



ORGANIZATION OF EASTERN CARIBBEAN STATES  
EDUCATION SECTOR STRATEGY



OECS PRIMARY GRADES' ASSESSMENT MANUAL



© 2018





## TABLE OF CONTENTS

CHAPTER 1 .....	7
IMPORTANT CONCEPTS.....	7
1.1 Purpose of Manual .....	7
1.2 Explanation of important concepts in this manual.....	8
1.3 The forms of assessment.....	11
Pre-instruction Assessment.....	11
Formative Assessment: Assessment during Instruction .....	11
Internal Assessment.....	12
Summative Assessment: Assessment after Instruction .....	12
Continuous Assessment.....	13
Principles governing good assessment practices .....	13
CHAPTER 2 .....	17
FRAMEWORKS .....	17
2.1 Bloom's Taxonomy .....	17
Knowledge.....	19
Comprehension.....	19
Application .....	19
Analysis .....	19
Synthesis.....	19
Evaluation.....	20
2.2 General comments.....	20
2.3. Revisions to Bloom's Taxonomy.....	21
2.4 The Knowledge Dimension of the Revised Taxonomy.....	22
2.5 Using the Revised Taxonomy to Classify Learning Outcomes and Test Items.....	25
2.6 Test items .....	26
2.7 Planning Tests Using the Revised Taxonomy .....	28
2.8 The Affective Domain .....	31
CHAPTER 3 .....	32
TEST SPECIFICATIONS/ASSESSMENT BLUEPRINT .....	32
3.1 Introduction.....	32

3.2	Improved Definition for Table of Specification.....	33
3.4	Purpose, Guidelines and Steps for Preparing Test Blueprints .....	36
3.5	Ten Steps in Preparing Test Blue Print or Table of Specifications: .....	37
	Planning End-of-Term/Year Examinations .....	37
3.6	Additional steps when preparing Unit tests, Term Tests and Annual Examinations ..	38
3.7	Summary .....	39
CHAPTER 4 .....		41
ASSESSMENT TOOLS PART 1 .....		41
4.1	Classroom Assessment Techniques .....	41
	Techniques for Assessing Course-Related Knowledge & Skills.....	41
	Techniques for Assessing Learner Attitudes, Values, and Self-Awareness.....	43
	Techniques for Assessing Learner Reactions to Instruction .....	44
4.2	Multiple Choice Varieties .....	44
	The standard multiple-choice item .....	45
	Constructing Multiple-Choice Items .....	46
4.3	Matching items.....	47
4.4	Comparison between matching items and classification items.....	49
	Guidelines for writing matching and classification items .....	49
4.5	Short Answer and Completion Items .....	49
4.6	Context Dependent Item Sets.....	50
	Guidelines for Constructing Context Dependent Item Sets.....	50
4.7	Questioning Techniques.....	52
	Find the errors.....	53
4.8	Summary .....	54
CHAPTER 5 .....		55
ASSESSMENT TOOLS - PART 2.....		55
5.0	Language Arts .....	55
5.1	Word Recognition -This item is appropriate for Grade 2.....	55
5.3	Reading Comprehension: .....	56
5.4	Grammar & Mechanics .....	57
CHAPTER 6 .....		60
ASSESSMENT TOOLS AND TECHNIQUES: PERFORMANCE ASSESSMENT .....		60

6.0.	Introduction: Performance assessment.....	60
6.1.	Restricted versus Extended Performance Tasks .....	60
6.3	Process versus Product .....	64
	Guidelines to assist the teacher in deciding whether to assess process, product or both .....	64
6.4	Authentic Assessment.....	64
6.5	Guidelines for Developing Performance/Authentic Tasks .....	65
	Additional considerations .....	65
6.6	Projects.....	66
6.6.1.	Characteristics of a Good Pupil Project.....	66
6.7	Cooperative Learning/ Group Projects.....	68
	Some strategies for assessing group work.....	69
CHAPTER 7 .....		71
SCORING RUBRICS .....		71
	Introduction.....	71
	Guidelines for Designing Rubrics .....	71
	Number of points on the scale .....	72
CHAPTER 8 .....		79
PROTFOLIO ASSESSMENT .....		79
8.1	Types of portfolios .....	79
	All-inclusive Portfolio .....	79
	Developmental Portfolio.....	79
	Showcase Portfolio .....	80
	Evaluation Portfolio.....	80
8.2	Contents of Portfolios .....	80
8.3	Eight Steps in Designing a Portfolio Plan.....	82
8.4	Important Questions to Ask When Designing a Portfolio System .....	82
CHAPTER 9 .....		84
THREE IMPORTANT PSYCHOMETRIC ISSUES .....		84
9.1	Fairness.....	84
	Key components of fairness .....	85
9.2	Validity.....	85
9.3	Reliability.....	86

Factors that affect the reliability of assessments .....	87
CHAPTER 10 .....	88
Caribbean Primary Exit Assessment.....	88
GLOSSARY .....	92
<i>Affective domain</i> .....	92
<i>Assessment</i> .....	92
BIBLIOGRAPHY.....	99
APPENDIX.....	102
Summary of the Quellmalz Taxonomy .....	103
Thinking Skills for Different Subject Matter in Quellmalz Taxonomy .....	104
The Core Thinking Skills .....	105
An Outline of Bloom’s Taxonomy of Educational Objectives: Cognitive Domain.....	106
Examples of Standards Aligned to Bloom’s Taxonomy Levels.....	108
Barrett’s Taxonomy.....	109

## **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.

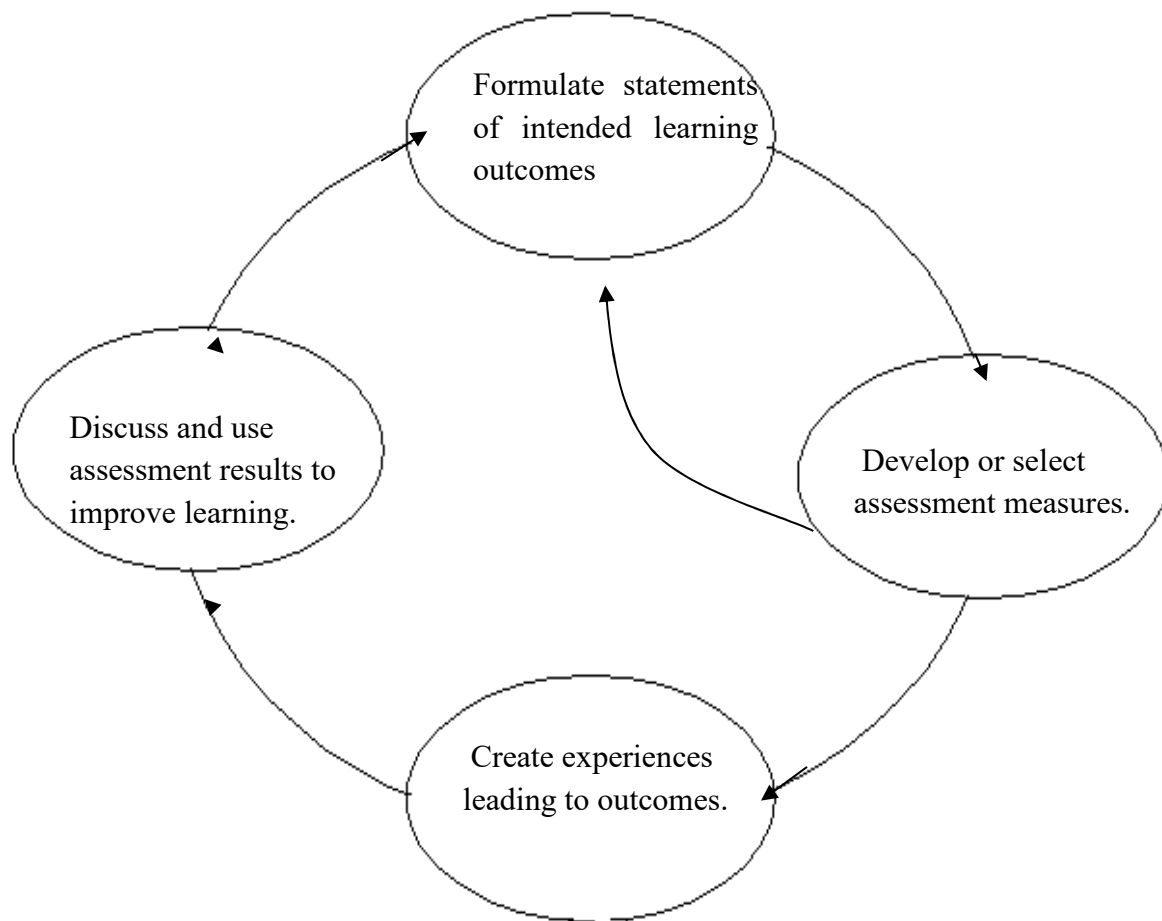
## **Assessment**

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 than 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 ASSESSMENT of LEARNING, for LEARNING, and as LEARNING**

<b>Assessment OF Learning</b>	<b>Assessment FOR Learning</b>	<b>Assessment AS Learning</b>
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.

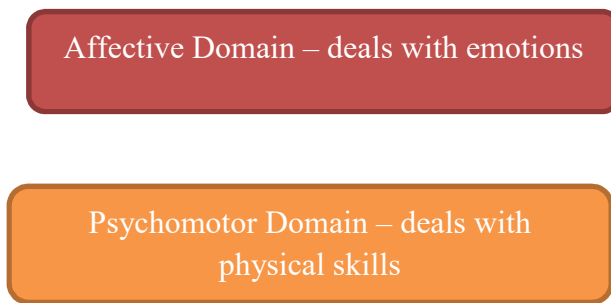
Several educators have proposed various lists of these skills that are useful for classroom instruction and assessment. Some of these are described as taxonomies, whilst others may be referred to as classification schemes or frameworks for organizing teaching and assessment activities. Three of these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

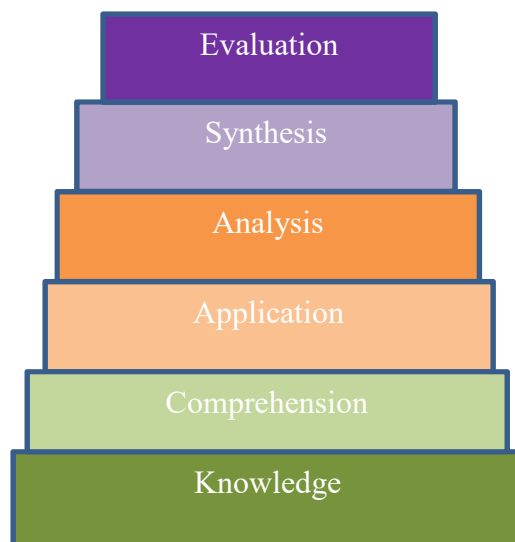
Bloom's Taxonomy divides the way a person learns into three domains:

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
<b>1.0 FACTUAL KNOWLEDGE</b> – The basic elements pupils must 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
<b>2.0 CONCEPTUAL KNOWLEDGE</b> – 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
<b>3.0 PROCEDURAL KNOWLEDGE</b> – 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
<b>4.0 METACOGNITIVE KNOWLEDGE</b> – 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	<b>REMEMBER</b> – 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	<b>UNDERSTAND</b> – 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	<b>APPLY</b> – 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	<b>ANALYZE</b> – 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	<b>EVALUATE</b> – 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	<b>CREATE</b> – 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 is 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 is 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**

