to the reading material.*

5. *Lexical comprehension means knowing the meaning of key vocabulary words*

Spelling. This subtest consists of 100 items sampling behaviors from kindergarten level through high school level. Initial items assess the student's ability to distinguish a printed letter of the alphabet from pictured objects and to associate letter symbols with speech sounds. More difficult items assess the student's ability to identify, from a response bank of four words, the correct spelling of a word read aloud by the examiner.

General Information: This subtest consists of 100 questions presented orally, which the student must answer orally. Items assess the extent to which the student has learned facts in social studies, science, sports, and the fi ne arts.

Written Expression: This subtest assesses written language skills at two levels. Level I, appropriate for students in kindergarten and first grade, is a measure of prewriting skills, such as skill in copying and writing letters, words, and sentences from dictation. At Level II, the student writes a story in response to a picture prompt.

3.3.2 Perceptual Abilities

Perceptual abilities determine how individuals perceive information and how they respond. These abilities can be subdivided into at least four general areas: visual-perceptual, auditory-perceptual, perceptual-motor skill, and attention. Assessing a student in these areas is intended to determine strengths and weaknesses in information and sensory processing and can help the assessment team gain an understanding of how the child learns best. The idea of "perceptual deficits" has long been linked to learning disabilities. It is important to realize that research results in this area have been mixed and controversial, and offer only small support for including evaluation of perceptual abilities in any assessment battery or approach (Overton, 1992). Linguistic issues, rather than perceptual abilities, may more often explain learning deficits.

Nevertheless, since assessing perceptual abilities continues to be part of the evaluation process at present, we will briefly discuss them below.

Visual-Perceptual Ability: Visual perception includes the ability to discriminate between two or more visual stimuli, locate a particular figure within a larger scene, and understand position in space. Perceptual skills include detecting specific colors, shapes, and sizes. In reading, it requires the ability to detect the visual features of a letter or word so that the 26 letters of the alphabet can be distinguished from each other. The student must also discriminate between ten written digits.

Auditory-Perceptual Ability: Auditory perception includes the ability to detect certain auditory features such as changes in volume, discrimination of vowel or consonant sounds, and non-phonemic sound discrimination (e.g., the sound of a bell from the sound of a buzzer). In a school setting, then, the student would need the ability to discriminate between different sounds, identify spoken words that are the same or different, and hear sounds in order.

Perceptual-Motor Ability: Most assessments include one or more measures of perceptual-motor ability. It has been an assumption of many educators that perceptual-motor or visual-motor problems are often associated with learning problems and, therefore, should be included in most assessment batteries (Salvia & Ysseldyke, 1991). Historically, tests of perceptual motor skill have been second only to intelligence tests in terms of use in the assessment of school-aged children.

Tests of perceptual-motor skill or perceptual motor integration most often ask students to copy geometric designs that are placed in front of them. This requires the child to see the design, attend to and remember the relevant features, and then carry out the motor actions necessary to reproduce the design on paper.

Attention. The ability to focus on a given activity for extended periods is important if a student is to take in information or complete the day-to-day tasks in school. Keogh and Margolis (1976) have suggested three phases of attention: the ability to (a) come to attention; (b) focus attention; and (c) maintain attention. The issue of "selective" attention must also be considered here. Students must be able to attend, *and* they must be able to sustain attention on the most relevant stimuli. For example, a student must be able to attend to the teacher's words rather than to his or her clothing. Difficulties in any of the three phases of attention can interfere with a student's ability to learn or share what he or she knows in a consistent fashion.

While the ability to attend effectively is seldom assessed through a formalized instrument, information related to attention can be gathered through classroom observations and observations of test behaviors.

Assessing Perceptual Ability. As was mentioned above, assessing perceptual abilities is not without its controversies. There are certainly a number of issues that need to be considered when addressing this area.

