ORGANIZATION OF EASTERN CARIBBEAN STATES

EDUCATION SECTOR STRATEGY

OECS PRIMARY GRADES' ASSESSMENT MANUAL

TABLE OF CONTENTS

CHAPTER 1

IMPORTANT CONCEPTS

1.1 Purpose of Manual

The purpose of this manual is to present a set of principles and procedures that classroom teachers should know and apply in teaching pupils in primary schools in the Caribbean. In particular, teachers of the OECS territories are urged to use these principles and procedures as they prepare their pupils for regional standardized assessments. It is critical that teachers learn how to construct meaningful tasks. They also need to be proficient in scoring them using acceptable procedures.

Also teachers need to obtain focused and systematic feedback from pupils. This is necessary for teachers to do a proper job in educating those in their charge. Educators today recognize that evaluating pupils' achievement by a one shot examination at the end of the term or year is not a fair method of assessing pupils' competencies and achievement. Teachers need to obtain constant feedback of the pupils' progress and pupils need constant feedback from the teachers on how well they are doing.

The OECS Assessment Framework requires teachers to engage pupils in certain internal assessment tasks, particularly projects, writing portfolios and book reports. The assessment concepts and skills covered in this manual should assist teachers in this regard. Another objective of the manual is to highlight a set of assessment and evaluation principles and procedures that will enable teachers to use multiple sources of evidence to provide a clearer picture of how well their pupils are doing with a view to help them to improve and move on to higher heights.

In addition, samples of tasks for pupils in various grades (Grades K to 6) and a variety of sample scoring rubrics that can be used to score tasks are provided. Teachers should feel free to use these as prototypes and may modify them for use in their classrooms. Teachers need to avoid the one size fits all approach. Adjustments are needed to meet specific class situations.

Formative assessment takes place during instruction. Its main focus is to provide feedback to help you and the pupils make adjustments in order to attain the objectives of your programme. A fuller discussion on this point will follow later.

The booklet is organized as follows:

In the remainder of Chapter 1, several important concepts that teachers ought to understand are explained. In Chapter 2, Frameworks for organizing teaching, learning and assessment are discussed.

Chapter 3 explains the importance of Tables of Specification and illustrates how they can be used to construct tests. Chapter 4 provides guidelines on how to construct various test items. Several examples are provided so that teachers can hopefully find one to match their context. The remaining chapters deal with how to construct a variety of tasks, including performance assessment tasks, projects, and investigations. Several types of scoring rubrics as well as important psychometric issues that affect assessments are covered.

Whilst the purpose of the manual is to ensure that teachers acquire the skills needed to prepare pupils adequately for the OECS Primary Exit Assessments, nonetheless this writer assumes that teachers are professionals and that they will not blindly follow an example without adjusting it where the context in which they are working suggests that an adjustment is in order.

1.2 Explanation of important concepts in this manual

Measurement

Measurement may be perceived as the process by which one assigns numbers to attributes, traits, characteristics, or behaviours of persons using clearly defined explicit rules. In the classroom situation, measurement is perceived as a systematic procedure for assigning numbers or scores to represent the degree or level to which a trait, attribute, skill or knowledge of subject matter is present in an individual's performance on a test: For example, if a vocabulary test is administered, the measurement procedure that is used to arrive at the score or the mark must be explicit enough that another can understand how the score was derived.

Hopkins (1998) emphasises that measurement involves a process by which things are differentiated and described. The process is not linked exclusively to the use of highly developed refined instruments. For example, one can tell whether the day is 'hot' or 'cold' without using an instrument. Yes, it would be more accurate if we use an instrument to measure the temperature. This makes the point that the validity of the results of a measurement procedure has much to do with the process and the quality of the instrument used. When therefore we create an instrument to measure reading ability or writing skills or any set of knowledge or skills the validity of the results is a function of the quality of the instrument used as well as the procedure used to arrive at a score.

Evaluation

Evaluation is conceptualized as a process of making a value judgement about the worth of a pupil's writing, performance, reading ability or some other cognitive skill or behaviour. It is a general process of making decisions about some educational or psychological attribute based on data. Nitko and Brookhart (2007) illustrate this by the following example. Evaluation may or may not be based on measurements or test results. It sometimes occurs in the absence of tests or objective data, but this is subject to bias. Tests and measurements help to reduce the amount of subjectivity in the evaluation.

Evaluation may be formative or summative. Formative evaluation involves judging the worth of pupils' performance or achievement while learning is taking place. During the course of the instruction, we make informal evaluation of whether pupils understand or not, and adjust our instruction accordingly. Summative evaluation on the other hand, involves judging the quality of performance or achievement at the end of the instructional period. This may be a day, a few weeks, a term or a year. Report cards, marks or letter grades may be given to represent our summative evaluation of pupils' achievement.

Assessment

Assessment is considered as a process for obtaining information to be used to make decisions about pupils' learning, the curricula or programmes or educational policy. It is concerned with the totality of the educational setting. It is an inclusive term and subsumes measurement and evaluation. Assessment focuses not only on the nature of learning, but also on what is learnt and how it is learnt.

Current views on assessment emphasize the importance of assessment **for** learning and assessment **as** learning. The final examination and end-of- unit test are examples **of** assessment of learning. In this case the focal point of interest is how much has been learned at this point. This is of course necessary for certification purposes and there is nothing wrong with that. However, current best practices suggest that teachers should also be interested in the formative aspects of assessment, where the teacher and pupil alike use the assessment results to identify strengths and weaknesses, so that the pupil can plan to improve and the teacher can plan to use other teaching strategies to ensure better understanding of the work. Table 1.1 provides a summary of these points. In short:

Assessment is the process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what pupils know, understand, and can do with their knowledge as a result of their educational experiences; the process culminates when assessment results are used to improve subsequent learning (Huba & Freed, 2000, p.10).

It is emphasized that assessment is really a process, not a test or instrument. It is a process of collecting information in order to make educational decisions. Huba and Freed illustrate the assessment process in the figure below (Figure 1.1).

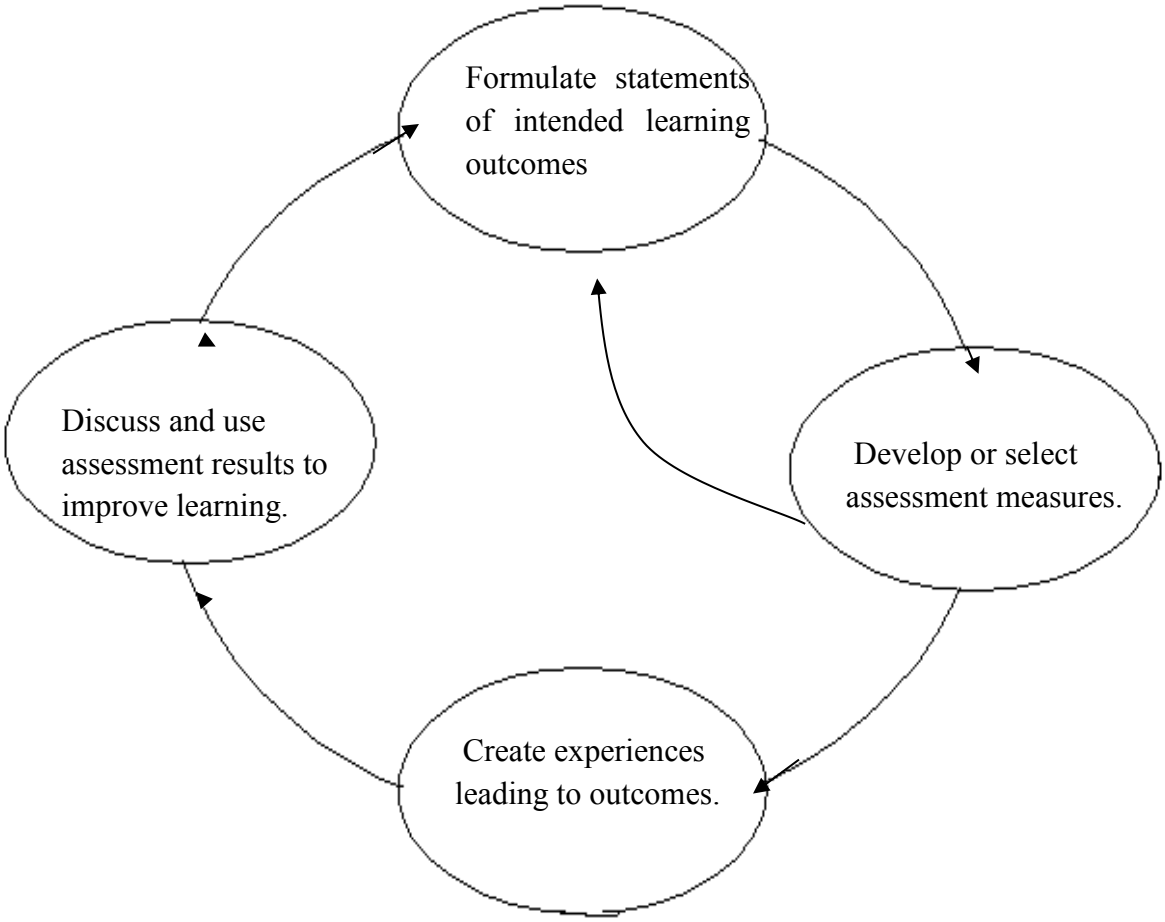


Figure 1. 1 The assessment process. Source: Huba and Freed, 2000. p.10.