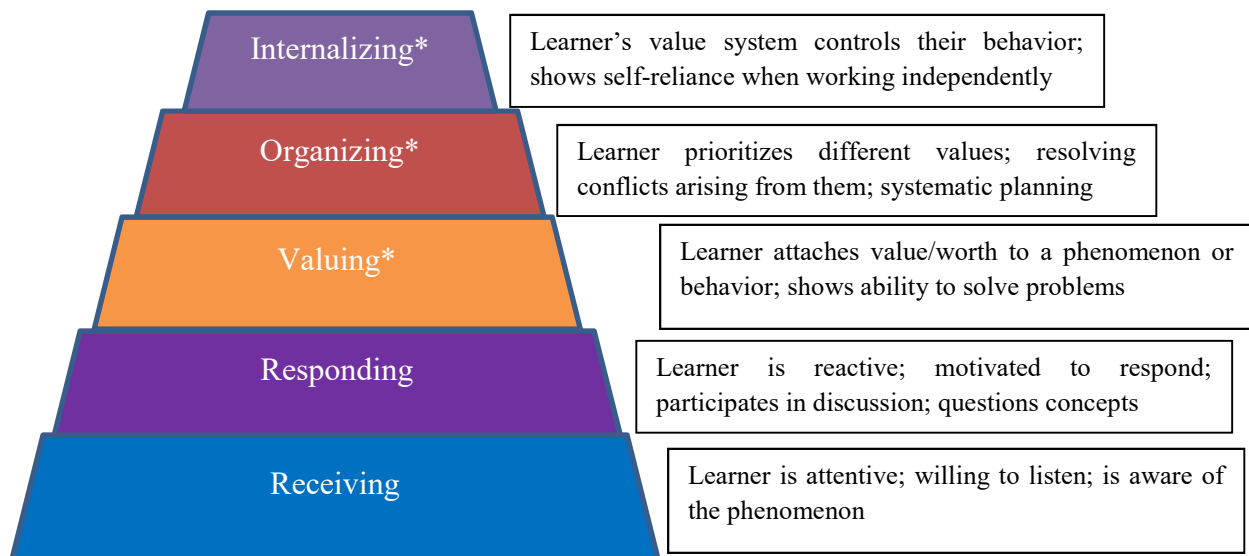
*Using the standards in Example 2 and the suggested framework at Table 2.7,*

- (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 blueprints 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**

<b>Content</b>	<b>Recall &amp; Comprehension</b>	<b>Application</b>	<b>Higher Order</b>	<b>Total</b>	<b>%</b>
Topic 1	2	1	-	3	<b>15%</b>
Topic 2	2	3	3	8	<b>40%</b>
Topic 3	2	4	3	9	<b>45%</b>
<b>Totals</b>	<b>6</b>	<b>8</b>	<b>6</b>	<b>20</b>	<b>-</b>
<b>Total %</b>	<b>30%</b>	<b>40%</b>	<b>30%</b>	<b>-</b>	<b>100%</b>

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**

<b>Content Area</b>	<b>Procedural Skills</b>	<b>Understanding</b>	<b>Application</b>	<b>Totals</b>
Simple Fractions	5	10	5	<b>20</b>
Mixed Fractions	5	15	10	<b>30</b>
Decimals	5	10	5	<b>20</b>
Decimal & Percentages	10	10	10	<b>30</b>
<b>Totals</b>	<b>25</b>	<b>45</b>	<b>30</b>	<b>100</b>

**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**

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

### 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	<b>25</b>
Paper 2 Constructed Response	30	50	20	100	<b>50</b>
Paper 3 Assignments /*	15	20	15	50	<b>25</b>
<b>Totals</b>	<b>60</b>	<b>90</b>	<b>50</b>	<b>200</b>	<b>\</b>
<b>%</b>	<b>30</b>	<b>45</b>	<b>25</b>		<b>100</b>

*\*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 2<sup>nd</sup>, 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.

<b>Example 2a:</b> Harrison's Cave is	<b>Example 2b:</b> 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.

<i>Column A</i>	<i>Column B</i>
____(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 from 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**

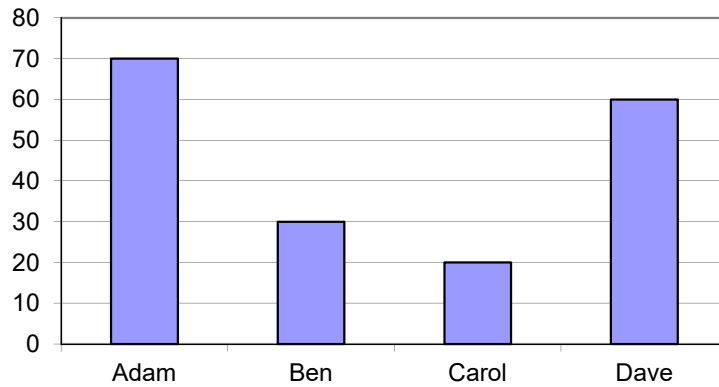
Context-dependent item sets consist of a series of test items based on common introductory material. The introductory material may be a paragraph, a picture, a graph, a chart or map. The items may consist of a series of selected-response items, completion items or short answer items. One may even ask 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**

<b>Taxonomy Level</b>	<b>Question Stems</b>
<b>REMEMBER (KNOWLEDGE)</b> - 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?
<b>UNDERSTAND (COMPREHENSION)</b> - 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.
<b>APPLY</b> - 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
<b>ANALYZE</b> - breaking down into parts, forms	What is the function of . . .? What's fact? Opinion? What assumptions. . .? What conclusions? What does the author believe?
<b>EVALUATE</b> - according to some set of criteria, and state why	Which is more important, moral, better, logical, valid, appropriate? Find the errors
<b>CREATE (SYNTHESIS)</b> - 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
<b>RECEIVE</b> – willing to listen and open to new experiences	How do you perceive .....? Discuss how your emotions when...? How would you feel if....?
<b>RESPOND</b> – actively responding to an activity	Can you clarify what....? How would you react if...? How best can you assist your group/team in .....?
<b>VALUE</b> – attaching value to something and expressing personal opinions	How would you justify.....? Discuss how you would deal with....? How would you persuade....?
<b>ORGANIZE</b> – to express personal views, beliefs, or opinions	Discuss how this would impact....? How would you defend.....?
<b>INTERNALIZE VALUES (CHARACTERIZATION)</b> – 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

### 5.3 Reading Comprehension:

Read the story very carefully. Circle the letters beside the best answers for the questions.

It was the summer holidays. Kim and his sister Ali enjoyed long hours of sleep in the mornings. They watched television, went to music classes weekly and played all kinds of indoor and outdoor games together. They were playing a game when the phone rang. Mom answered. She was very happy after speaking on the phone. Soon she was busy baking pies and singing loudly. She had not seen Aunt Sue for a long time. Aunt Sue went to live in Canada when Mom was a child. Kim and Ali had never seen their Mom so happy.

The next morning, the phone rang again. Before we could reach the phone, Mom answered.

"I am so sorry," she said, as she shook her head. After the call Mom looked sad.

1. The story took place during
  - (A) the Easter holidays
  - (B) the Christmas holidays
  - (C) the Summer holidays
2. Kim and his sister Ali enjoyed long hours of sleep in the mornings. This means that
  - (A) they woke up late in the afternoons.
  - (B) they sleep all day.
  - (C) they woke up later than usual in the mornings.
3. How often did the children go to music classes?
  - (A) Every day
  - (B) Every week
  - (C) Every month
4. Two words in line 3 that are opposite in meaning are
  - (A) all and together
  - (B) played and games
  - (C) indoor and outdoor



## 5.4 Grammar & Mechanics

1. Write the part of speech for each of the words in italics BELOW. Use the lines provided.

1. She sings extremely *well*. \_\_\_\_\_
2. Tamara *set* the table while Georgette prepared the breakfast. \_\_\_\_\_
3. The thief disappeared as the *police* officer approached him. \_\_\_\_\_

2. Use the conjunctions below to correctly complete each of the sentences. Do not use any conjunction twice.

but      while      because      although      since      until      and

1. Mr. Broomes would not allow me back in class \_\_\_\_\_ I promise to behave.
2. We ran quickly to school \_\_\_\_\_ we were still late.
3. The game was cancelled \_\_\_\_\_ it was raining.
4. I will not take it \_\_\_\_\_ you want me to.

3. Rewrite the sentences putting the verbs underlined in the simple past tense.

1. Sam paints the house.
2. She will sit patiently for the bus to arrive.
3. I am going to town to buy a pair of shoes.
4. Nathan is throwing the ball to his friend.

4. Complete each sentence by inserting into the blank space the adjective formed from the noun given.

1. **danger**      The cliff is too \_\_\_\_\_ to climb.
2. **favour**      Mother always cooks my \_\_\_\_\_ meal on Sundays.
3. **greed**      The \_\_\_\_\_ dog chewed hungrily on the bone.
4. **wealth**      A \_\_\_\_\_ family lives in the mansion on the hill.

5. Complete each of the following by underlining the correct word from the brackets.

1. Hasn't (no one, anyone) see my pen.
2. Before we had (hanged, hung) the clothes on the line, it began to rain.
3. I have the (most prettiest, prettiest) dress.

6. Vocabulary

Circle the word that is OPPOSITE in meaning to the underlined word in each of the following sentences.

- i. This pupil always submits untidy work to his teacher.
- A. clean                      B. complete                      C. dirty                      D. late
- ii. Grandma drank the bitter bush tea.
- A. salty                      B. sweet                      C. strong                      D. sour

## 7. Homophones

Complete each sentence by underlining the correct word in the brackets.

1. Do not (wring, ring) the clothes before hanging them on the line.
2. The dog chased the (herd, heard) of cattle across the field.
3. Shawn tied the blue sash around her (waist, waste).
4. The Mathematics classes lasted for an (hour, our).

## 8. Plurals

Change each word in the brackets to its plural form. Write the plural in the space provided.

1. Put the \_\_\_\_\_ into the water. (potato)
2. We learnt about \_\_\_\_\_ in our science classes. (pulley)
3. Are several pupils waiting to play the \_\_\_\_\_? (piano)
4. The \_\_\_\_\_ were all dressed in blue. (maid-of-honour)

9. Circle the letter beside the word that means the same or almost the same as the underlined words.

- i. You have enough time to complete the examination.
- A. adequate      B. extra      C. little      D. much
- ii. The President pardons the prisoner.
- (A.) chastises                      (B). forgives                      (C). condemns

**1. Rewrite and punctuate the following passage on the lines below.**

Barnardo was startled by what he heard Could it be true that this ten year old boy really had no home Are there other boys like you Jim without a home asked Barnardo  
Oh yes lots of them was the reply  
Surely the boy must be telling lies thought Barnardo but suppose he is telling the truth.  
I will give you a meal and a bed for the night Jim Will you show me where these boys sleep

*Story by Geoffery Hanks*

---

---

---

**2. From the passage below, identify FIVE common nouns and FIVE proper nouns.**

Michael and Sharon crashed into David Smith's stall in the middle of Reeds Avenue with their football. Mr. Smith was angry at first but he forgave them, but said he would inform their parents of the incident. Sharon was afraid because her mother told her she could play with the ball, but only if they she was very careful. So Sharon and Michael went through the back gates of the park and sat on a bench so they could stay out of trouble.

Common Nouns	Proper Nouns
1.	1.
2.	2.
3.	3.
4.	4.
5.	5.

## CHAPTER 6

### ASSESSMENT TOOLS AND TECHNIQUES: PERFORMANCE ASSESSMENT

#### 6.0. Introduction: Performance assessment

This chapter deals with several types of performance based assessment strategies. Mc Millan (2004) describes performance assessment as “one in which the teacher observes and makes judgment about the pupil’s demonstration of a skill or competency in creating a product, constructing a response, or making a presentation” (p.198). He further stated that the term *performance* is shorthand for *performance based* or *performances-and-product*. Nitko (1996) described performance assessment as a procedure in which teachers use work assignments to obtain information about how well pupils have mastered the course material.

Performance assessment tasks require pupils to produce, construct or perform a task. The tasks may consist of practical activities such as painting a scene or baking a cake, or they may involve paper-and- pencil writing tasks. Performance assessment is an umbrella term used for a number of assessment modes in which the pupils are required to compose or construct their own response to a test item. It could involve practical work, oral work or written work. These assessment tasks require pupils to draw skills from a various aspects of the content domain, and to demonstrate higher order thinking skills, such as analysis, reasoning and problem solving.

#### 6.1. Restricted versus Extended Performance Tasks

Some writers (Gronlund, 2003; McMillan, 2004; Payne, 2003) make a distinction between restricted performance tasks and extended performance tasks. A restricted performance task is designed to measure a specific or limited amount of knowledge or a limited number of skills. Extended performance tasks on the other hand require numerous skills and concepts to perform the tasks. The distinction between restricted and extended tasks is arbitrary and can vary according to individuals. It is a matter of judgement. However, in general the time taken to complete a task and the number of concepts and skills required to do the task can be used as a guides in making the distinction.