The first issue relates to the importance of ensuring that a student's apparent perceptual difficulties are not actually the result of a lack of visual or auditory acuity (as opposed to a difficulty with processing stimuli). Before beginning an assessment of perceptual ability, then, the student's eyesight and hearing should be tested (Overton, 1992; Swanson & Watson, 1989). This can be part of the assessment process, with the school referring the student to the appropriate facilities for such screenings.

The second issue is related to the relevance of such measures to the goals of assessment. There has been little to suggest that direct training in perceptual skills improves academic performance (Salvia & Ysseldyke, 1991; Vellutino, 1979). If there is little applicability, then it seems reasonable to question whether formal tests of perceptual skill are necessary as part of the assessment battery.

The third issue is related to the validity and reliability of the perceptual test measures. There is some suggestion that tests purported to measure perceptual abilities may actually measure other factors such as language or verbal memory skill (Vellutino, 1979). Information gained from tests thought to measure perceptual processing may actually result in incorrect explanations for learning problems.

This may lead those working with the child towards strategies that are not useful (perceptual training such as copying designs) and away from ones that may be helpful, such as training in phonological processing. There are also concerns that many of the instruments currently available do not meet acceptable standards of reliability and validity (Swanson & Watson, 1989, p. 217), making their use of questionable value.

3.3.4.Academic achievement: Academic achievement refers to how well the child is performing in core skill areas such as reading, mathematics, and writing. Assessment batteries typically include an individual measure of academic achievement, although it is important to realize that standardized achievement tests may be inappropriate for use with immigrant or minority group children. Information about the child's placement (i.e., below, at, or above) in his or her peer group and knowledge about the specific skills the child possesses are important both for the planning and evaluation of instruction.

Reading. Reading is an extension of the language process. It provides a way for individuals to exchange information. Reading also represents the means by which much of the information presented in school is learned and is the academic area most often implicated in school failure. Reading, like language, is an extremely complex process, a process that is, for many, so natural or fluent that many of the subskills are not recognized or identified as a part of the process. Identifying these subskills is important, however, if an adequate assessment in this area is to occur.

Pre-reading skills include:

- ✓ *general language competence;*
- ✓ *understanding that reading is a means of exchanging ideas (e.g., the ability to "read" pictures);*
- ✓ *the ability to complete rhymes and identify words that do not rhyme;*
- ✓ *the ability to distinguish between verbal and nonverbal sounds, recognize when words are the same or different, and segment and blend language sounds; and*
- ✓ *the ability to store and retrieve sounds one has heard.*

Having opportunities for abundant language experiences, while not a skill, is also important to the development of pre-reading and later reading ability. Reading skills can be divided into two general categories: word recognition and comprehension. A number of skills are used when attempting to identify, pronounce, or retrieve a word. Four types of analyses can be used by the child: visual analysis (i.e., the use of visual features), contextual analysis (i.e., using the surrounding words for clues about a given word), phonological analysis (i.e., using information about the sounds in the word), and structural

analysis (i.e., recognizing and giving meaning to specific word parts such as prefixes, suffixes, or syllables). Phonological analysis appears particularly important as children attempt to gain reading skill. It allows the child to decode (i.e., read) a word he or she has never seen before, either in isolation or in context.

This is not possible with visual, contextual, or structural analysis alone. The ability, then, to engage in phonemic analysis is important to becoming a proficient reader and, therefore, is an area that should be considered in any assessment of any child who is struggling with reading. Gaining meaning from text (comprehension) is the most common goal of reading. The general approach of the reader (active or passive), use of prior knowledge, and contextual analysis are all skills that appear related to comprehension. The ability to grasp literal information and to predict, interpret, critically analyze, or create new ideas in response to a paragraph are examples of the use of context at the comprehension level. Listening comprehension also appears to be related to reading comprehension, particularly at the higher reading skill levels (Stanovich, 1982).

The assessment of reading, then, needs to address the ability of the child to recognize individual words and to comprehend text. Assessment instruments should be selected that assure that test content and test tasks are as similar as possible to school reading tasks. Both formal and informal assessment may be useful here.