1.3 The forms of assessment

Pre-instruction Assessment

Teachers frequently have to assess pupils at the beginning of instruction, for example when faced with a new group of pupils at the beginning of a new unit of work. The teacher would want to know if the pupils have the prerequisite skills on which to build the concepts and principles to be taught in the unit or course of study. This type of *pre-test* is called a *readiness test.* Sometimes too, the teacher wants to know if the work is known already, in which case teaching the unit would be a waste of time. The information gathered from a pre-test helps the teacher to make sensible pre-instructional decisions, to set appropriate learning targets, and plan instructional activities in a meaningful way.

Strategies that can be used for this type of assessment include, but limited to:

1. Informal chats with previous instructors,
2. Examining pupils' records or data on previous tests.
3. Questioning pupils prior to teaching.
4. Administering diagnostic tests, pre-tests, oral tests.
5. Engaging pupils in classroom discussions. –

Instructional assessment is vital for effective instructional planning.

Formative Assessment: Assessment during Instruction

"Formative assessment exists for exactly one reason: to enhance pupils' learning" (Popham, 2008, p. 70). Formative assessment takes place during instruction. Formative assessment is a process used by teachers during instruction to provide feedback to adjust ongoing teaching and learning. During the course of instruction teachers would want to know if they are getting across the instruction effectively and if pupils are achieving the learning targets. Some strategies may be considered as informal assessments. Teachers observe pupils, ask questions, interview pupils and form some assessment of the standard of the pupils' work. Class discussions also provide information to teachers about what pupils understand or do not understand. However, sometimes the assessment is more formal, as in the form of written exercises and practice tests.

When reading problems are so persistent that re-teaching is unlikely to correct the problems, then teachers need to resort to *diagnostic testing*. Diagnostic tests are designed to probe more deeply into the causes of reading deficiencies. Diagnostic tests will usually consist of several

exercises arranged in some ordered sequence in such a way that successive sets of examples are increasingly more challenging that the preceding ones.

Mc Millan (2007) advises that the goal of formative assessment is "the improvement of pupil motivation and learning" (p.118). He further advises that "teachers need to employ a circular continuing process involving their evaluation of pupil work and behaviour, feedback to pupils and instructional correctives". The correctives must consist of new strategies and approaches. Correctives must be qualitatively different from the initial teaching.

Internal Assessment

Internal assessment refers to assessment carried out by the classroom teacher. The tasks and mark schemes may be created by the teachers themselves. There are cases however where a test is constructed by an external agent but it is administered and scored by the classroom teacher. For the purpose of this booklet however, internal assessment is used to **describe classroom tests or assignments constructed and scored by the classroom teacher.**

Summative Assessment: Assessment after Instruction

This type of assessment focuses on attainment or achievement levels at the end of a unit, a set of units or at the end of a course. The purpose is to assess pupils' competence, abilities, knowledge or skills at the end of a particular period of work. The score is used to indicate pupils' competencies at the point in time.

It is important to recognize that the purpose of the assessment is the overriding factor in distinguishing between formative and summative assessment. It can be argued that tests given during a course can be used for summative purposes, to indicate pupils' state of knowledge or competency at the end of a course. Likewise, it can be argued that a test given at the end of a unit course can be used as formative assessment in the sense that the results are used to evaluate pupils' strengths and weaknesses and the results are used to improve subsequent teaching and learning. In other words, the teacher then uses the information to re-teach sections of the course.

Summative assessments are used to determine the mastery level of pupils at specific points in time. The information is used to determine which pupils are ready for the next stage in the instructional process or for the next level of a programme. Since this information gathered from these tests is often used to make critical decisions, it is important that teachers learn how to construct valid, fair and reliable assessment instruments.

Continuous Assessment

The suggested OECS Assessment system is built on principles of continuous assessment. Nitko (2004) defines continuous assessment as a daily process by which teachers gather information about pupils' progress in achieving the learning targets prescribed in the curriculum. Halliday (2005) points out that there are three key points about this definition. First, continuous assessment is a daily process; it is ongoing. Second, it involves gathering information about progress or lack of progress. Third, the progress expected is geared towards prescribed learning targets or expected learning outcomes. This definition helps us to see the link between assessment, curriculum and instruction. Macintosh (1974) suggests that continuous assessment is a continual updating of teachers' judgements about pupils. The word updating implies that teachers use systematic record keeping in order to up-date their records of pupils' progress. Primary school teachers are urged to use continuous assessment strategies to evaluate pupils' performance.

Principles governing good assessment practices

Halliday (2005) lists the following as key principles governing good assessment practice:

1. Assessments should be closely aligned to the course content and the learning outcomes.
2. Classroom assessment should be based on the subject matter taught in the course.
3. Assessment tasks should reflect what is important to learn in the particular course.
4. Assessments should support good instructional practice and should enhance learning.
5. Assessment must be fair. Assessments should be designed to give pupils a fair chance to show what they know and can do.
6. Methods of assessment should take into account different learning styles of pupils.
7. Multiple sources of information from a variety of assessment formats should be used
8. The assessment process should be open and transparent. All the important rules that will be applied in scoring and grading should be made available for scrutiny.
9. Accommodation strategies must be developed to facilitate pupils who are physically and visually challenged.
10. Assessments should be used to provide positive feedback to pupils.

The following table summarises the purposes of assessment.

Table 1.1 ASSESMENT of LEARNING, for LEARNING, and as LEARNING

Assessment OF Learning	Assessment FOR Learning	Assessment AS Learning
Summative	Formative	Assessment used to engage learning
Certifies learning	Describes need for future learning	Fosters pupil self-monitoring of learning
Conducted at end of unit or term or course	Conducted during instruction of a unit	Conducted during instruction of a unit
Scoring may focus on rank order of pupils	Scoring focuses on achievement	Scoring focuses on achievement
Used mainly to report on progress at specific times	Used to modify instruction	Used to show importance of unit
Specific periods	Ongoing	During instruction
Used to report to parents or other bodies	Used to give feedback to pupil	Fosters pupil self-monitoring
Can decrease pupil motivation	Enhances pupil motivation	Enhances pupil motivation
Summary judgements	Diagnostic	Diagnostic
Focus on reliability	Focus on validity	Focus on validity
Delayed feedback	Immediate feed back	Immediate feedback

CHAPTER 2

FRAMEWORKS

When crafting instruments to be used in the assessment process, it is important to focus not only on the knowledge and understanding of concepts and principles, but on higher order thinking skills as well. Teachers are required to help pupils to develop a wide range of higher order thinking skills, such as problem solving, deductive and inductive reasoning, analyzing and evaluating arguments for example. These skills can be identified in the various taxonomies or cognitive frameworks for organizing classroom instruction and assessment.

Several educators have proposed various lists of these skills that are useful for classroom instruction and assessment. Some of these are described as taxonomies, whilst others may be referred to as classification schemes or frameworks for organizing teaching and assessment activities. Three of these are discussed in the next few pages.

2.1 Bloom's Taxonomy

One of the most widely used taxonomies for organizing school was developed by Benjamin Bloom and a group of educators in 1956. This taxonomy has dominated the education field for over 60 years and has been used to organize curriculum activities, assessment schemes and evaluation projects.

Bloom's Taxonomy divides the way a person learns into three domains:

Cognitive Domain – deals with thinking

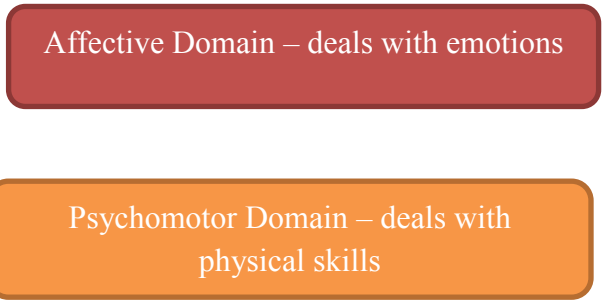


Figure 2.1 Bloom's Taxonomy

The Cognitive Domain emphasises intellectual outcomes and corresponds to the development of mental skills. It is arranged in six levels of increasing cognition or levels of thinking. Two versions of the cognitive domain taxonomy are used. The original version described states of cognition. The revised version, developed by Krathwhol and Anderson in 2001, made a few changes to the model, the most significant being the change from states of cognition (nouns) to thinking as a hierarchy of actions (verbs). In addition, Evaluation was reassigned to a lower level on the hierarchy and Synthesis was subsumed in the new upper level called Creation.