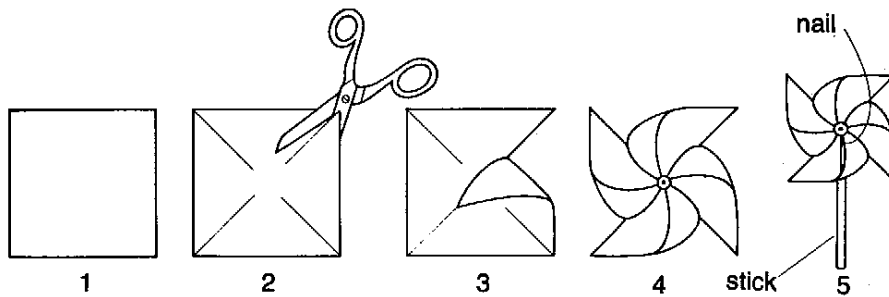
**Table 6.1 Restricted performance versus extended performance tasks**

<b>Restricted - Performance Task</b>	<b>Extended Performance Task.</b>
Construct a graph using given data	Paint a scene of a street carnival
Make one-minute speech on the importance of exercising frequently	Write an essay on "Christmas Morning in Queen's Park in Bridgetown".
Write a paragraph on a given topic	Design a stool for the science laboratory
Test the amount of chlorine in a given bottle of water	Repair a broken- down automobile

## 6.2 Science examples: Restricted performance assessment tasks

### Suitable for Grade 1 - 3

1. Make a windmill. Use a piece of card, a stick and a nail. The pictures will show you how to do it. Use the windmill to see how strong the wind is.



2. Test the following objects to see if they sink or float.

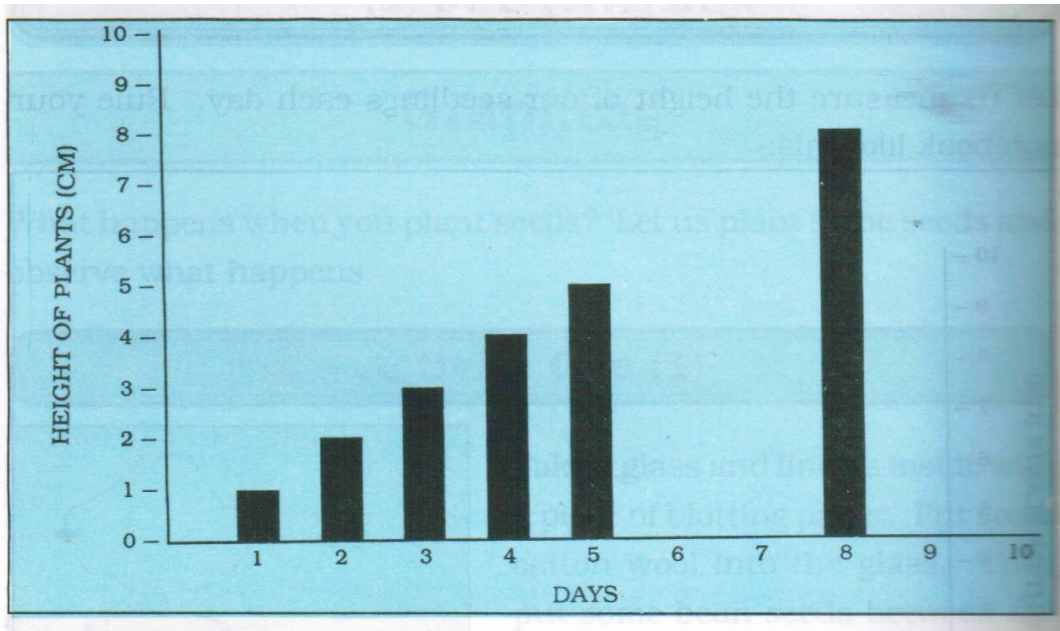
(a) Put each object into the bowl of water.

(b) Complete the table, using a tick (✓), show if an item sinks or floats.

Object	SINK	FLOAT
Sponge		
Pebble		
Marble		
Cotton reel		
Pencil		
Eraser		
Paper clip		
Chalk		
Bottle cap		

N.B. Suitable for Grades 4 to 6

3. The chart below shows the growth of a seedling. The height of the seedling was not recorded on some days. Answer the following questions.



What was the height of the plant on Day 3? \_\_\_\_\_

What do you think was the height of the plant on Day 6? \_\_\_\_\_

What do you think was the height of the plant on Day 7? \_\_\_\_\_

What do you think will be the height of the plant on the Day 10? \_\_\_\_\_

*Adapted from Thomas, M. (1995) Caribbean Science Series. Trinidad. Charran Publishing House*

### 6.3 Process versus Product

Performance based assessments may require pupils to demonstrate a sequence of skills or require pupils to focus on the outcome or product of an activity. Hence when we are dealing with performance based tasks we need to be clear about what we are going to assess, whether the process or the product or both. This is of particular importance when tasks such as role playing, simulation exercises, laboratory work, physical activities, speaking engagements, music, painting and so forth are being assessed.

#### Products

Two kinds of products can be assessed: Written products and Physical products.

Examples of **written products** are term papers, essays, journals and laboratory reports. Examples of **physical products** are industrial art projects, paintings, sculpture and clothing.

#### *Guidelines to assist the teacher in deciding whether to assess process, product or both*

- (i) When the steps involved in making the product can be retrieved, focus on the product.
- (ii) When the important characteristics of the product can be measured accurately or objectively, focus on the product.
- (iii) When it is not possible to assess the process, focus on the product.
- (iv) When the process is extremely critical, then the process should be assessed (e.g. patient care and hospital simulations).
- (v) When both process and product are important focus on both.

### 6.4 Authentic Assessment

Not all performance tasks are authentic. Authentic refers to the context in which the response to the task is performed (Gallagher, 1998). Authentic tasks should be relevant to the curriculum activities and as much as possible to real life situations. Wiggins (1983) suggests that “Authentic tasks are either replicas of or analogous to the kinds of problems faced by adult citizens and consumers or professionals in the field” (p.229). The major difference between performance assessment tasks and authentic assessment tasks is the extent to which the task approaches a real life task. Table 6.2 below illustrates the difference between performance tasks and authentic tasks.



**Table 6.2 Performance and Authentic Tasks**

<b>Performance task</b>	<b>Authentic Task</b>
Write an essay on “How we can improve our school”	Identify one major problem in your school that the Headmaster and staff should resolve urgently. Explain how the problem can be solved. Use whatever method of communication you think is best suited to the situation.
Discuss the qualities of a good citizen	You are the elected representative of your town. Write a speech you will deliver in parliament, outlining your plans to improve the social services in your community.

### **6.5 Guidelines for Developing Performance/Authentic Tasks**

Integrate the most essential aspects of the content with the most essential skills.

1. Place the task in a real-world context as far as it is possible to do so.
2. Structure the task to assess multiple learning targets.
3. Structure the task to help pupils to succeed.
4. Ensure that the task is feasible within the constraints given.
5. Construct tasks that allow for multiple solutions.
6. Ensure that the task is clearly specified.
7. Make the task challenging and stimulating to pupils.
8. Include explicitly stated scoring criteria as part of the task.

(Adapted from Mc Millan,1997)

#### ***Additional considerations***

It must be stressed that the tasks must be matched to the learning outcomes stated in the syllabus. Tasks should not only emphasize cognitive skills, but psychomotor skills and should help to develop aspects of the affective domain as well.

1. Tasks are to be based on critical learning outcomes for the subject.
2. Strategies need to be put in place to check that the pupil does his or her own work.

3. The project should be limited to the available resources.
4. The standards required should be specified clearly.
5. Scoring rubrics should be provided
6. The specific dimensions on which the project will be evaluated should be specified.
7. The weights (marks) for each dimension should be clearly stated.
8. All deadlines should be clearly specified.
9. Strategies to discourage plagiarism (e.g. periodic interviews/ oral presentations) should be put in place.
10. Where possible provide exemplars for pupil to study.

## **6.6 Projects**

A project is an extended performance task. Nitko (2004) defines a project as ‘a long-term activity that results in a pupil product’ (p.523). He suggested that the product may be a model, a functional object, a substantial report or a collection. Deere (1974) argued that a project is an activity that knits together various packets of materials; it comprises various kinds of tasks and hence integrates several skills in the process. He recognised that projects vary considerably, but concluded that, in general, a project is “a teaching/learning activity which requires the pupil to determine one or more of the following: his strategy, his resources, his target; which presents a task which is not artificially compartmentalised or idealised; which allows a range of solutions rather than a unique answer” (p.106)

### **6.6.1. Characteristics of a Good Pupil Project**

1. It focuses on multiple learning outcomes.
2. It includes the integration of understanding, skills and strategies.
3. It deals with problems and activities that relate to out-of-school life.
4. It involves the active participation of pupils in all the phases of the project.
5. It provides for self-assessment and independent learning.
6. It requires performance skills that are generalizable to similar situations.
7. It is feasible within the constraints of pupil’s present knowledge, time limits, and available resources and equipment.
8. It is both challenging and motivating to pupils.
9. It is fair and doable by all pupils
10. It provides collaboration between pupils and the teacher.

(Gronlund, 2003; p.148)

#### Example: Individual Project - Room to Decorate: Grades 4 to 6

Choose a room in your house to redecorate.

- (a) Based on the floor space, determine whether it is cheaper to tile or carpet the floor.
- (b) Decide whether to paint the walls or use wallpaper. Present a report explaining your decisions. You should submit appropriate drawings; indicate all necessary calculations, cost of materials and labour and any other important factors which affect your decisions.

#### Example: Statistics Project

Choose ONE of the following:

- (a) Vehicular accidents in your district or country.
- (b) A social problem in your district or country.

Carry out the following tasks:

- (i) Collect relevant monthly data for a period of 10 – 12 months OR annual data for a period of 10 – 12 years.
- (ii) Collect comparative data e.g., for a similar time period or for another district.
- (iii) Carry out the investigation and present the information from (i) and (ii), using appropriate tables, charts and diagrams and graphs as well as appropriate descriptions in words.
- (iv) Describe the information clearly and discuss any important features, trends or problems that the data suggest and compare the two sets of data. You may make predictions, discuss implications or make appropriate recommendations to the Head of Department to rectify or improve any problems you identify.

*\*May be done as an individual or group project*

#### Example: Science Performance Assessment

This performance task is grounded on scientific experiment. However, there is integration of Language Arts and Mathematics. The performance task will be conducted over a two-week period how a plant's surrounding environment affects its growth.

### *Performance Task*

1. Two plants will be placed in two different environments. One plant will be placed in the school's garden, while the other plant will remain in the classroom in a black plastic bag. The plants will be tagged A and B.
2. Students will be placed in groups of three and each group will be required to make journal entries on the growth and progress of each plant.
3. At the end of the two-week period, each group will prepare a report that includes the following:
  - a. Identify the changes made by both plants in chronological order.
  - b. Compare the times it takes for each plant to make changes.
  - c. Explain the needs of each plant under observation.
  - d. Discuss the effects the environment has on each plant.
4. The pupils should achieve the following learning targets by the end of the task:
  - a. Estimate the length of objects
  - b. Compare times
  - c. Write in journals to evaluate circumstances
  - d. Write reports on scientific experiments
  - e. Collect data where required
  - f. Illustrate data in an appropriate way
  - g. Interpret data and draw logical conclusions
  - h. Identify the needs of a growing plants
  - i. Study areas or environments and note plant adaptations
5. An analytic rubric will be used to pinpoint and highlight the strengths and weaknesses of the group projects. This will allow the teacher to give specific feedback to each group as well as to score the projects in a less subjective way when compared to a holistic rubric.

### **6.7 Cooperative Learning/ Group Projects**

Some projects are best done as group work. Educators today are putting increasing emphasis on cooperative learning. It is argued that in real life, citizens have to work with other persons; it is therefore critical to plan tasks that require pupils to work in groups. Nitko, (2004) pointed out that: "Research on cooperative learning indicates that pupils' achievement is best when the learning setting requires both group goals and individual accountability" (p.243). He also suggested that the two conditions are essential in order that the true benefit of cooperative learning can be realised. First, cooperative groups must have a group goal. Second, individual accountability is important. Therefore, when assessing group work, both an assessment of the group's performance and an assessment of the efforts of each individual member are to be done.

### ***Some strategies for assessing group work***

1. Structure the task so that each member of the group has a piece of the work to do. Hopefully peer pressure or the desire to cooperate would motivate each pupil to do his/her part.
2. Monitor the work closely to ensure that all pupils actively pursue the task.
3. Orally examine each pupil to establish their knowledge of the subject, and the degree of his/ her participation and contribution.
4. Develop an assessment scheme to enable members of the group to assess each other's contribution.
5. Develop a procedure for combining the group score and the individual's contribution score for each individual.
6. However, if the objective of the project is to develop cooperative learning rather than individual achievement, the average score for the group would be adequate.