Informal measures may include asking the student to:

✓ *read aloud, which permits the teacher to identify errors in decoding and to determine the student's fluency and accuracy when reading;*

✓ *answer questions after reading, to determine the student's ability to understand the main idea of the story, capture its details, or place events in sequence;*

✓ *paraphrase or re-tell the story in his or her own words;*

✓ *fill in missing words in a passage he or she has not read;*

✓ *identify which sentence out of several means the same thing as a sentence supplied by the teacher; and*

✓ *provide synonyms of selected words.*

Mathematics. Another critical area of school achievement is that of mathematics. The terms "mathematics" and "arithmetic" are often used interchangeably but actually mean different things. *Mathematics* refers to the study of numbers and their relationships to time, space, volume, and geometry, while *arithmetic* refers to the operations or computations performed.

These subskills related to mathematics include:

- ✓ problem-solving,
- ✓ the ability to perform mathematics in practical situations,
- ✓ performance of appropriate computational skills,
- ✓ use of mathematics to predict,
- ✓ understanding and use of concepts related to measurement,
- ✓ interpretation and construction of charts or graphs,
- ✓ ability to estimate,
- ✓ understanding and application of geometric concepts,
- ✓ ability to recognize the reasonableness of results, and
- ✓ computer knowledge.

For a student to learn and act on knowledge of mathematics, he or she must understand terms regarding amount or direction (i.e., language-based knowledge), understand that numbers stand for a quantity, hold multiple pieces of mathematical information in memory and perform mathematical operations (e.g., add, multiply) on them, and know that numbers can be manipulated in meaningful ways.

The assessment of mathematics should measure a student's ability in both calculation and reasoning (application). Like reading, an evaluation of mathematical understanding and performance should also be structured so that it closely matches the demands made on the child in the actual classroom situation. Assessment might begin by analyzing actual samples of the student's work and identifying specific errors and any apparent pattern to those errors. Curriculum-based assessment techniques are also useful, and can be combined with task analysis and error analysis to identify where, specifically, the student is having problems. Interviewing can be useful as well, and may include" asking the student to solve a problem and explain the steps used in the process" (Overton, 1992, p. 257). Such an approach can be invaluable in providing insight into a student's mathematical reasoning. Conducting several such interviews is important, however, to avoid drawing hasty conclusions about the nature of a student's difficulties. Observations can also provide productive information to the assessment team and should focus on student behavior during—and his or her approach to written assignments, working at the chalkboard, and classroom discussions.

Written Language. Written language is a complex form of communicating that consists of three general areas: spelling, handwriting, and written expression or composition. Like reading, writing tasks are an important part of the school curriculum and are often utilized in evaluating a student's understanding of

a given concept. Written language is directly tied to reading, listening, and speaking, and skills in all of these areas overlap.

Spelling has often been considered a difficult task (Henderson, 1985). In English, the difficulty arises because there is no one-to-one correspondence between letters and their representative sounds.

This can cause problems for the reader and may cause even greater problems for the speller. In spelling there are even fewer cues to aid in recreating a spoken word in print. As Lerner (1988) explains: Several clues aid the reader in recognizing a word in print: context, phonics, structural analysis, and configuration. There is no opportunity, however, to draw on peripheral clues in reproducing a word. Both language and reading experience appear to be important to the development of connections between letters and their sounds. Thus, knowledge of spelling patterns, analysis of word parts, and knowledge of syllable rules all need to be measured.

It is important to consider that any approach that does not require a child to independently reconstruct a word (e.g., one that simply asks a child to select a misspelled word from among a group of words) does little to give information about the child's ability to recreate accurate spelling in a sentence he or she is writing. Assessment of spelling is particularly well given to informal approaches such as curriculum-based measurement or interviews. A number of standardized, commercially available spelling tests are available as well.