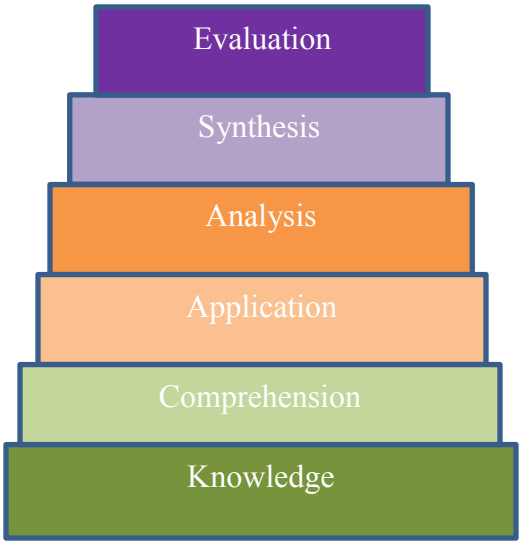


Figure 2.2 The Cognitive Taxonomy

A summary of the six levels of the taxonomy follows:

Knowledge

The Knowledge objectives emphasize the process of remembering what is learned. Pupils must recall specific facts, methods, definition of concepts, terms and terminology. Assessment tasks at this level usually require pupils to list, state, identify, describe or define specific concepts, principles or procedures. The importance of this is that whatever the body of knowledge one teaches or learns, one must know the basic concepts of that body of knowledge.

Comprehension

The Comprehension objectives emphasize the fact that pupils need to understand what has been communicated to them and should be able to restate that material in a different form. Pupils may be required to restate, paraphrase, explain or interpret the learned material. Mathematics, for example, processes requiring pupils to translate from one mathematical mode to another. Examples include converting fractions to decimals and vice versa, and translating words to symbolic (algebraic) expressions. Action words used to describe learning targets at this level include estimate, explain, predict, infer, rewrite, summarize, and give examples of the particular concept.

Application

Application requires pupils to apply principles, theories and algorithms learned in the discipline. Assessment tasks at the application level may require pupils to solve problems and implement principles and procedures established in the discipline. Some educators would insist that the problems should not be routine ones, but different from those encountered during instruction. Action words used at this level may include solve, show, compute, demonstrate, prepare, produce, use and apply.

Analysis

Tasks at this level require the breakdown of materials into constituent parts, identifying the relationships of those parts or detecting the way in which these parts relate to each other. Analytical skills are required in a variety of situations - in arguing a case, in experiments and so forth. Action words used in learning targets at this level include differentiate, distinguish, illustrate, outline, infer, relate, identify, and break down.

Synthesis

The Synthesis objectives require the putting together of two or more elements or parts in such a way as to constitute a pattern or structure not presented in that form previously. Pupils may be required to integrate, organize, combine, construct, or design something. Action words used in learning targets at this level include combine, compose, create, construct, design and devise.

Evaluation

This involves appraising, assessing or judging the worth of something. The judgement may require the use of internal evidence or the use of external criteria. Action words used in learning targets at this level include compare, contrast, appraise, explain, justify and support.

2.2 General comments

It should be understood however, that the action word used does not of itself identify the taxonomic level. One can be asked to explain something at the comprehension level, at the analysis level and even at the evaluation level. It depends on what is required, whether the pupil is merely asked to explain a term, or to explain the reasons for a phenomenon that involves analyzing issues or whether in the explanation the pupil has to make judgements.

Second, although many educators accept that the six levels of Bloom's taxonomy are useful, they prefer to think of them as forming a classification scheme, rather than a taxonomy or hierarchy of skills. For example, some educators would argue that '*infer*' and '*predict*', which are usually placed at comprehension level, in Bloom's taxonomy, involve higher order skills that require a considerable amount of analysis and synthesis. In addition, research has shown that teachers classify items differently. There is no unique way of classifying the objectives. As a result, several educators have developed other schemes with slightly different emphases. One curriculum framework that has similarities to Bloom's taxonomy, but emphasizes some important differences is one prepared by Quellmalz and other educators. **(*See Appendix)***

Table 2.1 Taxonomy of Educational Objectives - Summary

1.0	KNOWLEDGE
1.1	Knowledge of specifics
1.2	Knowledge of ways and means of dealing with specifics
1.3	Knowledge of universal and abstractions in a field
2.0	COMPREHENSION
2.1	Translation
2.2	Interpretation
2.3	Extrapolation
3.0	APPLICATION
4.0	ANALYSIS
4.1	Analysis of elements

4.2	Analysis of relationships
4.3	Analysis of organizational principles
5.0	SYNTHESIS
5.1	Production of a unique communication
5.2	Production of a plan, or proposed set of operations
5.3	Derivation of a set of abstract relations
6.0	EVALUATION
6.1	Judgement in terms of internal evidence
6.2	Judgement in terms of external criteria

The Quellmalz taxonomy and a comparison with that of Bloom's is provided in the appendices so that teachers can gain some insight into the various skills on which one needs to focus in teaching and assessing pupils. Marzano's Taxonomy of Core Thinking Skills is also provided in the Appendix. The teacher should note the vast variety of skills identified in this taxonomy.

Many teachers tend to use this taxonomy because it is relatively easy to use. Teachers should also consider using the revised taxonomy which is discussed below. It may seem a lot to consider, but some helpful suggestions on how to use it are provided in the ensuing pages.

2.3. Revisions to Bloom's Taxonomy

Two revisions of Bloom's Taxonomy have been proposed in 2001. Both revisions seek to take into consideration advances in the thinking of cognitive psychologists and the results of research undertaken by several educators on the use of cognitive frameworks to organize learning and assessment. In particular, it is now recognized that there are different types of knowledge. It is also recognized that the ability to organize one's learning is an important factor in learning. Thus meta-cognitive skills are regarded as critical to teaching, learning and assessment.

Space does not permit a full discussion of the two revisions. However, the revised version, made a few changes to the model, the most significant being the change from states of cognition (nouns) to thinking as a hierarchy of actions (verbs). In addition, Evaluation was reassigned to a lower level on the hierarchy and Synthesis was subsumed in the new upper level called Creation.

Figure 2.3 Revised Cognitive Taxonomy

In this manual we confine our discussion to the revision proposed by Anderson and Krathwohl (2001). This revision is selected because it utilizes the major ideas of the original Bloom's Taxonomy and improves upon them. Krathwohl, one of the key authors of the old taxonomy, is also one of the key authors of the revision. Teachers can more easily relate to this revision than the other revision proposed by Marzano et al (2001). Moreover, the Anderson and Krathwohl version is easier for teachers to use from a practical point of view.

2.4 The Knowledge Dimension of the Revised Taxonomy

The new taxonomy divides *knowledge* in four subtypes: ***Factual Knowledge***, ***Conceptual Knowledge***, ***Procedural Knowledge*** and ***Meta-cognitive Knowledge***. In addition, there are six cognitive processes that are similar to those of the old taxonomy. The different types of knowledge and brief explanations are given in Table 3.1, which is taken from Anderson and Krathwohl (2001), p. 46. Descriptions of the six (6) skills: *Remember, Understand, Apply, Analyse, Evaluate* and *Create* are given in Table 3.2, which is taken from Anderson and Krathwohl (2001), p. 31.

To apply these new ideas, the reader should consider the matrix set out in Table 3.3. Each learning target or each assessment task should fit in the matrix as one of the four (4) types of knowledge that requires one of the six (6) levels of skills.