Problem for discussion:

*How would you assess individually candidates involved in group work?*

*Showing how mathematics relates to real life will require open-ended tasks that often focus on writing and communication skills as in the following examples.*

Example: Group Project\_Statistics, Consumer Arithmetic, Health science

The management of your school is considering adding a snack section to the cafeteria offerings. You are asked to assist in the preparatory planning.

- (a) Ask a sample of 100 pupils to give you their 10 most preferred snacks.
- (b) Ascertain the cost price (for bulk rate) and the normal selling price of those snacks.
- (c) Advise the management which snacks to stock and the price at which to sell them.

Present your report, showing how you selected the sample, and all relevant data used to help you arrive at your conclusion.

**Activity for Teachers:**

1. For each of the examples above, identify the standards, strands and sub strands to which the tasks **relate**.
- 2 Suggest projects/ investigations involving statistics, consumer arithmetic, geometry, or any other units in the syllabus you choose.

Group investigation (Consumer Arithmetic, Home Economics and Science -Grades 5 & 6)

With the major improvements in our electrical system consumers now have a choice between a variety of appliances. For instance, consumers can choose between (i) an electrical heater or a solar water heater, (ii) an electrical stove/ range or a gas stove/ range and (iii) an electrical BBQ, gas BBQ or a coal BBQ. {Source: Anguilla workshop 1992}

For this project/investigation pupils are required to investigate if it is more economical to purchase and use an electrical stove/range or to purchase a gas stove/range.

Hint: In doing this project, pupils should consider

- (i) A typical household of 3 to 6 persons.
- (ii) Average number of hours these appliances are used each day.
- (iii) The average of maintenance costs for the different kinds of appliances.
- (iv) The cost of a bottle of propane gas and the average time it lasts.

Submitted by Participants at a Primary School Teachers Workshop conducted in 1992 by James Halliday

## CHAPTER 7

### SCORING RUBRICS

#### Introduction

This chapter deals with scoring rubrics. A rubric describes the scoring rules to mark a performance task. Huba & Freed (2000) suggest that the rubrics should explain to the pupil the criteria against which their work will be judged. They are the guidelines for scoring tasks. Rubrics set out the criteria for scoring; they describe the facets or dimensions of pupil performance that are used for judging the level of achievement (Mc Millan, 2007)

In order to specify the criteria, we need to summarize the dimensions of the performance that we require. Next we try to develop a qualitative scale to show the differences in the level of performance that we expect. We may consider terms like Excellent, Good, Adequate, Weak, Unsatisfactory. Or we may use terms like Outstanding, Very Good, Satisfactory and, Inadequate. Whatever terms we use, we must ensure that there is a qualitative difference in the terms used.

Some of the recommended descriptors from Huba & Freed (2000) that are useful for feedback purposes include the following: (i) Developing, Proficient and Exemplary (ii). "More than Adequate", "Adequate", and "Less than Adequate"(iii) Sophisticated, Competent, Partly Competent, Not yet Competent; (iv) Proficient, marginal, unacceptable (v) Exemplary, Proficient, Marginal, Unsatisfactory

The terms used will depend on whether the scoring is for summative assessment or for formative assessment, the nature of the task and the grade level of the pupils. A plethora of rubrics can be found on the internet and in text books. However, there are a few guidelines that teachers should think about.

#### ***Guidelines for Designing Rubrics***

Ensure that the type of rubric is suitable for the purpose of the assessment
Use terms or descriptors that shows qualitative differences
Use descriptors that will denote the quality of work
Avoid using normative terms like "average", "below average" etc.
State the criteria in words that communicate useful information to the pupil, the parent or other teachers

### ***Number of points on the scale***

In general, the number of points to use depends on the purpose of the rubric and whether it is for formative or summative purposes. Generally speaking, for primary school between 3 and 6 points are common. For very young children, three points may be adequate. One tends to want more discrimination at the upper end of Primary. This writer recommends the use of FOUR points at the lower end and SIX points at the upper end. It is strongly recommended that as far as possible teachers avoid three and five points, because of the central tendency problem. Many teachers tend to score most tasks at the middle point. Using an even number forces the marker to study the responses more keenly and give the appropriate grade.

A number of rubrics written by teachers are given in this chapter. Critique them and adapt them or modify them as you see fit.

#### **TASK for TEACHERS:**

*For each of the rubrics listed below:*

- (i) Determine the Grade level and type of task for which you think the rubric is appropriate.*
- (ii) Modify it for use in your class.*

### **(a) Analytic Scoring Rubric**

Scoring is on four qualities or dimensions

#### **A. Ideas and Development**

**4:** Response is clear, focused, well developed and connects the writing to the specified purpose. The writer develops the topic in a logical, organized way.

**3:** Response is clear and focused. Ideas though related to the purpose specified may be sketchy or overly general. The writer has defined but not thoroughly developed the topic, idea, or story.

**2:** The response does not maintain focus throughout. The writer has defined but not thoroughly developed the topic, idea, or story line; response may be unclear or sketchy or may read like a collection of thoughts from which no central idea emerges. Ideas are minimally related to the purpose specified in the prompt.



**1:** The response tends to be unfocused. The writer has not defined the topic. Ideas may not be related to the purpose specified in the prompt. Details may be irrelevant or too insufficient for the reader to construct meaning.

## **B. Organization**

**4:** The writer develops the topic in a logical, organized way. Supporting details are relevant and provide important information about the topic. The main idea stands out from the details.

**3:** Attempts to introduce and develop the topic in an organized way; adequate focus; adequate sequencing of ideas. Stays on the topic for the most part; Order and structure are present; appropriate conclusion.

**2:** May lack a clear organizational structure; little or no evidence of unity; limited sequencing and or transitions. Details may be randomly placed.

**1:** Organization is not evident; lacks logical direction.

## **C. Sentences and Paragraphs**

**4:** Ideas well organized into paragraphs; the main idea stands out from the details. Supporting details are relevant and provide important information about the topic; demonstrates use of varied and appropriate sentence structure; has few or no run on or fragment errors.

**3:** Ideas are organized into paragraphs. The writer makes general observations without using specific details or does not delineate the main idea from the details. Some supporting details are relevant but limited, overly general, or less important. Adequately demonstrates appropriate and varied sentence structure, or may use simple but accurate sentence structure. May contain a small number of run on or fragment errors which do not interfere with fluency.

**2:** Demonstrates little or no attempt at paragraphing. No distinction is made between main ideas and details or supporting details are minimal or irrelevant or. Writing may be choppy or repetitive. Portions of the writing may be unreadable; errors may impede communication in some parts of the response

**1:** No attempt at paragraphing. Supporting details are absent. The writer uses simple, repetitive sentence structures or many sentence fragments; many errors in structure impeding fluency.

#### **D. Conventions and Mechanics**

**4:** Demonstrates appropriate use of correct spelling, punctuation, capitalization, grammar usage. Errors are minor and do not affect readability.

**3:** Demonstrates adequate use of correct spelling, punctuation, capitalization, grammar usage. Errors may be more noticeable but do not significantly affect readability.

**2:** Demonstrates minimal use of correct spelling, punctuation, capitalization, grammar usage. Errors may be distracting and interfere with readability.

**1:** Demonstrates very limited use of correct spelling, punctuation, capitalization, grammar usage. Errors are numerous and severely impede readability.

*Used with permission from Analytic Rubric by Donna Roberts in Master of Education Thesis*

#### **(b) Holistic 5- Point Scoring Rubric**

**5** = Demonstrates strong composition skills. The response is clear, focused and well developed. Relevant details and choice of words support and enrich the central idea. Demonstrates careful and acceptable use of conventions and mechanics; any errors in language usage, spelling, and mechanics, if present, do not impede meaning.

**4** = Demonstrates good composition skills. The response is clear and focused, although development and sentence structure may suffer minor flaws. Any errors in language usage, spelling, and mechanics, if present, do not impede meaning. The writing effectively accomplishes the goals of the assignment.

**3** = Essay contains satisfactory composition skills. The response is mostly clear and focused, although development shows some weaknesses and sentence structure may not be always clear and effective. Demonstrates competence in the use of mechanics; errors where noticeable do not impede readability; accomplishes the goals of the assignment with an overall effective approach.

**2** = Essay demonstrates acceptable competence in composition skills including adequate development and organization. The development of ideas may be commonplace and assumptions may be unsupported in more than one area or may not be clear and effective. Errors begin to impede readability. Minimally accomplishes the goals of the assignment.

**1** = Essay demonstrates weak competence in composition skills. Focus, development, organization and or sentence structure has flaws. Some ideas are unsupported. Some errors impede clarity and readability. Minimally accomplishes the goals of the assignment.

**0** = Composition skills are very weak. Significant errors in use of conventions and mechanics seriously affect readability. The reader finds it difficult to focus on the message and must reread for meaning; fails to accomplish the goals of the assignment.

*Adapted with permission from Holistic Rubric by Donna Roberts in Master of Education Thesis*

### **(c) Mathematics Holistic Scoring Rubric**

*Suitable for Grades 3 to 6*

**3** = Response is exemplary. Task is completed with no errors. Work is clear and complete. Demonstrates a thorough understanding of the mathematical concepts or procedures embodied in the task

**2** = Response is mostly correct. Task is completed with some minor errors. Work shows basic understanding of the problem and uses anticipated approaches. However, the response may contain minor flaws which reflect inattentive execution of the mathematical procedures, or misunderstanding of underlying mathematical concepts or problems.

**1** = Response is partially correct. Work demonstrates only a limited understanding of the mathematical concepts or procedures embodied in the task. Although pupil may address some of the conditions of the task, the pupil reached an inaccurate conclusion demonstrating misunderstanding of important aspects of the task.

*Adapted with permission from Holistic Rubric by Donna Roberts in Master of Education Thesis*

### **(d) Three - Point Rubric for Problem Solving**

*For Grades 3 to 6*

**3 = Thorough/ Insightful Use of Skills/ Strategies:**

The skills and strategies show some evidence of insightful thinking to explore the problem. Pupil's work is clear and focused. Skills/ Strategies are appropriate and demonstrate some insightful thinking. The response gives possible extensions or generalizations to the solution of the problem.

## 2 = Partial Use of Skills/Strategies

The skills and strategies have some focus, but clarity is limited. The pupil applies a strategy which is only partially useful. Strategy is not fully executed. The pupil starts the problem appropriately, but changes to an incorrect focus. Recognizes the pattern or relationship, but expands it incorrectly.

## 1 = Limited Evidence of Skills/ Strategies

The skills and strategies lack a central focus and the details are sketchy or not present. The procedures are not recorded (i.e., only the solution is present); strategies are random; The pupil does not fully explore the problem, looking for patterns or relationships. The pupil fails to see alternative solutions that the problem requires.

*Used with permission from Holistic Rubric by Donna Roberts in Master of Education Thesis*

(e) **Mathematics Analytic Rubric**

**MATHS SCORING RUBRIC FOR ANALYTIC TRAITS**

**3 TRAITS / DIMENSIONS 4 LEVELS OF QUALITY**

	<b>STRATEGIC KNOWLEDGE (REASONING) D1</b>	<b>CONCEPTUAL KNOWLEDGE (ACCURACY) D2</b>	<b>COMMUNICATION (EXPLANATION) D3</b>
4	Shows a sound understanding of the strategies, principles and procedures required for the solution, reasons well; solution is complete or very near completion	Gives accurate or very near accurate calculations measurements & drawings; mathematical symbols and representations are used correctly; executes algorithms well.	Explanations are clear, logical and coherent; arguments and reasons for decisions are sound. Makes effective use of language, graphs, tables and diagrams to convey ideas
3	Shows a good understanding of the strategies, principles and procedures required for the solution, reasons fairly well; solution is fairly complete.	Gives mostly accurate calculations, measurements & drawings; mathematical symbols & representations are fairly correct; executes algorithms correctly for the most part.	Explanations are clear for the most part; arguments and reasons are for the most part sound. Makes relatively good use of language, graphs, tables and diagrams to convey ideas
2	Shows some understanding of the strategies, principles and procedures required for the solution, reasoning is partly sound, partial solution is given.	Calculations, drawings & measurements are barely adequate; mathematical symbols & representations are correct some of the time. Starts algorithms but does not follow through completely to the end.	Explanations are given but are weak/ unclear at points; arguments/ reasons are not too strong. Use of language, graphs, tables and diagrams partly helpful unclear.
1	Shows little understanding of the strategies, principles and procedures needed to solve the problem, reasoning is inconsistent or inappropriate. Solution is unsatisfactory.	Very few of the calculations, measurements and drawings are accurate; mathematical symbols and representations are rarely correct. Unable to execute algorithms correctly.	Explanations are unclear; arguments/ reasons are weak. Use of language, graphs, tables and diagrams was not helpful.