Handwriting refers to the actual motor activity that is involved in writing. Most students are taught manuscript (printing) initially and then move to cursive writing. There are those who advocate that only manuscript or only cursive should be taught (Reid & Hresko, 1981). In truth, problems may appear among students in either system. Wiederhold et al. (1978) have suggested a number of areas which may be assessed related to both manuscript and cursive writing. The assessment of manuscript includes evaluating the position of the hand and paper, size of letters and the proportion of letters to each other, quality of the actual pencil lines, the amount and regularity of the slant of the letters, letter formation and alignment, letter or word spacing, and speed of production. Cursive writing can be considered according to many of the same qualities but should also include an evaluation of the way in which letters are connected.

Composition refers to the more creative parts of written expression.

Alley and Deshler (1979) suggest three general areas that need to be addressed in any assessment of written expression:

  ✓ *the student's attitude toward writing;*

✓ *ability of the individual to express content (e.g., skill in describing or reporting events, or in expressing views or feelings); and*

✓ *the student's ability to "craft" a paragraph (e.g., the student's ability to organize, sequence, choose effective words, use punctuation and capitalization, or take notes).*

Both formal and informal measures of assessment of written expression are available and should be considered in a thorough evaluation. Analyzing work samples produced by the student can be particularly useful, as can interviewing the student regarding his or her perceptions of the writing process.

3.3.5 Behavioral, Emotional and Social Development: Behavior—how a student conducts himself or herself in school—is often a key factor in educational performance. Certainly, behavior that is off-target academically or socially—inattention, being out of seat, talking too much, hitting or biting, skipping school—can detract from learning. When a student's behavior appears to be interfering with school performance and relationships with others, or when that behavior is maladaptive, bizarre, or dangerous, it becomes important to assess the student's behavior (when the behavior occurs, how often, and for what reasons) as well as his or her emotional and social development. Wallace, Larsen, and Elksnin (1992) "stress the need to take an *ecological* perspective when assessing a student's nonacademic behaviors in order to obtain a complete picture and examine the relationship between the behavior and the environment". Negative or inappropriate behaviors may occur for different reasons. One child may be disruptive in class because of attention deficit disorder. A second child may exhibit similar behaviors due to a mental illness, while another's inappropriate behavior may be linked to environmental factors such as his or her parents' recent divorce. Still another child may be disruptive only in one or two classes, for reasons associated with the way instruction is organized (e.g., a predominance of small group, large group, or self-paced activities) or something in that environment which the student finds disturbing.

Thus, identifying *why* a child is exhibiting certain behaviors is an important part of the assessment process. The reasons why, if they can be determined, will influence whether or not the child is determined eligible for special education services and, if so, will certainly affect the nature of decisions made regarding educational and other interventions.

Assessing Problem Behavior. For children exhibiting signs of emotional, social, or behavioral problems, the assessment team will generally conduct a *behavioral assessment.* The goal of behavioral assessment is to gain an increased understanding of how environmental factors may be influencing the child's

behavior. This includes identifying (a) what expectations and rules are established by significant others in the settings where the problem behavior occurs, and (b) what "specific variables in a particular situation . . . may be maintaining problem behaviors" (Berdine & Meyer, 1987,). This knowledge will then be used directly in designing intervention strategies. "Behavioral assessment depends on keen observation and precise measurement" (Swanson & Watson, 1989,). Assessment is tied to observing a specific situation (e.g., how the child responds during lunch or reading) at a particular point in time. It is important that a behavioral assessment involve multiple measures and take place in various settings (e.g., the classroom, school playground, chorus, home) and at different times during the day (e.g., morning, afternoon, and night). The ability to observe and record behavior, select the most appropriate places to observe the child, and find efficient and clear means of interpreting results are all critical in behavioral assessment. Collectively, the observations should provide information which:

(a) *pinpoints and quantifies the nature of the behavior problem (including what variables in the environment are contributing to or maintaining the behavior);*

(b) allows eligibility and placement decisions to be made;

(c) illuminates what type of instruction or intervention is needed; and

(d) provides baseline information against which progress can be measured once intervention begins.