Table 2.2 Main Types and Subtypes of the Knowledge Dimension

MAJOR TYPES AND SUBTYPES		EXAMPLES
1.0	**FACTUAL KNOWLEDGE** – The basic elements pupils musk know to be acquainted with a discipline or solve problems in it	
1.1	Knowledge of terminology	Technical vocabulary, musical symbols
1.2	Knowledge of specific details and elements	Major natural resources, reliable sources of information
2.0	**CONCEPTUAL KNOWLEDGE** – The interrelationships among the basic elements within a larger structure that enable them to function together	
2.1	Knowledge of classifications and categories	Periods of geological time, forms of business ownership
2.2	Knowledge of principles and generalizations	Pythagorean theorem, law of supply and demand
2.3	Knowledge of theories, models, and structures	Theory of evolution, structure of Congress
3.0	**PROCEDURAL KNOWLEDGE** – How to do something, methods of inquiry, and criteria for using skills, algorithms, techniques, and methods	
3.1	Knowledge of subject-specific skills and algorithms	Skills used in painting with watercolors, whole-number division algorithm
3.2	Knowledge of subject-specific techniques and methods	Interviewing techniques, scientific method
3.4	Knowledge of criteria for determining when to use appropriate procedures	Criteria used to determine when to apply a procedure involving Newton's second law, criteria used to judge the feasibility of using a particular method to estimate business costs
4.0	**METACOGNITIVE KNOWLEDGE** – Knowledge of cognition in general as well as awareness and knowledge of one's own cognition	
4.1	Strategic knowledge	Knowledge of outlining as a means of capturing the structure of a unit of subject matter in a text book, knowledge of the use of heuristics
4.2	Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge	Knowledge of the types of tests particular teachers administer, knowledge of the cognitive demands of different tasks
4.3	Self-knowledge	Knowledge that critiquing essays is a personal strength, whereas writing essays is a personal weakness; awareness of one's own knowledge level

Table 2.3: Categories of the Cognitive Process Dimension and Related Cognitive Processes

PROCESS CATEGORIES		COGNITIVE PROCESSES AND EXAMPLES
1.0	**REMEMBER** – Retrieve relevant knowledge from long-term memory	
1.1	Recognizing	E.g. Recognize the dates of important events in U.S. history
1.2	Recalling	E.g. Recall the dates of important events in U.S. history
2.0	**UNDERSTAND** – Construct meaning from instructional messages, including oral, written, and graphic communication	
2.1	Interpreting	E.g. Paraphrase important speeches and documents
2.2	Exemplifying	E.g. Give examples of various artistic painting styles
2.3	Classifying	E.g. Classify observed or described cases of mental disorders
2.4	Summarizing	E.g. Write a short summary of the events portrayed on videotapes
2.5	Inferring	E.g. In learning a foreign language, infer grammatical principles from examples
2.6	Comparing	E.g. Compare historical events to contemporary situations
2.7	Explaining	E.g. Explain the causes of important eighteenth-century events in France
3.0	**APPLY** – Carry out or use a procedure in a given situation	
3.1	Executing	E.g. Divide one whole number by another whole number, both with multiple digits
3.2	Implementing	E.g. Determine in which situations Newton's second law is appropriate
4.0	**ANALYZE** – Break material into constituent parts and determine how parts relate to one another and to an overall structure or purpose	
4.1	Differentiating	E.g. Distinguish between relevant and irrelevant numbers in a mathematical word problem
4.2	Organizing	E.g. Structure evidence in a historical description into evidence for and against a particular historical explanation
4.3	Attributing	E.g. Determine the point of view of the author of an essay in terms of his or her political perspective
5.0	**EVALUATE** – Make judgements based on criteria and standards	
5.1	Checking	E.g. Determine whether a scientist's conclusions follow from observed data
5.2	Critiquing	E.g. Judge which of two methods is the best way to solve a given problem
6.0	**CREATE** – Put elements together to form a coherent or functional whole; reorganize elements into a new pattern or structure	
6.1	Generating	E.g. Generate hypothesis to account for and observed phenomenon
6.2	Planning	E.g. Plan a research paper on a given historical topic
6.3	Producing	E.g. Build habitats for certain species for certain purposes

For classroom work, the two sets of dimensions can be collapsed in a matrix shown in Table 2.4.

Table 2.4: Suggested Plan for using the Anderson and Krathwohl Taxonomy

KNOWLEDGE DIMENSION		Cognitive Process Dimension					
		1	2	3	4	5	6
		Remember	Understand	Apply	Analyze	Evaluate	Create
A	Factual Knowledge						
B	Conceptual Knowledge						
C	Procedural Knowledge						
D	Meta-cognitive Knowledge						

This matrix is recommended for use in primary and secondary classes (Halliday, 2004).

2.5 Using the Revised Taxonomy to Classify Learning Outcomes and Test Items

Table 2.4 shows the categories of the revised taxonomy consisting of six (6) verbs on the horizontal dimension and four (4) nouns on the vertical dimension. To determine how to classify a learning objective or a test item, the teacher would ask first: What it is that is the main focus of the objective or test item, that is, what type of knowledge is required? One of the following will be the appropriate answer.

1) Factual knowledge
2) Conceptual knowledge

3) Procedural knowledge
4) Meta-cognitive (which includes strategic) knowledge

Second, the teacher should consider what is the major skill emphasized in a particular objective or test item. One of the following should be the appropriate answer.

(1) recall information, remember
(2) understand
(3) apply principles
(4) analyze information
(5) evaluate data
(6) create a plan or solution.

The answers to these two questions would enable the teacher to classify the objective or test item. The answer will be identified by a numeral to identify the skill and a letter of the alphabet to identify the type of knowledge. (e. g 1A , 3B or alternately A1 and 3B)

For example: Consider the following standard:

Identify the names of fractional parts when a pictorial representation in given.

In this standard, the names of fractional parts are the focus of attention. This is clearly factual knowledge that is required. Now pupils have to be able to identify the names and match them to pictorial representations. There will be some of these that pupils will remember and recall from memory. Hence that objective can be classified as **A1** in the taxonomy matrix. If however the pictorial representation is not one that can be easily recalled, but requires some level of comprehension, perhaps a step is necessary to make it a recognizable one. In that case the classification level would be **A2**.

Suppose we have to classify the standard/objective '*Select sources of information to write a project report'.* Conceptual Understanding in involved and the skill here involves more, the action word 'select' requires one to '*analyze*' first. We therefore assign this objective to cell **B4**.

2.6 Test items

Mathematics: Calculate 20% of $260.

This task requires Procedural Knowledge as the pupil is required to apply the procedure correctly for finding the relevant percentage value. The item might be classified as C3.

Here is an example adapted from Anderson & Krathwohl (2001 p. 211). The task is: *Deliver a report to the class on a particular project.* This task requires not only Factual Knowledge, but also

Conceptual Knowledge. The speaker must use several skills, the highest one is the ability to '*Create*'; hence, we can classify that objective as both **A6** and **B6**.

Task No 1

Study the objectives below and the suggested cell in the taxonomy that the item fits. Discuss with a colleague or two, the suggested fit and you disagree with the classification, write the one you think is appropriate.

Task 2. Identify the strand and standards in the mathematics standard booklet that the six examples in Table 2.5 match most closely.

Table 2.5 Examples showing how to use the new taxonomy

Objective		Suggestion	Revision
1.	Add two fractions with like denominators	C2	
2.	Differentiate between 2D shapes	B4	
3.	Make change from $500	C4	
4.	Use latitude and longitude to find a country on a map of the Caribbean	B3	
5.	Draw a sketch map of St. Lucia	C6	
6.	Locate the OECS territories on a map of the Caribbean	C3	

Task No. 3. How would you classify the mathematics items below?

1. Add two more terms to the following:

(a) 32, 42, 52, ____, ____.

(b) 2, 4, 7, 11, ____, ____.

2. Determine the area of a circle where the diameter is 14 cm.

Note: Keep in mind that different teachers might classify an item differently depending on the level of pupils they teach and the curriculum practices they employ. There may not be always total agreement. The important point is that teachers try to cover a wide a range of skills as possible in their classroom instruction and assessment.

2.7 Planning Tests Using the Revised Taxonomy

It might be difficult in some cases to use a 4 x 6 matrix to plan the day-to-day classroom quizzes and tests. In this case a simple 2 x 3 matrix is recommended. By combining adjacent parts of the taxonomy table, it may be useful to work with the following table (Halliday, 2004).

Table 2.6 Suggested framework for planning tests

		Lower Order	Middle Order	Higher Order
		Remember Understand	Apply Analyze	Evaluate Create
FACK	Factual and Conceptual Knowledge			
PAMS	Procedural and Strategic/ Meta-Cognitive Knowledge			

This simple matrix is recommended for use in schools (Halliday, 2004).

In Table 2.6, Factual and Conceptual knowledge have been combined into one group, using the acronym **FACK**. This first category represents the body of facts and concepts that should be known and understood. The second category required knowledge of procedures, meta-cognitive and strategic knowledge. This is given the acronym **PAMS**. For the six (6) skills, I recommend the following: "*Remember"* and "*Understand"* are designated '***Lower Order',*** *Apply* and *Analyze* are designated '***Middle Order',*** and *Evaluate* and *Create are* designated '***Higher Order***'.