Created by J. Halliday (2013)

(f) **CPEA Rubric**

**Rubric for Scoring a Book Report**

<b>Dimensions</b>	<b>Weight</b>	<b>Marks Allocated</b>
<b>Organisation</b>	<b>4</b>	
<ul style="list-style-type: none"><li>• Good organisation, details are logically ordered</li><li>• Sharp sense of beginning and end</li></ul>	3 - 4	
<ul style="list-style-type: none"><li>• Organised, details are sometimes mis-ordered</li><li>• Clear beginning and end</li></ul>	2	
<ul style="list-style-type: none"><li>• Some organisation, details jump around</li><li>• Start and end unclear</li></ul>	1	
<b>Quality of Information</b>	<b>4</b>	
<ul style="list-style-type: none"><li>• Supporting details specific to subject</li></ul>	3 – 4	
<ul style="list-style-type: none"><li>• Some details are not supporting the subject</li></ul>	2	
<ul style="list-style-type: none"><li>• Details are insufficient</li></ul>	1	
<b>Personal Response / Creativity</b>	<b>4</b>	
<ul style="list-style-type: none"><li>• Very original presentation of response to book</li></ul>	3 – 4	
<ul style="list-style-type: none"><li>• Fairly organized presentation</li></ul>	2	
<ul style="list-style-type: none"><li>• Minimal originality in personal response</li></ul>	1	
	<b>TOTAL</b>	

Now you should be on your way to create rubrics for the various performance tasks that you give to your pupils whether for feedback purposes or as summative assessments for the mark sheets and school reports.

## CHAPTER 8

### PROTFOLIO ASSESSMENT

Nittko (2004) defined a portfolio as a 'limited collection of a pupil's work used for assessment purposes either to present the pupil's best work(s) or demonstrate the pupil's educational growth over a given period of time' (p. 522). This definition implies that a portfolio is different from a folio that may be described as a collection of a pupil's work. This definition also implies that a careful selection of the pieces of work a pupil does must take place.

Gronlund (2003) argues strongly that a collection of work in a folder is not necessarily a portfolio. There needs to be some clearly specified rule governing the selection of the pieces. There must be some organizing principle that is used to put the pieces together. However, Huber and Freed (2000) describe a type of portfolio referred to as the *all-inclusive portfolio*. This is a collection of work that a pupil produces in a course. We see then that writers differ on what constitutes a portfolio.

#### 8.1 Types of portfolios

The literature also shows that different writers use different terms to describe various types of portfolios. For the purpose of our discussions, in this manual we will consider four types: all-inclusive, developmental, showcase and evaluation portfolios.

##### ***All-inclusive Portfolio***

As stated above, this is a collection of work a pupil produces over a period of time, say for a unit of work or a term. It is a complete record of pupils' achievement, for teachers and pupils to review. It can include projects, papers, homework, laboratory reports and classroom tests. It may also include painting, photographs, sketches, and articles that are relevant to the particular course of study that pupils find interesting and useful. Many teachers begin with this type of portfolio and later make selections to develop other types of portfolios.

##### ***Developmental Portfolio***

A developmental portfolio contains examples of pupils' work along with comments that demonstrate how the pupil has performed over a given period of time. It usually consists of first drafts of pupils' assignments as well as improved drafts after teachers have provided feedback. This approach is useful for formative evaluation purposes. The pupils' role is critical here since

pupils must focus on their weaknesses and attempt to improve their work. Pupils' self-evaluation and self-management skills are also critical to the success of this type of portfolio. Nitko (2004) describes this type as '*a growth and learning progress portfolio*'

This type of portfolio is useful in developing skills in many areas of classroom work such as essay writing, language skills and literacy studies. For example, teachers might give pupils a descriptive essay or a persuasive essay to write. They may discuss the strengths and weaknesses with pupils, and after a period of teaching specific concepts to correct areas of weakness, provide pupils with the opportunity to write a second essay so as to assess pupils' growth and development. The process may be repeated and a third opportunity provided for further improvement.

### ***Showcase Portfolio***

This involves the presentation of pupils' best work. The best works of a pupil are selected to provide evidence of the level of achievement attained. The portfolio can be used to provide evidence to parents at teacher-parent conferences, or it can be used as a showcase for the school's exhibition to demonstrate what the school is doing, or it may be used to show the teacher of the next class what the pupils are capable of doing.

### ***Evaluation Portfolio***

This involves specially selected pieces of work that are used to assess pupil learning, usually to give a grade for the course or part of the course. Since different pupils might do different assignments, it is important that clear criteria be provided for the selection of pieces, whether the teacher does the selection or the pupil does the selection.

## **8.2 Contents of Portfolios**

There are several ways of organizing portfolios. Portfolios can be tailored to meet specific purposes. The content and purposes may change from one situation to another. Fenwick and Parsons (2004, p.149) provide a list of pieces of pupils' work that may be included in portfolios:

1. Formal written papers, articles, descriptions, case studies
2. Assessment inventories (e.g. personality or learning style inventories)
3. Photographs or drawings of learner-created products, artwork
4. Video tapes of learner presentations
5. Journal, memos, personal responses
6. Written reports, attestations to learner performance by others, e.g. peers, colleagues, supervisors
7. Written observations of performance contributed by the instructor



**Table 8.2. Sample Ideas for Portfolios**

LANGUAGE ARTS	SOCIAL STUDIES
<p>Letter to classmates inviting them to a birthday party</p> <p>Application for a vacation job</p> <p>Different types of composition/essay</p> <p>Different types of articles</p> <p>Instruction on how to do something</p> <p>Short stories</p> <p>Poems</p> <p>Pictures that tell a story</p> <p>Selected classroom tests</p> <p>Self -assessment on reading / writing skills</p> <p>Plans for improving skills</p>	<p>Map of your country showing the parishes and historical sites</p> <p>Notes on the agricultural products of Grenada</p> <p>Articles on the contribution of these products to the economic development of your community/country</p> <p>Essays on the national heroes and their contribution to the social/ economic development of Grenada</p> <p>Report on types of groups and their roles at the community and national levels</p> <p>Articles on CARICOM</p> <p>Selected classroom tests</p>
MATHEMATICS	SCIENCE
<p>Selected classroom assignments</p> <p>Selected homework assignments</p> <p>Investigations and projects completed</p> <p>End of unit tests</p> <p>End of term tests</p> <p>Challenging exercises completed</p>	<p>Report on laboratory work</p> <p>Articles on topics covered in class</p> <p>Notes on topics discussed in class</p> <p>Selected classroom tests</p> <p>Projects and investigations</p> <p>End of unit tests</p>

### 8.3 Eight Steps in Designing a Portfolio Plan

1. Determine the criterion and/or standards for the portfolio assessment system
2. Translate the criteria and/or standards into observable behaviours
3. Using the criteria, examine the scope and sequence of the curriculum to determine an approximate time frame for collecting evidence and completing evaluations
4. Determine the stakeholders in the portfolio model
5. Determine the types of evidence to be collected
6. State the method by which the evidence will be transformed into scores
7. Establish a system for reporting assessment information and decisions
8. Establish a series of exemplary pieces/portfolios to provide a comparison

*Adapted from: Shaklee, Barbour, Ambrose & Hansford (1999; p. 66).*

### 8.4 Important Questions to Ask When Designing a Portfolio System

**Why?** Why is a portfolio system being undertaken? This has to do with the purpose of the portfolio system. Two main purposes were discussed when the definition of portfolio was given at the beginning of this chapter. But there are others. McMillan (2004) identified five (5) of these:

- (i) To improve pupils' learning
- (ii) To encourage pupils to take responsibility for their own learning
- (iii) To evaluate the programme effectiveness
- (iv) To develop pupils' self-assessment skills
- (v) To grade pupils' accomplishment

**How?** How will the assessment be done? This is a critical point. Teachers have to decide if they are going to use checklists, rating scales, descriptive scoring rubrics, observation schedules or anecdotal records. Multiple methods should be used in the process of evaluating a portfolio. Allied to this is this question is **WHEN** will the assessments be done.

**When?** Teachers must decide if they are going to score each piece as it is written and give pupils an opportunity to redo assignments to show improvement. In other words, how much weight should they put on development aspects and how much weight on the summative aspects. It is important to emphasize as well that pupils must have multiple opportunities to demonstrate the important skills that are being assessed

**Who?** Will pupils be asked to evaluate their own work? Will they be asked to evaluate the work of their peers? Are there any tasks that parents or other adults will be asked to evaluate? If the answer to any of these questions is 'yes', then the next question is '*What part will these assessments play in the final grade?*' In some cases, the teachers may wish the pupils to employ self-assessment and peer-assessment strategies. But if the portfolio is high stakes, then it would be more appropriate for the classroom teacher and perhaps a colleague to do the assessment.

**What?** The **WHAT** question has many parts. First, what should be included? This has to be specified at the beginning. When dealing with an all-inclusive portfolio, an important consideration is which pieces should be graded? Will pupils have to select their best pieces of work for grading purposes? What is the grading policy? Pupils need to know the answers to these questions before they begin to put together the portfolio.

**Where?** Where will the portfolios be stored? Is there sufficient space to keep portfolios over a number of years or will teachers only keep a sample and the remainder returned to pupils?

## **CHAPTER 9**

### **THREE IMPORTANT PSYCHOMETRIC ISSUES**

There are three important qualities that are required in order if teachers and other critical stakeholders are to make sound judgments and decisions using the data from tests and assessments. First, assessments must be fair. Second, the assessments must be valid. Third, the scores must be reliable. Each of these requires several pages to treat the concept in full. We will confine our discussion to a short treatise on each.

#### **9.1 Fairness**

Nitko & Brookhart (2007) suggest that a fair assessment or test is one that provides scores that (a) are interpreted and used appropriately for specific purposes (b) do not have negative or adverse consequences as a result of the way they are interpreted or use and (c) promote appropriate values (p.512). The test or assignment must be designed in a way that provides all pupils an opportunity to demonstrate their knowledge and achievement.

For a test to be fair, the teachers must test what they have taught or the information that is in the appointed text book to which the pupils have access and are directed to study. The test cannot contain tasks that pupils did not have an opportunity to study. The test must be based on concepts and skills in the designated syllabus that pupils had the opportunity to learn.

The assessments must reflect the relative emphases that the teacher used during instruction. It is common to hear pupils complaining that a teacher spent three weeks on a topic, but it was not tested. However, the same teacher spent one day on a certain topic and two substantial questions came on that topic.

A fair test must not contain any material that is blatantly offensive to any group of pupils or offends any ethnic, racial or gender group. Moreover, the reading level of the tasks should be appropriate for the grade level. Fairness extends to the availability of materials required for the test, classroom conditions when the test is written, the temperature of the room and the general physical conditions under which the test is taken.

### ***Key components of fairness***

1. Items match important learning targets
2. Items match the type of tasks to which pupils were exposed during instruction
3. Items are based on information that pupils had opportunity to learn
4. Items do not favour one group over another
5. The language and format are appropriate to the grade level
6. The language is not offensive to particular group or groups
7. Adequate time is given for pupils to complete the tasks comfortably
8. The difficulty level is appropriate to the age group tested
9. Accommodation is made for pupils with special needs

## **9.2 Validity**

Nitko & Brookhart (2007) defined validity as the soundness of the interpretation and uses pupils' assessment results. Therefore, the assessor must combine evidence from a variety of sources that demonstrate that the interpretations are consistent and correct. This is why valid assessments can only be based on evidence gathered from a variety of sources. These include more than one task, more than one type of tasks, and tasks tested on more than one occasion.

Traditionally, psychometricians speak of content validity, construct validity, criterion validity, face validity, and more recently, consequential validity. These are briefly illustrated in the table below.

**Table 9.1 Major considerations in validation** (Adapted from Linn & Miller, 2005; p.72)

Consideration	Procedure	Meaning
Content validity	Compare the assessment tasks to the specifications	How well do the assessment tasks represent the domain of tasks; are important content emphasized
Construct validity	Are the critical cognitive processes represented?	How well does pupils' performance enable you to assess the critical processes or constructs?
Criterion related validity	Compare assessment results with another measure of performance on a similar test either concurrently or at a later date	How well does performance on this assessment compare or predicts performance on a concurrent or future test?
Consequential validity	Evaluate the effects of the use of the assessment results on teachers and pupils	How well do the assessment results accomplish the intended purposes and avoid unintended effects?
Face Validity	Compare test items to the types of practice items to which pupils were exposed	How much do the test items look similar to what pupils have been taught?