Interviews are also a useful means of gathering information about a child's behavior. Parents and significant others may be able to offer insight into the nature and history of the child's difficulties. The child may also be an excellent source of information. Of primary interest here is determining the child's "awareness of the problem behaviors and their controlling variables, degree of motivation to change, and skill at behavioral self-control"

Assessing Adaptive Behavior. Other aspects of behavior may be important to assess as well. ***Adaptive behavior*** is a frequent focus of assessment, and is a required area of assessment when a classification of mental retardation is being considered for a student. Adaptive behavior refers to the effectiveness or degree with which individuals meet the standards of personal independence and social responsibility expected for age and cultural groups.

When assessing a person's adaptive behavior, examiners may investigate his or her strengths and weaknesses in a variety of different skill areas, such as: communication, self-care, home living, social skills, community use, self-direction, health and safety, functional academics, leisure, and work. According to the American Association on Mental Retardation (1992), these are the skills with which individuals most often require assistance or some specialized support.

The IDEA specifies "deficits in adaptive behavior" as one of the two characteristics necessary for a student to be classified as having mental retardation (the other characteristic being "significantly subaverage general intellectual functioning. Measuring a student's adaptive behavior, however, should not be limited to only those students suspected of having mental retardation; this type of assessment has much to offer the decision-making associated with students with other disabilities as well, particularly in regards to IEP development and instructional and transition planning.

Many commercially-developed adaptive behavior instruments exist to help educators evaluate a student's adaptive skills. Using these instruments typically does not require the student to be involved directly; rather, examiners record information collected from a third person who is familiar with the student (e.g., parent, teacher, direct service provider) and who can report what types of adaptive skills the student has mastered and which he or she has not. Unfortunately, there is some concern that many of the available adaptive behavior scales do not meet the technical requirements of good instrumentation [for example, reliability and validity may not reported by the publisher (Berdine & Meyer, 1987; AAMR, 1992)] and that there may be bias inherent in assessing the behavior of children who are culturally or linguistically different from the majority culture. Therefore, care must be taken with the selection of the adaptive behavior scale to be used. It is also a good idea to use other methods to collect information about the student's skills, such as direct observation and interviewing the student. For minority students, it is imperative to develop an understanding of what types of behavior are considered adaptive (and, thus, appropriate) in the minority culture, before making judgments about the particular functioning of a student.

Assessing Emotional and Social Development. No child lives in a vacuum. His or her relative freedom from internal and external stressors, ability to interact with others comfortably, and ability to respond consistently and positively in the learning environment all are important for the child to benefit maximally from school experiences.

In assessing a child's emotional and social adjustment, questions need to be answered related to the child's intrapersonal and interpersonal experience. Assessment of the child's *intrapersonal* world involves knowledge about how the child views him or herself, how the child responds emotionally, how much conflict or anxiety he or she is currently experiencing, the degree to which the individual believes that personal behaviors can actually make a difference in his or her own life, his or her tolerance for frustration, and general activity level.

Interpersonal characteristics are related to how the individual views the world and other people. Such characteristics are developed in response to the child's experiences within the environment. If the child sees the world as a hostile place and views people as untrustworthy, negative interactive patterns and behaviors may emerge. The development of the child's intrapersonal experience and interpersonal behaviors is, at least in part, related to the way basic physiological and psychological needs (e.g., to be fed, feel safe, belong, be productive, unique, empowered) are being met. If a child is abused, ignored, or neglected, there are often negative behavioral, cognitive, and emotional outcomes.

Problem behaviors such as tantrums, aggression toward others, or withdrawal may result from the child's emotional and social turmoil. However, as was mentioned above, it is important to remember that negative behaviors may arise from vastly different reasons than experiences of abuse or neglect (e.g., biochemical or physiological factors).