The number of items or the number of marks we assign to each of these six (6) cells will depend on the nature of the topic or unit we are assessing and the nature of the subject matter itself. In Mathematics, Science and practical subjects, it may be possible to assign about an equal number to each cell, whereas for history and social studies, FACK would probably require a higher weighting than PAMS, especially at the Primary and lower Secondary levels. Teachers would need to plan beforehand how to fill the cells. Two suggested examples follow.

Example 1: Social Studies Unit - Topic: Population

(i) Define various terms used in studying populations.
   (a) Census
   (b) Enumerator
   (c) Demography

(ii) Describe how population is measured.

(iii) List the factors that affect population changes e.g.

   a. Health standards
   b. Birth rates
   c. Death rates
   d. Migration
   e. Hazards/Disasters

(iii) Interpret data from graphs, tables or other graphic representations of population growth.
(iv) Use maps to show and interpret population density.
(v) Explain how governments use population data to plan developments for a country.

Table 2.7 Suggested framework for planning a Social Studies Test for Grade 5

	Lower Order	Middle Order	Higher Order
Factual and Conceptual Knowledge	Population terms (i) a, b and c	Population measurement (ii)	Interpret graphs, tables….population growth (iv)
Procedural and Strategic/ Meta-Cognitive Knowledge	Factors that affect population changes (iii) a, b, c, d, e	How governments use population data (vi)	Use maps to show and interpret density (v)

In Table 2.7, the topics are listed in the matrix to indicate the cognitive level that might be needed to deal with them.

Task: 4. *Discuss Table 2.7 with a colleague and if you have a different point of view, produce a revised table and give reasons for your decisions*.

Example 2: Mathematics Unit **-** Topic: Geometry

(i) Identify different types of angles
   (a) Acute angle
   (b) Right angle
   (c) Obtuse angle

(ii) List the relationship between the sides and angles of special quadrilaterals e.g.
   (a) Parallelogram
   (b) Trapezium
   (c) Kite
   (d) Rectangle
   (e) Square

Task No 5

Using the standards in Example 2 and the suggested framework at Table 2.7,

(i) Assign the standards to the various cells

(ii) Construct test items to fill these cells

(iii) Compare yours with the suggested on given in Table 2.8

2.8 The Affective Domain

The Affective Domain addresses the manner in which the learner deals with the subject matter emotionally. Manner may include feelings, values, motivation, appreciation, enthusiasm, and attitudes.

This domain lists the behaviours or reactions in a five-level hierarchy, starting from the simplest to most complex behaviour. The Affective domain can have a significant impact on student learning**.**

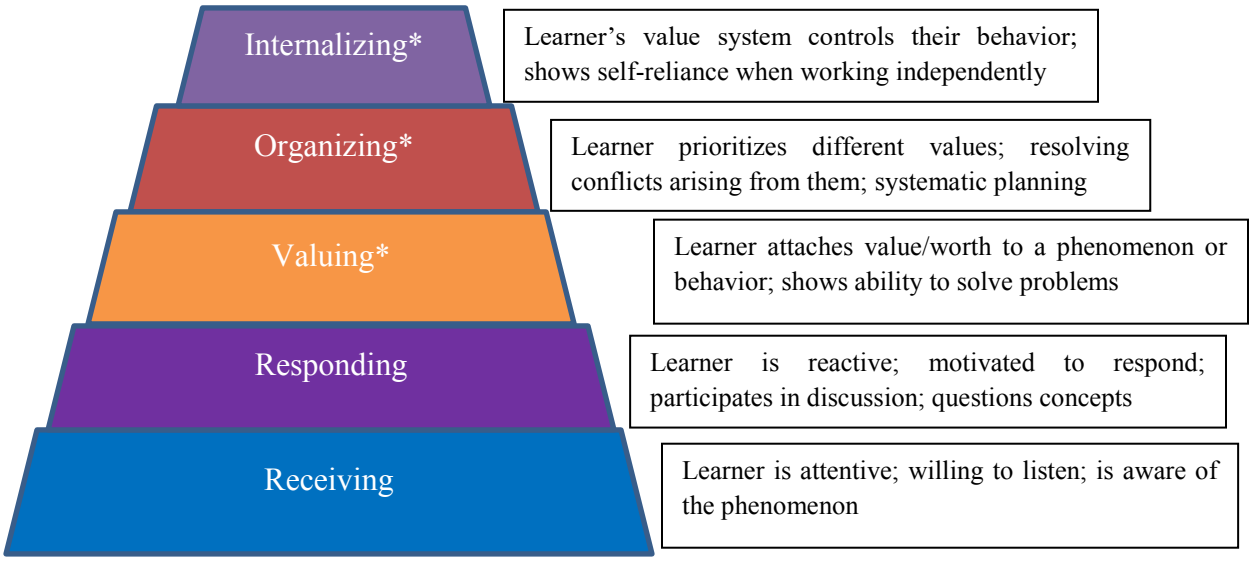


Figure 2.4 The Affective Taxonomy

CHAPTER 3

TEST SPECIFICATIONS/ASSESSMENT BLUEPRINT

3.1 Introduction

When planning tests or assessments, we need a test blueprint or table of specifications. Just as a builder needs a blueprint to use in order to construct a good dwelling, so too the examiner or test constructor needs test blueprints to construct high quality tests. Popham (2004) observes that the assessment blueprint serves as a basis for setting the number of assessment tasks and for assuring that the assessment will have the desired emphasis and balance. Popham (1978) identifies two main features of test specifications. First, they communicate to test users what the test measures. Second, they outline the details of the behavioural domain so that item writers can generate appropriate items. Test specifications are used to enhance the validity and reliability of the assessments. It also helps to ensure that the assessments are fair. {Issues of validity, reliability and fairness are discussed in Chapter 9.

There are many different ways to format a table of specifications. It may in some cases consist simply of a one-way grid showing the topics and objectives to be tested and the weight or number of marks to be assigned to each.

Table 3.1: A One-Way Assessment Grid: Language Arts: Comprehension

Reading Skill	No of items or Marks
A. Identifies details in a passage	10
B. Identifies the main ideas in a passage	10
C. Identifies the sequence of actions or events	10
D. Identifies relationships expressed in a passage	10
E. Identifies inferences drawn from a passage	10
Source: Linn & Miller; 2005 p. 144	

In the example above the content is spelt out by using specific standards. Usually some test blue prints will list the content or topics instead. Second, each objective was given the same weight or level of importance. This need not be so in every blueprint. Some topics or objectives may be

more important than others; hence the number of items can vary. The numbers could have been 10, 12, 8, 8, 12 instead of 10,10,10,10, 10 for the total number of marks. The example above is useful for a reading skills test, but the format is not appropriate for most classroom tests or examinations. We need the cognitive level for the items. All items need not be pegged at the knowledge level. Table 3.2 illustrates how a two-way grid is used for test specifications when the weight (i.e. the number of marks) for the topics /objectives) differ.

3.2 Improved Definition for Table of Specification

The most popular table of test specifications or assessment blueprint is a two-way table or grid in which the content/objectives are listed as one dimension, usually vertically and the intellectual skills or cognitive levels of functioning are listed on the horizontal part of the grid. The elements of the grid consist of the number of items or the number or proportion of marks to be allocated to that element.

Ward and Murray-Ward (1999) defined an assessment blueprint as an outline of content and skills to be covered on an assessment, along with the proportion of the assessment to be devoted to each topic and skill. They also pointed out that the method of assessment should also be indicated. Gronlund (2006) suggests that a table of specification is a two-way grid or chart that specifies the number or percentage of test items or assessment tasks to be designed for each area of content and each type of intended learning outcome (p. 224).

Usually the content, topics or objectives are written on the vertical dimension of the chart (grid) and the cognitive taxonomic levels or skills are written on the horizontal dimension of the chart (grid). This is a generic example; please insert the topics or concepts to suit your syllabus.