Constructing a table of specifications is one important device that can be used to ensure that assessments have good content, construct and criterion related evidence of validity. Using the same table of specifications can go a long way to ensure that these psychometric qualities are covered in creating parallel tests or assessments.

### 9.3 Reliability

Reliability is another critical characteristic of sound assessments. Linn & Miller (2005; p.69) suggest that: "Reliability is concerned with the consistency, stability, and dependability of scores". Psychometricians have derived various methods of estimating the reliability of test scores. There are methods relating to the internal consistency of the test, parallel forms reliability, generalizability indices, to name a few. We will not focus on these in this manual Two of the procedures are provided in the Appendices for those interested in the procedures.

However, it is important to bear in mind the important factors that impact on the reliability of test scores.

***Factors that affect the reliability of assessments***

CONSTRUCTION OF TESTS	SCORING OF TESTS
1. Number of items; not too few; or too many	1. A marking key is required
2. Difficulty level of the items. Use a balance of easy & challenging items	2. Marking when one is tired; marking under poor lighting
3. Items need to be carefully worded	3. Central tendency errors; avoiding high or low scores
4. Balance coverage of the syllabus as taught	4. Marking too easy or too hard
5. Poor presentation of test items	5. Halo effect; bearing in mind whose pupil script it is you are marking
6. Inadequate time	6. Severity errors
7. Write out the answers as you prepare the question. For the pupil double the time it took you to answer the questions	7. Use a scoring key that has been checked by at least one other colleague

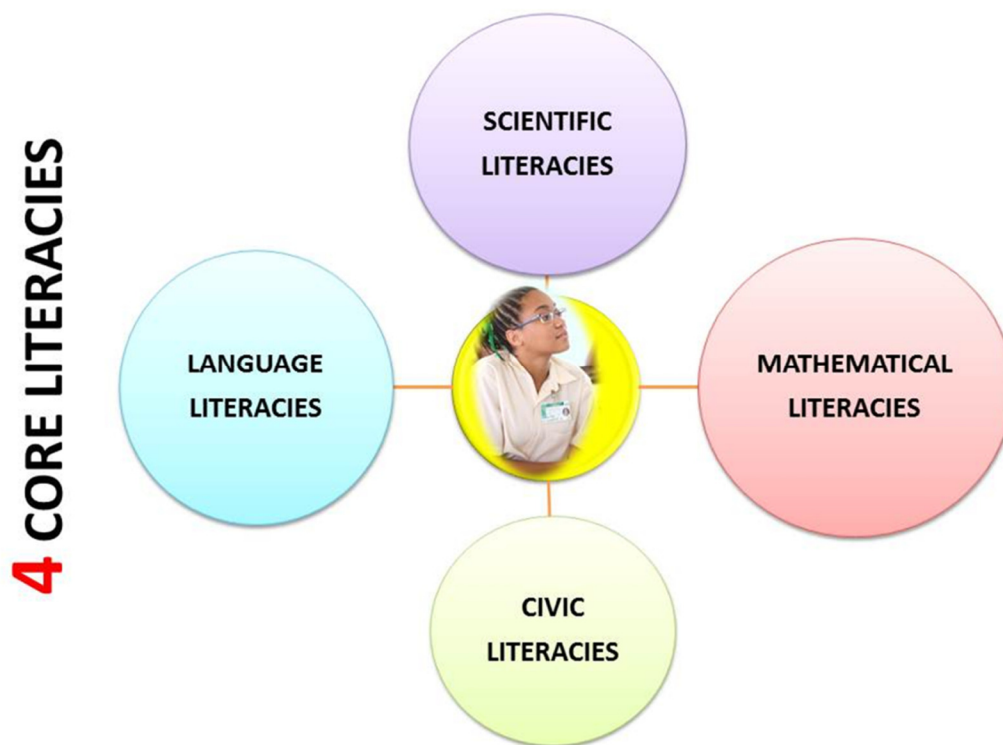
Reliability is critical. If the test is not marked reliably, one cannot have confidence in the results. One cannot generalize over the results. Consider the case of a doctor certified but unable to treat patients or an engineer who builds a bridge which collapses and kills hundreds of people. So too, consider the harm done to hundreds of pupils who are given unfair tests, or are marked unfairly and consider the psychological harm done to thousands of pupils, because the test is not valid, or the marking is not done in a reliable manner.

It is important therefore that continuous assessment strategies employed meet the rigorous criteria for validity, reliability and fairness.

## CHAPTER 10

### Caribbean Primary Exit Assessment

The Caribbean Primary Exit Assessment (CPEA) incorporates internal assessment and external assessment strategies. The internal assessments require pupils to do several classroom tasks. These include but are not limited to (1) project (2) book report (3) writing portfolio. In preparing pupils for these assessments, teachers are expected to use formative assessment strategies to give pupils feedback as they progress in doing these tasks. In particular, teachers are expected to give practice using 'can do' skills as they prepare them for the external assessment in English, Mathematics, Science and Civics or (History and Social Studies). They will also be required to give end of unit tests (summative tests) and submit these to the examination board.



Teachers are also expected to engage pupils in self-assessment activities. Pupils will be required to work in groups for specific tasks, assess each other and provide feedback (peer assessment). Also opportunities should be provided for pupils to do self-reflection, assess their own progress, and plan strategies to rectify their weaknesses. The programme therefore encompasses a variety of strategies in both formative and summative assessments.



The proposed assessment scheme for the OECS is the Caribbean Primary Exit Assessment model which uses internal and external assessments, portfolios, projects and book reviews. The internal assessments, conducted by the classroom teacher, will contribute 40% of the marks and the external assessments will consist of written external examination papers that account for 60% of the marks.

### Internal Assessments

These are divided into two:

- In subset A, the teacher and the pupil work on projects account for 40 marks; second a set of practice exercises, referred to as can-do skills account for 40%, and thirdly, pupils' self-assessment tasks account for 20 marks – making a total of 100 marks.
- In subset B, teachers' tests contribute 25 % of the marks for each of the subjects – English, Mathematics, Science and Social Studies.

### External Assessment

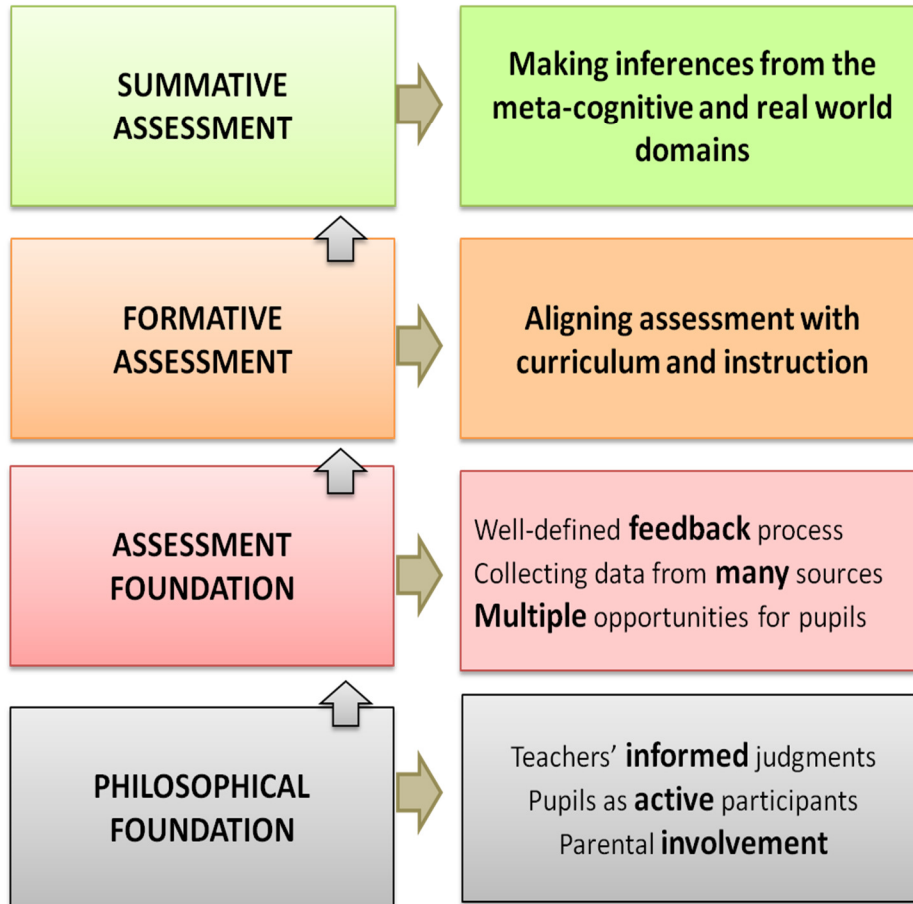
An examination for each of the subjects (English, Mathematics, Science and Social Studies) contributes 100 marks for each subject.

The table below illustrates the weights given to the various assessment components.

**Table 1.2 CPEA Structure**

Internal 40%				External 60%	
Teacher & Pupil		Teacher Prepared Tests		CXC	
Project	40	English	25	English	100
Practice Can Do Skills	40	Mathematics	25	Mathematics	100
Self-Assessment	20	Science	25	Science	100
		Social Studies	25	Social Studies	100
100		100		400	
200					

## CXC APPROACH TO CPEA



### **Assessment Tasks associated with the CPEA**

1. Project: The project should focus on something that is relevant and meaningful to study in pupils' individual community. It must incorporate knowledge and skills from all four areas: Language, mathematics, science and civics. The project is usually done as a group task to encourage collaborative attitudes, but there will be opportunity for individual input as well.
2. Portfolio: Selected writing pieces and other relevant compositions by pupils
3. Book review: Comments by pupils on an interesting (useful) book that they have read. Guidelines or subheads for pupils to use are provided.

All of these assessment tasks have been discussed in previous chapters.

## GLOSSARY

Term	Meaning
<i>Affective domain</i>	This relates to behaviours, attitudes, feelings, values and interests and personality issues that affect learning.
<i>Assessment</i>	The process of gathering information about students' attainment for the purpose of making educational decisions.
<i>Authentic Assessment</i>	Assessment approach that involves performance based tasks designed in such a way that the task is as close to real life situations as possible. In some cases, simulation type exercises are used, since they approximate the real situation.
<i>Alternative Assessment</i>	Assessment devices other than multiple-choice and other selected-response type items
<i>Continuous Assessment</i>	Consists of a variety of devices and procedures to evaluate students' performance for the purpose of improving student learning and the effectiveness of the instructional programme as well as other educational decisions.
<i>Internal assessment</i>	Refers to assessment of students' work by the classroom teacher.
<i>Performance Assessment</i>	(A short form for performance-based assessment) is a system of assessment that requires students to demonstrate what they know or can do by creating a response verbally or in writing, making a presentation, constructing a product or performing a specific task.
<i>Portfolio Assessment</i>	A technique that involves students assembling pieces of work according to prescribed principles and the evaluation of those pieces using agreed predetermined criteria. Some writers insist that the pieces should include evidence of students' reflection on their work; others do not.