There are many instruments available to assess a child's emotional and social functioning. Salvia and Ysseldyke (1991) suggest several ways in which personality variables may be measured. The use of Rating Scales was discussed above under "Behavioral Assessment" and is applicable here as well. A second approach, using

projective techniques, asks students to respond to vague or ambiguous stimuli such as inkblots or pictures, draw pictures, or express themselves through the use of puppets or dolls. The responses are then interpreted by a person trained in such procedures. A third approach is to administer personality inventories or questionnaires that vary in their focus. Some may measure self-concept or learning style, while others are intended to indicate the possible presence of mental illness. These latter instruments are generally lengthy and present the individual taking them with a substantial reading load, both in terms of how much there is to read and in terms of how complex and abstract many of the ideas are. Thus, many such inventories are not suitable for individuals with low literacy. Furthermore, as Berdine and Meyer (1987) remark, "Many of these measures suffer from technical inadequacies and yield esoteric results that are difficult to translate into treatment goals" . For this reason, while information gathered through these instruments may help the assessment team understand the student more fully, information collected through approaches such as direct observation and interviews may be more useful and reliable.

4. Roles of Assessment and Possible Services for CWSNs and their Parents

4.1 Roles of Assessment in Decision-making

1. ***Evaluation Decisions:*** *Information collected in the assessment process can provide detailed information of a student's strengths, weaknesses, and overall progress.*

2. *Diagnostic Decisions*: Information collected in the assessment process can provide detailed information of the specific nature of the student's problems or disability.

3. *Eligibility Decisions*: Information collected in the assessment process can provide detailed information of whether a child is eligible for special education services.

4.*IEP (IEP Development) Decisions*: Information collected in the assessment process can provide detailed information so that an Individualized Education Program (IEP) may be developed

5.*Educational Placement Decisions:* Information collected in the assessment process can provide detailed information so that appropriate decisions may be made about the child's educational placement

6. *Instructional Planning Decisions:* Information collected in the assessment process is critical in planning instruction appropriate to the child's special social, academic, physical, and management needs

4.2 Possible Services for CWSNs and their Parents

✓ **Audiology:** This service includes identifying the hearing loss, determining the need for amplification, and providing auditory training, aural rehabilitation, speech reading, and other services. Guidance and counseling for children, parents' and teachers included.

✓ **Counseling**

  ✓ *Social workers, psychologists, guidance counselors, and other trained personnel assist the family in understanding the special needs of the child and enhancing the child's development.*

  ✓ *Early identification and assessment of disabilities in children*

  ✓ *These services include identifying a disability as early as possible in a child's life.*

♣ **Medical** services are provided by a licensed physician to determine a child's medically related disability.

♣ **Occupational Therapy**: These services are provided to improve, develop or restore the child's functional ability.

♣ **Orientation and mobility services**: Students who are blind or visually impaired can receive these services to attain systematic orientation and safe movement within the school, home, and community environments.

♣ **Parent counseling and training**: These services are designed to assist parents in understanding the special needs of their child, provide parents with information about child development, and help parents to acquire the necessary skills to support their child's IEP.

♣ **Physical Therapy:** These services promote sensorimotor function.

♣ **Recreation:** These services include: assessment of leisure functions, therapeutic recreation programs, and leisure education.

♣ **Psychological services**: These services include:

   ✓ *Administering psychological and educational tests and other assessment procedures;*

   ✓ *Interpreting assessment results;*

   ✓ *Obtaining, integrating, and interpreting information about behavior and conditions relating to learning;*

   ✓ *Consulting with other staff members in planning school programs;*

   ✓ *Providing psychological counseling to children and parents; and*

   ✓ *Assisting in developing positive behavioral intervention strategies.*

♣ **Rehabilitation counseling**: Group or individual counseling sessions focus on career development, employment preparation, achieving independence, and integration in the workplace and community.

♣ **School health:** These services are provided by the school nurse.