Table 3.2: Table of Specifications – A Two-Way Classification Grid

Content	**Recall & Comprehension**	**Application**	**Higher Order**	**Total**	**%**
Topic 1	2	1	-	3	**15%**
Topic 2	2	3	3	8	**40%**
Topic 3	2	4	3	9	**45%**
Totals	**6**	**8**	**6**	**20**	**-**
Total %	**30%**	**40%**	**30%**	**-**	**100%**

Readers should note that the elements of the test plan should include the following:

i. Content topics (or the standards/strands to assess)
ii. Types of thinking skills/cognitive processes to assess
iii. Specific learning targets to assess
iv. The number of tasks (or points) for each standard/strand or learning target to be assessed

Further in deciding how many tasks or the proportion of marks to allocate to each element, the test constructor should consider the following five criteria:

i. The importance of each learning target
ii. The importance of the content
iii. The time available for the test
iv. The nature of the task
v. The emphasis you placed on them during your teaching

Readers should note that in Table 3.3 the number for each cell in the matrix is given. It is assumed that these are objective test items. However, sometimes the items are constructed response items. The numbers in the table can reflect the weightings or the number of marks. For the various items/tasks, some practitioners use the percentage of items or marks whilst others simply state the number of items/marks. Yet other practitioners provide both the numbers and the percentages in the tables. The following example is adapted from Linn & Miller (2005) p.144

Table 3.3: Table of Specifications – Arithmetic Test – Grade 5

Content Area	**Procedural Skills**	**Understanding**	**Application**	**Totals**
Simple Fractions	5	10	5	**20**
Mixed Fractions	5	15	10	**30**
Decimals	5	10	5	**20**
Decimal &Percentages	1 0	10	10	**30**
Totals	**25**	**45**	**30**	**100**

Tasks:

1. *Discuss the table above with a colleague and decide what changes, if any, you would make. Give reasons for your decision*

2. *Choose a different grade level, modify the table (content & numbers) to suit the grade level that you have chosen. Justify your choices.*

3. Discuss the table of specifications below (Table 3.4) and suggest alternative numbers for each cell. Give reasons for your choice.

4. Suppose you have to prepare an end –of- unit test on Decimals alone, provide a feasible table of specifications.

5. Study the Table of Specifications for Social Studies given in Table 3.4 below and adapt it to match the syllabus for your class.

Table3.4: Table of Specifications: Primary Social Studies National Assessment

Levels / Items/Topics	Type of Item	Knowledge / Recall	Application	Analysis & Evaluation	Totals
1. Culture & Customs	Matching	5	-	-	**5**
2. Historical sites	Fill in the Blanks	5	15	10	**30**
3. Organizations	Short answer	5	10	5	**20**
4. Tourism	Short answer	5	-	5	**10**
5. Map Work	Map work	-	10	-	**10**
6. Resources	Multiple choice	15	5	5	**25**
Total		**35**	**20**	**15**	**70**
%		**50%**	**30%**	**20%**	**100%**

3.4 Purpose, Guidelines and Steps for Preparing Test Blueprints

Purpose: There are several reasons why a test blue print is required

i. It ensures that topics and cognitive levels of learning are represented in the test.
ii. It is used to ensure congruence between instructional objectives and the test content of a classroom test, or to ensure congruence between the syllabus objectives and the test content in the case of an external high stakes test.
iii. To ensure balance in skills tested.
iv. To ensure that a representative sample of objectives are tested.
v. To facilitate the development of alternate test forms.

Guidelines to follow in planning the table of specification

i. Choose a cognitive framework that best represents your assessment philosophy.
ii. Ensure the numbers in each cell in the grid represents the relative emphasis of each content area.
iii. Ensure that the numbers in each cell reflect the relative emphasis of the cognitive levels of learning that you employed in your instructional activities.

3.5 Ten Steps in Preparing Test Blue Print or Table of Specifications:

i. Select the content to be covered in test (Topic Headings or Specific Objectives).
ii. Determine the cognitive framework (e.g. Bloom's Taxonomy) or list of skills/behaviours.
iii. Determine the item format or formats to be used.
iv. Determine the overall weight (with percentages) of items for each topic.
v. Determine the number and percentage of items for each standard and cognitive level
vi. Determine the number/percentage of items for each cell in the grid.
vii. Consider the importance of each topic/objective and the time spent teaching it.
viii. Have a colleague review your table of specifications, and give feedback.
ix. Study the feedback from colleagues.
x. Revise your first choice.

Planning End-of-Term/Year Examinations

When preparing term or annual examinations, various formats and modes of assessment would be necessary. One approach is to decide how each component should be constructed and then prepare a table of specifications for each component. An overall assessment plan must first be done. Table 3.5 below illustrates an example of a Test plan before the individual papers are prepared.

Table 3.5. Table of Specifications – End-of-Term Examination

Assessment Method	Intellectual Skills			Totals	%
	Knowledge/ Comprehension	Higher Order	Practical Component		
Paper 1 Multiple Choice	15	20	15	50	**25**
Paper 2 Constructed Response	30	50	20	100	**50**
Paper 3 Assignments /*	15	20	15	50	**25**
Totals	**60**	**90**	**50**	**200**	****
%	**30**	**45**	**25**		**100**

***These may include performance assessment tasks: projects/ investigations etc*.**

3. 6 Additional steps when preparing Unit tests, Term Tests and Annual Examinations

i. Since multiple modes of assessments are required for these tests, guidelines are needed for each assessment mode.
ii. The relationship between each mode and the overall assessment plan must be made clear.
iii. Some practitioners may go further and require sample tasks to amplify the assessment specifications.

You may use the following checklist to guide you in preparing your test blue print.

Table 3.6 Checklist for Reviewing Specifications for Assessments

		Yes	No
1	Are the specifications in harmony with the purpose of the test/assessment		
2	Do the specifications indicate the nature and limits of the achievement domain?		
3	Do the specifications indicate the types of learning outcomes to be measured?		
4	Do the specifications indicate the sample of learning outcomes to be measured?		
5	Is the number of test items of assessment tasks indicated for each subdivision?		
6	Are the types of items and tasks to be used appropriate for the outcomes to be measured?		
7	Is the distribution of items and tasks adequate for the types of interpretation to be made?		
8	If sample items and tasks are included, do they illustrate the desired attributes?		
9	Do the specifications as a whole indicate a representative sample of instructionally relevant tasks that fit the use of the results?		

3.7 Summary

In this chapter we studied various frameworks for organizing curriculum and assessment. We acknowledged that one can classify learning outcomes and test items in different ways. But it is important to choose a cognitive framework that is in keeping with an acceptable philosophy of teaching the subject area and one that is of suitable for the age level that one is teaching. This writer suggests that one can combine adjacent levels in some of these frameworks in order to make it easier for classroom application. For example, it is not always easy to distinguish

knowledge from comprehension in some cases and for test purposes they may be combined to form one level. Also, some experts differ on whether analysis must take place before one can compare, predict, make inferences and extrapolate. It might therefore be appropriate to classify all these skills as higher order thinking. In fact, for testing purposes; analysis, synthesis and evaluation may be combined and described as higher order thinking skills. However, it is suggested that teachers working in the same age group should use the same scheme.

CHAPTER 4

ASSESSMENT TOOLS PART 1

In this chapter we examine some test items and tasks that teachers use to measure pupil achievement. Classroom Assessment is a systematic approach to formative evaluation, used by teachers to determine how much and how well students are learning.

4.1 Classroom Assessment Techniques

Classroom Assessment Techniques (CATs) and other informal assessment tools provide key information during the academic year regarding teaching and learning so that changes can be made as necessary. CATs are simple tools (instruments, forms, strategies, activities) used for collecting information on student learning in order to improve it. CATs are easy to design, administer and analyze, and have the added benefit of involving students in their own learning.

CATs are divided into three broad categories:

- Techniques for assessing course-related knowledge and skills;
- Techniques for assessing learner attitudes, values and self-awareness; and
- Techniques for assessing learner reactions to instruction

Techniques for Assessing Course-Related Knowledge & Skills

Assessing Prior Knowledge, Recall, and Understanding - The CATs in this group are recommended to assess declarative learning, the content of a particular subject.

Background Knowledge Probe: short, simple questionnaires prepared by instructors for use at the beginning of a course or at the start of new units or topics; can serve as a pre-test.

Focused Listing: focuses students' attention on a single important term, name, or concept from a lesson or class session and directs students to list ideas related to the "focus."

Empty Outlines: in a limited amount of time students complete an empty or partially completed outline of an in-class presentation or homework assignment.

Minute Paper: perhaps the most frequently used CAT; students answer 2 questions (What was the most important thing you learned during this class? And What important question remains unanswered?).

Muddiest Point: - Technique consists of asking students to jot down a quick response to one question: "What was the muddiest point in _____?" with the focus on the lesson, a discussion, a homework assignment, a play, or a video. Used to provide information on what students find least clear or most confusing about a particular lesson or topic.

I. ***Assessing Skill in Analysis and Critical Thinking*** **-** The CATs in this group focus on analysis—the breaking down of information, questions, or problems to facilitate understanding and problem solving.

Pro and Con Grid: pupils list pros/cons, costs/benefits, advantages/disadvantages of an issue, question or value of competing claims.

Content, Form, and Function Outlines: in an outline form, students analyse the "what" (content), "how" (form), and "why" (function) of a particular message (e.g. poem, newspaper story, advertisement, story).