<b>Term</b>	<b>Meaning</b>
<i>School-based Assessment</i>	Assessment of students' progress / performance carried out by the classroom teacher during the course of instruction. The assessment tasks may be entirely controlled by the teachers or some may be done following specific guidelines set out by an external agency.
<i>Benchmark</i>	A term used to identify content standards for a particular age group or age groups. They represent the broad aims of a programme or course for a grade or set of grades. Standards are broken down into benchmarks and benchmarks are broken down into specific objectives or learning targets.
<i>Checklist</i>	A device used to assess the existence of various attributes. It consists of a set of specific elements and a place for marking whether or not each element is present.
<i>Cognitive</i>	Checklists can be used to assess characteristics of process, procedures or product, or behavioural or attitudinal attributes. Refers to mental abilities such as recall, comprehension, application, analysis, synthesis, problem solving etc. It involves skills dealing with memory and processing of information.
<i>Cognitive Domain</i>	Refers to behaviour related to thinking, memorising, reasoning, analysing and problem solving.
<i>Completion Item</i>	Refers to an assessment item that requires the pupil to supply a word or phrase to complete a statement so that it makes sense.
<i>Construct validity</i>	Refers to the degree to which a test measures certain traits or abilities that it is intended to measure.
<i>Content validity</i>	Refers to the degree to which a test samples the subject matter. This may include behaviour, knowledge, skills and attitudes. To the extent that the test is a representative sample of all areas of the course to that extent there is content validity.
<i>Correlation</i>	The extent to which two variables vary concomitantly – that is, as one increases in value, the other increases or as one

Term	Meaning
	decreases in value the other decreases in value. The extent of the correlation is usually expressed by a correlation coefficient that has a value between is and positive or negative.
<i>Criterion related validity</i>	The extent to which there is agreement between the test score of an instrument and one or more external instruments that measure the same attributes. We can speak of concurrent validity – if we are interested in assessing the extent to which two or more instruments measure the same attributes or we can speak of predictive validity if we are concerned with the extent to which an instrument ranks students at one point in time to the way in which a similar instrument ranks students subsequently.
<i>Essay</i>	Refers to an assessment item that requires a response in several sentences to a direct or indirect instruction.
<i>Evaluation</i>	A systematic process of determining or judging the merit or worth of students' product, performance or behaviour or for judging the value of an educational programme; policy, process or procedure.
<i>Formative evaluation</i>	The process of judging the worth of students' performance while instructional is in process. Oral questions, observations or tests may be used to provide the teacher with information whereby a judgement is made on the students' progress and hence what courses of action are needed based on that judgment.
<i>Summative evaluation</i>	Refers to evaluation of students' performance or achievement at the end of a unit, course of work or programme. Summative evaluation focuses on the attainment or achievement levels.
<i>Goals</i>	General statements of what students are expected to know and be able to do after a period of instruction lasting for several weeks, a term or even a year. The statements usually represent

<b>Term</b>	<b>Meaning</b>
	the philosophy that the policy makes of curriculum directors hope to achieve by setting up the course.
<i>Halo effect</i>	A type of rater effect in which the teacher or marker evaluating a set of tasks for several students allow their evaluation to be influenced by selected pieces of information about students rather than by the relevant aspects of the students' work.
<i>Meta-cognitive</i>	A term used to describe the psychological processes that enable individuals to direct and guide their learning activities. These processes include planning, monitoring and evaluating the behaviour and skills, the ability to evaluate progress and to re-direct or refocus their planning and activities in directions that are likely to ensure success.
<i>Objectives</i>	A statement that specifies on observable and measurable terms what students are expected to know or be able to do as a result of instruction.
<i>Behavioural objective</i>	An educational outcome that specifies the condition under which the knowledge, behaviour or skill learnt is to be exhibited.
<i>Educational objectives</i>	These are general statements of what students are expected to learn to be able to do usually after a lesson.
<i>General objectives</i>	Broad or general statements that indicate the focus of a course of programme. The term is similar to educational goals. Global objective may be used as a synonym for general objective.
<i>Global objective</i>	See objective—general objective
<i>Specific objective</i>	States in specific terms what students are expected to do at the end of a lesson. Learning targets, learning outcomes and

<b>Term</b>	<b>Meaning</b>
	behavioural objectives are terms used almost in the same way as the term specific objective.
<i>Psychomotor Domain</i>	Range of skills needed to function and perform ordinary practical tasks.
<i>Portfolio</i>	A collection of students' work, artefacts or assignments according to some prescribed principles for purpose of demonstrating the level of attainment or skill achieved.
<i>Project</i>	An extended performance task that results in long-term activity and engages behaviours from one or more of the three educational domains, cognitive, affective and psychomotor domains.
<i>Reliability</i>	This is the consistency with which repeated tests measure the same traits or skills. Implied in the term is the extent to which stakeholders can depend on the test to indicate students' level of achievement.
<i>Rubric</i>	Also referred to as scoring rubric or descriptive scoring rubric. A coherent set of rules or guidelines for scoring or rating assignments, performance-based tasks, achievement, behaviour, attitudes etc.
<i>Generic scoring rubric</i>	A set of rules or guidelines for scoring a wide range of similar tasks, assignments or performances.
<i>Holistic rubric</i>	Consists of a set of guidelines to evaluate a performance of level of attainment as a whole.
<i>Task-specific scoring rubric</i>	A set of guidelines for scoring a specific task. The guidelines will not be applicable to any other task.



<b>Term</b>	<b>Meaning</b>
<i>Standards</i>	Statements about what students are expected to learn from a programme of study. Content standards express the knowledge and skills that students are expected to master. They represent the overall goals of a course or programme. Performance standards refers to the degree or level at which students are expected to perform.
<i>Short Answer</i>	An assessment item that poses a task in question format, an incomplete statement format, or as an instruction, so that the response can be given as, one or two statements, a number, a short calculation or by mathematical statements.
<i>Table of specification</i>	<p>This is usually a two-way grid that summaries how a test instrument should be developed. Usually content or objectives are expressed on the vertical dimension and cognitive processes on the horizontal domain. The test developer or planner must decide, using professional judgment or established guidelines, how to assign various items to the cells of the grid.</p> <p>The Revision of Bloom's taxonomy however, requires considerable adjustment to the traditional way of preparing a task of specifications.</p>
<i>Task</i>	<p>An assessment activity or assignment that is to be done by a student or group of students.</p> <p>Task is a comprehensive term. A test item or a question can be referred to as a task.</p>
<i>Taxonomy</i>	A classification system in which the items are arranged in a systematic, ordered manner. Bloom's taxonomy is the most well- known taxonomy or classification scheme in education.
<i>Test</i>	An instrument consisting of a set of items, questions or tasks that are served to measure one's ability or achievement.
<i>Criterion-referenced test</i>	A test that focuses on the objectives of the content. Its purpose is to determine how well students master these objectives. It is

Term	Meaning
	scored according to pre-established standards of mastering. The performance of other students is not of primary concern in grading a student's performance; rather it is the student's performance in relation to the established standards.
<i>Norm-referenced test</i>	A test designed to maximise discrimination among examinees. The original sample on which the test is scored provides a distribution that is used to compare future samples. The original sample becomes the 'norm' group.
<i>Standardised tests</i>	<p>Tests that meet specific criteria:</p> <ul style="list-style-type: none"> <li>(i) They are constructed using prescribed guidelines.</li> <li>(ii) They are administered under prescribed conditions.</li> <li>(iii) They are scored using definite rules.</li> <li>(iv) The results are compared to a reference group.</li> </ul> <p>Usually the reliability and validity evidence of commercialised standardised tests as well as information on the norm or reference group are provided with the tests.</p>
<i>Trait</i>	A characteristic or attribute of an individual. In its broadest sense a trait can refer to ability. In most educational applications however, one thinks of narrower aspects of ability such as reasoning ability, spatial ability. In psychology self-esteem, motivation, are traits, sometimes called constructs.
<i>Validity</i>	Refers to the soundness or accuracy with which one can make inferences from the test scores or assessment devices. The main types are: content validity, construct validity and criterion related validity. More recently consequential validity has been added.

## BIBLIOGRAPHY

- Airasian, Peter (2000). *Assessment in the Classroom: A Concise Approach*, 2<sup>nd</sup> edition. Boston, MA: McGraw Hill, Higher Education.
- Anderson, Lorin W. & Krathwohl, David R. (eds.) (2001). *A Taxonomy for Learning, Teaching and Assessing*. NY: Addison Wesley Longman Inc.
- Arter, Judith & McTighe (2001). *Scoring Rubrics in the Classroom*. Thousand Oaks, CA: Corwin Press, INC – Sage Publication Company
- Benjamin, Bloom (ed.) (1956). *Taxonomy of Educational Objectives, Book 1 – Cognitive Domain*. London: Longman.
- Broomes, Desmond R & Halliday, James A. (1991). Major Issues in Assessing Mathematics Performance at 16+ Level: A Caribbean Perspective in Mogens Niss (Editor). *Cases of Assessment in Mathematics Education- An ICMI Study*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Capper, Joanne (1996). *Testing to Learn – Learning to Test*. Newark, Delaware: International Reading Association, Academy for Educational Development (AED), Washington.
- Carey, Lou (2001). *Measuring and Evaluating School Learning: 3<sup>rd</sup> Edition*. Needham Heights, MA: Allyn & Bacon.
- Chatteriji, Madhabi (2003). *Designing Educational Tools for Educational Assessment*. Belmont, CA: A Pearson Education Inc.
- Corbin, Kristin Chatter Houses in Ministry of Education, Youth Affairs and Sports, *New Horizons* p 15-16.
- Davis, Barbara, C. (2001). *Tools for Teaching*. San Francisco, CA: Jossey & Bass
- Deere MT (1974). Assessment of Project Work. In H.G McIntosh (Editor). *Techniques and Problems of Assessment – A Practical Handbook for Teachers*. Hill St, London: Edward Arnold Pub.
- Ebel, Robert. L (1979). *Essentials of Educational Measurement, 3<sup>rd</sup> Edition*. Englewood Cliffs, New Jersey: Prentice Hall.
- Glatthorn, Allan A. (1999). *Performance Standards and Authentic Learning*. Larchmont, NY: Eye on Education.

- Gronlund, Norman E. (2006). *Assessment of Pupil Achievement*, 8<sup>th</sup> edition. Boston, MA: Pearson Education Inc.
- Haladyna, Thomas M. (1997). *Writing Test Items to Evaluate Higher Order Thinking*. Needham Heights, MA: Allyn & Bacon.
- Halliday, James A. (2000). *New Directions for mathematics Assessment in Caribbean Secondary Schools- Authentic Assessment Tasks* Unpublished.
- Halliday, James A. (2004) *Manual on Continuous Assessment*. Prepared for the Grenada Ministry of Education Assessment Project – Unpublished.
- Halliday, James A. (2005). *Measuring Learning and Achievement in Secondary Schools: A Booklet* prepared for Ministry of Education, Barbados Unpublished.
- Halliday, James A. (2010). *Principles and Procedures for Testing, Measurement and Assessment in Schools*. Unpublished.
- Halliday James A (2012). *The impact of Holistic and Analytic Scoring on Rater Reliability and the Dependability of mathematics Performance Test Scores* Ph D thesis for UWI
- Hart, Diane (1994). *Authentic Assessment: A Handbook for Educators*. Melono, Park, CA: Addison Wesley Pub. Co.
- Huba, Mary E. & Freed, Joanne (2000). *Learner Centred Assessment on College Campuses: Shifting the Focus from Teaching to Learning*. Needham Heights, MA: Allyn & Bacon.
- Lesh & Lamon (1992). 'Interpreting Responses to Problems with Several Levels and Types of Correct Answers' in R. Lesh & Lamon (eds.), *Assessment of Authentic Performance in*
- Linn R L& Miller (2005) *Measurement and Assessment in Teaching* Upper Saddle River, NJ: Pearson Prentice Hall
- Maki, Peggy (2004). *Assessing for Learning*. Sterling, VA: Stylus Publishing Company
- Marzano, Robert (2001). *Designing a New Taxonomy of Educational Objectives*. Thousand Oaks, CA: Corwin Press Inc., A Sage Publication Company.
- Marzano, Robert (2010). *Formative Assessment and Standards Based Grading: Classroom Strategies That Work*. Bloomington, IN: Marzano Research Laboratory.
- McMillan, James H. (2007). *Classroom Assessment: Principles and Practice for Effective Instructors*, 4<sup>th</sup> edition. Boston, MA: Pearson Educational Inc.
- National Council of Teachers of Mathematics (1995). *Assessment Standards for School Mathematics*. Reston, VA: NCTM.

- National Council of Teachers of Mathematics (2000). *Assessment for School Mathematics: A Practical Handbook*. Reston, VA: NCTM
- Nitko, Anthony & Brookhart, Susan. (2007). *Educational Assessment of Pupils*, 5<sup>th</sup> edition. Upper Saddle Rd., NJ: Pearson Prentice Hall Inc.
- Nitko, Anthony (1999). 'Curriculum-based Continuous Assessment: A Framework for Concepts, Procedures and Policy' in *Patricia Broadfoot (ed.) Assessment in Education: Principles and Practice*. Oxfordshire, UK: Carfax Publishing Co.
- Nitko, Anthony (1996). *Educational Assessment of Pupils*, 3<sup>rd</sup> edition. Englewood Cliffs, NJ: Prentice Hall Inc.
- Payne, David (2003). *Applied Educational Assessment*, 2<sup>nd</sup> edition. Belmont, CA: Wadsworth Thompson Learning
- Popham, James. (2008). *Transforming Assessment*. Alexandria, VA: Association for Supervision and Curriculum.
- Sax, Gilbert (1997). *Principles of Educational and Psychological Measurement and Evaluation*, 4<sup>th</sup> edition. Belmont, CA: Wadsworth Publishing Co.
- Shaklee, Barbara, Ambrose & Hansford (1999). *Designing and Using Portfolios*. Needham Heights, MA: Allyn & Bacon
- Selek, William R. & Galls, Michael A. (2002). *Fundamentals of Mathematics*, 9<sup>th</sup> edition. Upper Saddle Riva, NJ: Prentice Hall
- Ward, Anne & Murray-Ward, Mildred (1999). *Assessment in the Classroom*. Belmont, CA: Wadsworth Publishing Co.
- Worthen B.R, White K.R, Fan X, and Sudweeks R. (1999). *Assessment in Schools*. Reading, MA: Addison Wesley Longman.