♣ **Social work:** These services include preparing a social or developmental history on the child;

♣ **Counseling with the child and family**; Working with parents and others on those problems in the home, school, and community that affect the child's adjustment to school;

♣ **Mobilizing school and community resources** to enable the child to learn as effectively as possible; and

♣ **Assisting in developing positive behavioral interventions strategies.**

♣ **Speech-language pathology:** This area includes identifying children with speech or language impairments, diagnosis, referral for medical or other professional attention, provision of speech and language services, counseling and guidance of parents, children, and teachers regarding speech and language impairments.

♣ **Transportation:** These services consist of travel to and from school and between schools, travel in and around school buildings, specialized equipment, such as lifts, ramps, and so on.

4.3 Ethical Considerations in Assessment

   ✓ Right of notification before assessment

   ✓ Right of consent or agreement on the assessment process

   ✓ Right of evaluation using multiple measures

✓ Right to request for an independent evaluation

✓ Right of participation in the process of assessment

✓ Right to notification on proposed changes of educational program

✓ Right to access all educational records

✓ Right to educational records confidentiality

✓ Right of hearing the assessment results

4.3 Roles of Parents in the Assessment Process

While designing, conducting, interpreting, and paying for the assessment are the school system's responsibilities, parents have an important part to play before, during, and after the evaluation. The purpose of this section is to provide parents with suggestions for the range of ways in which they might involve themselves in the assessment of their child. The extent, to which parents involve themselves, however, is a personal decision and will vary from family to family.

Before the evaluation, parents*:*

✓ May initiate the evaluation process by requesting that the school system evaluate their child for the presence of a disability and the need for special education.

✓ Must be notified by the school, and give their consent, before any initial evaluation of the child may be conducted.

✓ May wish to talk with the person responsible for conducting the evaluation to find out what the evaluation will involve.

✓ May find it very useful to become informed about assessment issues in general and any specific issues relevant to their child (e.g., assessment of minority children, use of specific tests or assessment techniques with a specific disability).

✓ May need to advocate for a comprehensive evaluation—one that investigates all skill areas apparently affected by the suspected disability and that uses multiple means of collecting information (e.g., observations, interviews, alternative approaches).

✓ May suggest specific questions they would like to see addressed through the evaluation.

✓ Should inform the school of any accommodations the child will need (e.g., removing time limits from tests, conducting interviews/testing in the child's native language, adapting testing environment to child's specific physical and other needs)?

✓ Should inform the school if they themselves need an interpreter or other accommodations during any of their discussions with the school?

- ✓ May prepare their child for the evaluation process, explaining what will happen and, where necessary, reducing the child's anxiety. It may help the child to know that he or she will not be receiving a "grade" on the tests he or she will be taking but that the purpose behind any testing is to gather information to help the student succeed in school.

During the evaluation process, parents:

- ✓ Need to share with the school their insights into the child's background (developmental, medical, and academic) and past and present school performance.
- ✓ May wish to share with the school any prior school records, reports, tests, or evaluation information available on their child.
- ✓ May need to share information about cultural differences that can illuminate the educational team's understanding of the student?
- ✓ Need to make every effort to attend interviews the school may set up with them and provide information about their child.

After the evaluation, parents:

- ✓ Need to carefully consider the results that emerge from their child's evaluation, in light of their own observation and knowledge of the child. Do the results make sense in terms of the behaviors, skills, needs, and attitudes they have observed in their child? Are there gaps, inconsistencies, or unexpected findings in the results that parents feel are important to address, if a comprehensive picture of the student's strengths and needs is to be developed?
- ✓ May share their insights and concerns about the evaluation results with the school and suggest areas where additional information may be needed. Schools may or may not act upon parents' suggestions, and parents have certain recourses under law, should they feel strongly about pursuing the matter.
- ✓ Participate fully in the development of their child's Individualized Education Program (IEP), using information from the evaluation.