II. ***Assessing Skill in Synthesis and Creative Thinking*** -The CATs in this group focus on synthesis—each stimulate the student to create, and allow teachers to assess original intellectual products that result from a synthesis of course content and the pupils' intelligence, judgment, knowledge, and skills.

One-Sentence Summary: pupils answer the questions "Who does what to whom, when, where, how, and why?" (WDWWWWHW) about a given topic and then creates a single informative, grammatical, and long summary sentence.

Word Journal: involves a 2 part response; 1st the student summarizes a short text in a single word and 2nd, the student writes 1-2 paragraphs explaining the word choice.

Concept Maps: pupils draw or diagram the mental connections they make between a major concept and other concepts they have learned.

III. ***Assessing Skill in Application and Performance*** - The CATs in this group focus on students' abilities to apply important—sometimes referenced as conditional knowledge—knowing when and where to apply what know and can do.

Directed Paraphrasing: pupils paraphrase part of a lesson for a specific audience demonstrating ability to translate highly specialized information into language the clients or customers can understand.

Student-Generated Test Questions: pupils generate test questions and model answers for critical areas of learning.

Double-Entry Journal: using a T-chart, students will reserve one side for elements of the text that stood out to them, while the opposite side will be the explanation, analysis, and possible application of this portion of text.

Techniques for Assessing Learner Attitudes, Values, and Self-Awareness

IV. **Assessing Students' Awareness of Their Attitudes and Values** - The CATs in this group are designed to assist teachers in developing students' attitudes, opinions, values, and self-awareness within the course curriculum.

Classroom Opinion Polls: Students indicate degree of agreement or disagreement with a statement or prompt.

Double-entry Journals: Students record and respond to significant passages of text.

Everyday Ethical Dilemma: Students respond to a case study that poses a discipline-related ethical dilemma.

V. **Assessing Students' Self-Awareness as Learners** - The CATs in this group are recommended to help pupils express personal goals and clarify self-concept in order to make a connection between the articulated goals and those of the course.

Interest/Knowledge/Skills Checklists: Students complete a checklist survey to indicate their knowledge, skills and interest in various course topics.

Self-Assessment Ways of Learning: Students compare themselves with several different "learning styles" profiles to find the most likely match.

VI. **Assessing Course-Related Learning and Study Skills, Strategies, and Behaviours** - The CATs in this group focus both student and teacher attention on the behaviours the student actually engages in when trying to learn.

Study-Time Logs: Students complete a study log to record the quantity and quality of time spent studying for a specific course.

Process Analysis: Students outline the process they take in completing a specified assignment.

Diagnostic Learning Logs: Students write to learn by identifying, diagnosing, and prescribing solutions to their own learning problems.

Techniques for Assessing Learner Reactions to Instruction

VII. Assessing Learner Reactions to Teachers and Teaching - The CATs in this group are designed to provide context-specific feedback that can improve teaching within a particular topic/unit/subject.

Teacher-designed Feedback Forms: Students respond to specific questions through a focused feedback form about the effectiveness of a particular class session.

Group Instructional Feedback Technique: pupils respond to three questions related to the pupil's learning in the unit/subject.

VIII. **Assessing Learner Reactions to Class Activities, Assignments, and Materials -** The CATs in this group are designed to give teachers information that will help them improve their course materials and assignments.

RSQC2 (Recall, Summarize, Question, Connect and Comment): Students write brief statements that recall, summarize, question, connect and comment on meaningful points from previous class.

Group-Work Evaluation: Students complete a brief survey about how their group is functioning and make suggestions for improving the group process.

4.2 Multiple Choice Varieties

We turn our focus to selected–response type items. Selected-response items require pupils to choose an answer from a set of alternative choices. Selected-response items include multiple-choice items and alternative response type items. We will also study some supply type items such as short answer items, completion type items, context dependent item sets and interpretive exercises. Next, we take a brief look at some exercises used in English Language.

The standard multiple-choice item

Multiple-choice items may be classified as the standard multiple-choice item type, the matching type, classification type and the multiple-selection or combined response type.

The standard multiple-choice item consists of two parts: a stem and a set of options from which pupils must select the correct answer. The stem presents the problem situation or the question to be answered. The options include the key (the correct answer) and the distracters, which are plausible options that pupils with little or no understanding of the problem are likely to choose.

Format: The standard multiple-choice item may be presented as an *incomplete statement* or as a *question*. The examples below illustrate these two formats.

Example 1

1(a) Question Format	1(b) Incomplete Statement Format
Which island in the Lesser Antilles is the most easterly?	The most easterly island of the Lesser Antilles is
(A) Anguilla	(A) Anguilla
(B) Barbados	(B) Barbados
(C) St. Lucia	(C) St Lucia
(D) Tobago	(D) Tobago

In Examples 1(a) and 1(b) above, option B is the key; Options A, C and D are the distracters.

Some writers believe that it is often possible to use fewer words using the incomplete format than when using the question format. Ebel (1979), however, suggests that the question format enables the test constructor to present the problem in more specific terms. He also argues that the direct question helps the writer to focus more clearly on the purpose of the item and to avoid using irrelevant words and unrelated distracters. In the final analysis, the choice is that of the item writer, bearing in mind that specificity and clarity of the stem are important.

Constructing Multiple-Choice Items

Before introducing other varieties of multiple-choice items we need to examine some important principles that should be observed when constructing multiple-choice items. In the interest of space, the discussion is confined to a few of the principles.

1. The stem. The stem should be as clear and yet as precise as possible. It should present the problem to be solved or the concept or principle being assessed. Teachers should avoid indirect stems that do not indicate what is being tested.

Example 2a: Harrison's Cave is	**Example 2b:** Nelson dockyard is in
(A) in a gully	(A) St. George
(B) on a plantation	(B) St. John
(C) in St. Thomas	(C) St. Paul
(D) in St. George	(D) St. Peter

Presumably the teacher wants to test whether the children know the parish in which the cave is situated. The question would be better put as follows:

Harrison's Cave is situated in the parish of	In which parish is the Nelson dockyard situated?
(A) St. George	(A) St. George
(B) St. John	(B) St. John
(C) St. Michael	(C) St. Paul
(D) St. Thomas	(D) St. Peter

***2. Arrangement of options*.** Options should be arranged in some kind of logical order. For example, we can use ascending or descending order of magnitude if the options are numbers. If the options are words, we can use alphabetical order as in Example 3 above. In other cases, we can use sequential order, historical order or chronological order.

***3. Number of options*.** In most classroom tests three, four or five options are used. Three options are preferable for younger children as it reduces the reading time and the complexity of the item. It is not always easy to get a good fifth option.

4. Notes on Distracters

(a) All distracters should be plausible.
(b) Use common errors that pupils often make or errors they are likely to make if they read the question carelessly or do the calculations carelessly.
(c) Make distracters similar to the key in terms of length, structure and style.

4.3 Matching items

The matching format consists of three important points:

(a) A series of stems called 'premises'
(b) A series of answers or responses
(c) A set of directions for matching the premises and responses.

Example 4: Directions: Match the items in Column A with the responses in Column B by writing the letter in front of the appropriate response in column B, in the blank space provided before each item in column A.

Column A	*Column B*
____(1) Antigua	(A) British Virgin Islands
____(2) Dominica	(B) Leeward Islands
____(3) Grenada	(C) Windward Islands
____(4) St Kitts	(D) United States Virgin Islands
____(5) St Lucia	
____(6) St Thomas	
____(7) Tortola	

In the example above, we write 'B' in the first space in Column A, since Antigua is in the Leeward Islands. We write 'A' in the last blank space beside (No. 7), since Tortola belongs to the British Virgin Islands. We also write "A" beside No 4, since St. Kitts is a Leeward Island. Note also that in the example above, there are more *premises* than *responses*. This is an important principle to observe when constructing matching type items. In order to reduce guessing the number of premises should be more than the number of responses.

A good matching item functions in the same way as a series of multiple-choice items. As each premise is considered, all the options (responses) should serve as plausible distracters. Matching items may be presented in different formats. The format used in Example 4 above is used with younger children, whereas, the format used in Example 5 below, is often used in examinations for older children and adults.

Example 5*: Directions. The group of items below consists of four lettered headings followed by a list of numbered statements. For each numbered statement, select one heading that is closely related to it. Each heading may be used once, more than once, or not at all.*

Headings

(A) Hexagon

(B) Parallelogram

(C) Square

(D) Trapezium

Statements

(1) Only one pair of opposite sides is parallel.

(ii) Both pairs of opposite sides are parallel.

(iii) All four sides are equal.

(iv). This polygon has six sides.

(v) The diagonals of this figure are equal.