# Skills & Competencies to be Assessed

## APPENDIX

LANGUAGE	MATHEMATICS	SCIENCE	CIVIC
<ul style="list-style-type: none"> <li>listening with a high degree of understanding to instructions, descriptions, explanations and narration in Standard English, in a familiar accent and in the vocabulary and sentence structure appropriate to the age of the student;</li> <li>speaking by using words exactly and precisely for his age to communicate thoughts and feels; demonstrate spontaneity in speaking in a variety of situations;</li> <li>thinking creatively, critically and constructively;</li> <li>responding sensitively, to varied and meaningful literature and other forms of art at the appropriate level;</li> <li>reading effectively and for different purposes a variety of print or electronic media;</li> <li>expressing oneself in the following forms of writing: explanations, narratives, descriptions, letter writing and do so legibly; and</li> <li>using various forms of visual literacy to interpret and gain information</li> </ul>	<ul style="list-style-type: none"> <li>recognizing and using number concepts;</li> <li>understanding and using fractions, percentages, proportions and decimals;</li> <li>application of mathematical concepts, facts and algorithms related to number, money, measurement and geometry (space and shape);</li> <li>performing tasks involved in measurement;</li> <li>using tally charts and other methods to collect and record data and being able to read coordinates;</li> <li>approaching problem-solving and inquiry at a level appropriate to the age of the student; and</li> <li>applying principles to solve worded problems and everyday problems</li> </ul>	<ul style="list-style-type: none"> <li>observing or using the senses to find out about objects or events;</li> <li>classifying, based on properties;</li> <li>predicting or describing in advance the outcome of an event;</li> <li>measuring by using instruments;</li> <li>recording by using appropriate language and appropriate units and drawings; diagrams, tables, charts and graphs;</li> <li>interpreting by comparing and drawing conclusion;</li> <li>evaluating by assessing, making suggestions and conclusions;</li> <li>designing procedures to obtain information;</li> <li>demonstration of an appreciation of the value of science as a discipline, as well as the role of science in technological advances and in everyday decision making in the home, community and workplace</li> </ul>	<ul style="list-style-type: none"> <li>an understanding of power, authority and governance;</li> <li>an understanding of the creation, maintenance and change of government;</li> <li>an understanding of how governments can be kept responsive to citizens; and</li> <li>the ability to see things from another's point of view</li> </ul>

### Summary of the Quellmalz Taxonomy

Classification	Definition
Recall	The first level of Quellmalz Taxonomy is termed <b>Recall</b> . It is similar to the <b>Knowledge</b> Level in Bloom's Taxonomy. Recall tasks would require pupils to remember key facts, recognize definitions, concepts, rules and principles. Also placed at this level is the ability to associate concepts with related concepts. So this level touches some aspects of the <b>Comprehension</b> Level in Bloom's Taxonomy as well.
Analysis	The second level in Quellmalz's Taxonomy is <b>Analysis</b> . It is believed that analysis is usually needed before one can deal with the skills of comparison and making inferences. Analysis requires pupils to divide a whole in component parts, identify part/whole relationships and identify cause and effect relationships. This is much the same as the <b>Analysis</b> Level in the Bloom's Taxonomy.
Comparison	Now that the skills of <b>Analysis</b> are developed, pupils have the skills to make meaningful comparisons. Pupils can then recognise or explain similarities and differences. It is to be noted that this level involves some aspects of Bloom's <b>Analysis</b> . But it is given a separate category to emphasize that pupils are required to go beyond breaking up a whole into parts. <b>Comparison</b> has been thought of as a level of <b>Comprehension</b> in Bloom's Taxonomy. Placing <b>Comparison</b> above <b>Analysis</b> emphasizes that to do meaningful comparisons requires higher level thinking than their <b>Comprehension</b> , <b>Application</b> and <b>Analysis</b> as set out in Bloom's Taxonomy.
Inference	This category requires skills in deductions and inductive reasoning. Many <b>Application</b> problems do require the use of these skills. It is therefore, unnecessary to put <b>Application</b> as a separate category. To use deductive reasoning tasks pupils are given details or evidence and are expected to arrive at the generalization. Tasks that require pupils to integrate information, to hypothesize, predict and draw meaningful conclusions from evidence are included in this level. Hence the skills given under <b>Synthesis</b> in Bloom's Taxonomy are included in this level.
Evaluation	Tasks at this level require pupils to judge the quality or worth of something, assess its credibility, use established criteria to argue logically or to reach a conclusion. Pupils have to assemble evidence and explain interrelationships based on the evidence. Bloom's level of <b>Synthesis</b> and <b>Evaluation</b> would fall in Quellmalz's <b>Evaluation</b> level.

**Thinking Skills for Different Subject Matter in Quellmalz Taxonomy**

Category	<i>Subject Matter</i>		
	Science	Social Science	Literature
Analyse	Identify components of a process, features of animate and inanimate objects	Analyse components of arguments, elements of an even	Identify components of literacy, expository, and persuasive discourse
Compare	Compare properties of objects components of processes	Compare causes an/or effects of separate events; social, political, economics, geo- graphic features	Compare meanings, themes, plots, characters, settings, arguments
Infer	Draw conclusions; make predictions; pose hypo-thesis tests and explanations	Predict, hypothesis, conclude, interpret, using historical, social, political, economic, geographic information	Infer theme, significance, characters' motivations; interrelationships of literary elements
Evaluate	Evaluate soundness of procedures, credibility of conclusions significance of findings	Evaluate credibility of arguments, decisions, reports, significance	Evaluate believability, significance, form, completeness, clarity

Source: Nitko (1996, p.27) – Educational Assessment of Pupils (Second Edition), Englewood Cliffs, New Jersey: Prentice Hall



## The Core Thinking Skills

The Core Thinking Skills focus on the mental processes and cognitive skills involved in performing learning targets and assessment tasks.

SKILL CATEGORIES	COGNITIVE PROCESSES AND EXAMPLES
<b>1.0 Focusing Skills</b> - Attending to selected pieces of information and ignoring others.	
1.1 Defining problems	Clarifying needs, discrepancies or puzzling situations
1.2 Setting goals	Establishing direction and purpose
<b>2.0 Information-Gathering Skills</b> - Bringing to consciousness the relevant data needed for cognitive processing	
2.1 Observing	Obtaining information through one or more senses
2.2 Formulating questions	Seeking new information through inquiry
<b>3.0 Remembering Skills</b> - Storing and retrieving information	
3.1 Encoding	Storing information in long-term memory
3.2 Recalling	Retrieving information from long-term memory
<b>4.0 Organising Skills</b> - Arranging information so it can be used more effectively	
4.1 Comparing	Noting similarities and differences between or among entities
4.2 Classifying	Grouping and labelling entities on the basis of their attribute
4.3 Ordering	Sequencing entities according to a given criterion
4.4 Representing	Changing the form but not the substance of information
<b>5.0 Analysing Skills</b> - Clarifying existing information by examining parts and relationships	
5.1 Identifying attributes and components	Determining characteristics or parts of something
5.2 Identifying relationships and patterns	Recognising ways elements are related
5.3 Identifying main ideas	Identifying the central element; for example, the hierarchy of key ideas in a message or line of reasoning
5.4 Identifying errors	Recognising local fallacies and other mistakes and, where possible, correcting them
<b>6.0 Generating Skills</b> - Producing new information, meaning, or ideas	
6.1 Inferring	Going beyond available information to identify what reasonably may be true
6.2 Predicting	Anticipating new events, or the outcome of a situation
6.3 Elaborating	Explaining by adding details, examples, or other relevant information
<b>7.0 Integrating Skills</b> - Connecting and combining information	
7.1 Summarizing	Combining information efficiently into a cohesive statement
7.2 Restructuring	Restructuring
<b>8.0 Evaluating Skills</b> - Assessing the reasonableness and quality of ideas	
8.1 Establishing criteria	Setting standards for making judgements
8.2 Verifying	Confirming the accuracy of claims

### ***An Outline of Bloom's Taxonomy of Educational Objectives: Cognitive Domain***

<b>Level</b>	<b>Explanatory Examples</b>
<b>1.00 Knowledge</b>	
1.11 Knowledge of terminology	Simple definition required
1.12 Knowledge of specific facts	Dates, events, persons, places
1.21 Knowledge of conventions	Rules of etiquette, grammar
1.22 Knowledge of trends and sequences	Processes, decisions regarding time
1.23 Knowledge of classifications and categories	Classes, sets, divisions, arrangements useful to a given field
1.24 Knowledge of criteria	Criteria to evaluate facts, principles, etc.
1.25 Knowledge of methodology	Techniques and procedures
1.31 Knowledge of principles and generalizations	Abstractions that summarize observations of phenomena
1.32 Knowledge of theories and structures	Body of principles and generalizations
<b>2.00 Comprehension</b>	
2.10 Translation of one level of abstraction to another; translation from one form to another form	Restating a problem, reducing size of communication. Giving an example
2.20 Interpretation	Reordering and rearranging, qualifying
2.30 Extrapolation	Extension of data to past or future
<b>3.00 Application</b>	Applying data to new problems
<b>4.00 Analysis</b>	
4.10 Analysis of elements	Finding assumptions, distinguishing facts from opinion
4.20 Analysis of relationships	Relevant data, causal, finding fallacies
4.30 Analysis of organizational principles	Form and style, inferring author's purpose
<b>5.00 Synthesis</b>	
5.10 Production of a unique communication	Skill in writing, telling stories, composing
5.20 Production of a plan, or proposed set of operations	Proposing hypotheses, planning units, tool designing
5.30 Derivation of a set of abstract relations	Formulation of a theory
<b>6.00 Evaluation</b>	
6.10 Judgements in terms of internal evidence	Evaluation of logical accuracy, consistency, and other internal criteria

6.20 Judgements in terms of external criteria	Evaluating theories, judging by standards, weighing standards
---	---

*Source: Sax (1990 Table 3.2, p. 64)*

***Examples of Standards Aligned to Bloom's Taxonomy Levels***

Category	Learning targets for Assessment Teachers	Learning targets Language Arts Teachers
1.21	Define different types of tests	Explain the term 'decoding' as used in Reading
2.10	Distinguish between formative assessment and continuous assessment	Distinguish between a cloze test and a maze test
2.10	Explain the levels of Barrett's Reading taxonomy	Differentiate between phoneme and morpheme
3.00	Explain the importance of reliability in the context of high stakes testing	Use correlation analysis to determine rater consistency of a pair of scorers
3.00	Relate the purpose of testing to the type of test	Use adjectives in descriptive writing
4.20	Discuss the advantages of using norm referenced tests	Discuss the advantages and disadvantages of using a maze test
5.20	Apply Bloom's taxonomy in constructing classroom a test	Write a technical evaluation report on the pupils' performance on a test
6.10	Comment on the usefulness of the Revision of Bloom's taxonomy	Comment on the problems in using standardized Reading tests, constructed in North America, in Caribbean schools

## ***Barrett's Taxonomy***

<b>Literal Comprehension</b> <b><i>Recognition</i></b> Recognition of Details Recognition of Main Ideas Recognition of Sequence Recognition of Comparison Recognition of Cause and Effect Relationships Recognition of Character Traits <b><i>Recall</i></b> Recall of Details Recall of Main ideas Recall of Sequence Recall of Comparison Recall of Cause and Effect Relationships Recall of Character Traits
<b>Reorganization</b> Classifying Outlining Summarizing Synthesizing
<b>Inferential Comprehension</b> Inferring Supporting Details Inferring Main Ideas Inferring Sequence Inferring Comparisons Inferring Cause and Effect Relationships Inferring Character Traits Predicting Outcomes Interpreting Figurative Language
<b>Evaluation</b> Judgements of Reality or Fantasy Judgements of Fact or Opinion Judgements of Adequacy and Validity Judgements of Appropriateness
<b>Appreciation</b> Emotional Response to Context Identification with Characters or Incidents Reactions to Author's Use of Language Imagery, Vocabulary, Feelings Produced by the Author