The matching item format is useful for testing a wide variety of concepts and principles. It can be used to test pupils' understanding of homogeneous concepts and to find out if they can discriminate between closely related principles or ideas. Examples could include matching events with dates, terms with definitions, formulae with quantities and so forth.

4.4 Comparison between matching items and classification items

Closely related to matching items are classification items. Classification items require pupils to sort a set of objects or ideas according to some well-defined principle: for example, classes of animals or plants, cause – effect sequences, branches of government and so forth. Ebel (1974) points out that the main distinction between matching and classification items, is that in the classification items there is a one-to-one matching, whereas in the matching type format there are more premises than responses.

Guidelines for writing matching and classification items

1. Ensure that clear directions are provided. The directions should specify how the items are to be matched, how the answers are to be marked or given, as well as the number of times each option may be used.
2. Place all items, options and directions on the same page of the test.
3. Use lists or descriptors for premises that are relatively homogeneous.
4. The list of premises should be longer than the list of options in order to discourage random guessing.
5. As a general rule, use four to eight premises.
6. Arrange the options, and if possible the premises, in some kind of logical order (e.g. alphabetical or numerical).
7. If responses are numerals, place them in order form smallest to largest.
8. Keep responses as short as possible.

4.5 Short Answer and Completion Items

Both the short answer item and the completion item are supply-type items that require a short response, usually in the form of a few words or numbers or combination of both. A short-answer item is one that can be answered by a word, a few words, a phrase, a statement or even a number. Some writers, for example, Worthen, White, Fan and Sudweeks (1999) treat short-answer items and completion items as variants of the same type. Other writers (Carey, 2001; Sax, 1997; Trice, 2001) prefer to make a distinction between short-answer and completion items. Sax makes the following distinction between the two. The completion item requires a pupil to complete a sentence with a word or phrase, whereas the short-answer item poses a question that can be answered by a word or phrase.

The distinction between the short answer and an essay item can often be blurred. Even experts do not all agree on exactly what the difference is. Gallagher (1998) views a short-answer item as a short essay item requiring a brief response which could be a word, a phrase, a sentence or a short-list. Trice (2001) views a short-answer item as "a supply item that does not require organization" (p. 121). We can contrast the two terms by saying that an essay requires quite a bit of structure and organization for an acceptable response; whereas the short answer item might require a short listing of information, such as giving a definition or describing something in a word, a phrase, or in one or two sentences.

4.6 Context Dependent Item Sets

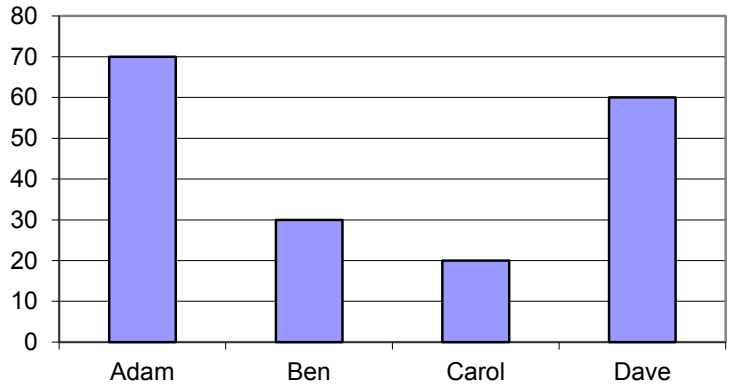
Context-dependent item sets consist of a series of test items based on common introductory material. The introductory material may be a paragraph, a picture, a graph, a chart or map. The items may consist of a series of selected- response items, completion items or short answer items. One may even ask more elaborate questions that require short essays using this format. The examples below will illustrate just the short answer format.

Guidelines for Constructing Context Dependent Item Sets

1. The material selected should relate to the intended learning outcomes.
2. It is advisable to use material that is new to the pupils. If one uses material that the examinees have seen before, then memory could influence their performance.
3. Keep the material as brief as possible.
4. Write questions to assess higher order as well as lower order their skills.
5. Ensure that information from one question does not provide clues for the response to another.
6. Write appropriate directions for each set of items.
7. Prepare the scoring key/rubric.
8. Have a colleague, review or critique the items and the scoring rubrics.
9. Revise the items and scoring rubrics based on comments received.

Example 1.

The graph below shows the scores obtained by four candidates on a test. Use the graph to answer items (i) – (iii).



(i) Calculate a median score.

(ii) Calculate the mean score

(iii) How many more marks than Carol did Adam score?

Table 4.6 Advantages and disadvantages of context-dependent item sets.

Advantages	Disadvantages
Very versatile; can be used to assess from low order to high order thinking	Difficult to ensure that high level thinking skills are tested
Can be used for several items	Time consuming to construct
Can be used both for selected response and constructed response formats	Requires a lot of reading. Hence reading ability may influence what is really being tested

4.7 Questioning Techniques

- Questioning has a variety of purposes in the classroom, including
- Actively involving students in the lesson
- Increasing motivation or interest
- Evaluating pupils' level of preparation
- Evaluating pupils' level of understanding
- Developing critical thinking skills
- Creating interest in the topic or lesson
- Assessing pupils' achievement or mastery of goals and objectives
- Stimulating independent learning

The questioning techniques utilised in instruction should reflect the key aspects of both the cognitive and affective domains and should be integrated with active learning strategies.

Examples of the question stems that can be incorporated in assessing pupils' performance in the cognitive and affective domains are shown in the tables below.

Table 4.7 Question Stems Based on the Cognitive Taxonomy

Taxonomy Level	Question Stems
REMEMBER (KNOWLEDGE) - drawing out factual answers, testing recall and recognition	Who? Where? Which One? What? How? What is the best one? Why? How much? When? What does It mean?
UNDERSTAND (COMPREHENSION) - translating, interpreting and extrapolating	Which are facts? What does this mean? Is this the same as. . .? Give an example. Select the best definition. Condense this paragraph What would happen if . . .? State in one word . . . Explain what is happening. What part doesn't fit? Explain what is meant.
APPLY - Knowing when to apply; why to apply; and recognizing patterns of transfer to situations that are new or unfamiliar	Predict what would happen if Choose the best statements that apply Judge the effects

Taxonomy Level	Question Stems
	What would result Tell what would happen Tell how, when, where, why
ANALYZE - breaking down into parts, forms	What is the function of . . .? What's fact? Opinion? What assumptions. . .? What conclusions? What does the author believe?
EVALUATE - according to some set of criteria, and state why	Which is more important, moral, better, logical, valid, appropriate? Find the errors
CREATE (SYNTHESIS) - combining elements into a pattern not clearly there before	How would you test. . .? Propose an alternative. Solve the following. How else would you . . .? State a rule

Table 4.8 Question Stems Based on the Affective Taxonomy

Taxonomy Level	Question Stems
RECEIVE – willing to listen and open to new experiences	How do you perceive …..? Discuss how your emotions when…? How would you feel if….?
RESPOND – actively responding to an activity	Can you clarify what….? How would you react if…? How best can you assist your group/team in ……?
VALUE – attaching value to something and expressing personal opinions	How would you justify…..? Discuss how you would deal with….? How would you persuade….?
ORGANIZE – to express personal views, beliefs, or opinions	Discuss how this would impact….? How would you defend…..?
INTERNALIZE VALUES (CHARACTERIZATION) – acting according to one's personal beliefs and values	How would you resolve….? What would influence you to….? What is the best way to carry out…?

4.8 Summary

There are many tools that can be used to assess pupil achievement. However, it is important to know which tool is best to use for a given situation. Just as a carpenter needs to know when to use a hammer, a chisel or a saw, so too the classroom teacher needs to know which tool is best suited for measuring pupil learning in various situations. Each type has advantages and disadvantages. Hence whenever a teacher selects a multiple- choice format, an alternative response type format, a short answer or context- dependent item sets, the teacher should ensure that the tool selected is the most appropriate one for the given purpose.

CHAPTER 5

ASSESSMENT TOOLS - PART 2

5.0 Language Arts

The item formats discussed so far can be used in almost any subject. However, Language Arts require a variety of formats to assess a wide range of skills. Some of these are illustrated below. The first two examples illustrate some exercises suitable for very young children (Grade 2). The other examples are selected from Grade 4 and Grade 6 examinations. The author is grateful to the Ministry of Education in Antigua for permission to use these items for illustrative purposes. [The example was selected from a manual prepared by Halliday (2008) for Ministry of Education, Antigua & Barbuda.]

5.1 Word Recognition -*This item is appropriate for Grade 2*

Circle the word that matches each picture.

	star	stem	stare
	snack	snail	snake
	flag	flap	flower
	plum	pear	pail
	church	school	shop