The University of the West Indies

STUDENT ASSESSMENT ESSENTIALS HANDBOOK

Prepared by

Instructional Development Units
Cave Hill, Mona and St. Augustine

April 2006

Our assessment efforts are handicapped in part because we are not really very clear about what we are trying to accomplish, and in part because we perpetuate questionable practices out of sheer habit, for convenience, or to fulfill purposes that are unrelated or at best tangential to the basic mission of our colleges and universities. (Alexander Astin, 1991).

Preface

The rapid growth of The University of the West Indies (UWI) over the past decade has created challenges for faculty members in a range of areas, including the conduct of student assessment. This challenge has been experienced in universities worldwide as the demand for and access to tertiary and higher education increases, and as development in scholarship relating to student assessment provides educators with new directions of thought and practice concerning this important issue.

The UWI is committed to ensuring that its teaching/learning environment is one which meets the standard required of a quality driven institution. This demands the use of curricula in which relevant and appropriate assessment of student learning is a critical component. This handbook has been prepared to assist faculty members to understand the major concepts which guide best practice in the conduct of student assessments. The concepts and practical explanations offered will allow faculty members new to the management of the educational process to gain insight into the requirements of quality assessment and also encourage all faculty members to employ these strategies. The overall aim is to establish at UWI a culture of student assessment which is valid, reliable and reflective of the institution's objectives of producing graduates with the knowledge, skills and competencies needed to function in a world which requires leaders who are open minded, able to think critically, solve problems and stimulate as well as cope with change. We hope that the handbook will be thoroughly read, frequently consulted and well used.

Preparation of the handbook has been spearheaded by the Managers of the Instructional Development Units on the three UWI campuses and we thank Dr. Anna-May Edwards-Henry, Dr. Beryl Allen and Dr. Jamillah Grant for their initiative in this regard, and their commitment to the task. The support and advice of members of the Implementation Task Force are also acknowledged, with thanks and appreciation.

Professor Elsa Leo-Rhynie
Chair
Implementation Task Force

April 2006

IDU WRITING TEAM

The following staff members from the Instructional Development Units were responsible for preparing the Student Assessment Essentials Handbook:

Dr. Anna-May Edwards-Henry
Programme Co-ordinator, IDU, St. Augustine

Dr. Beryl Allen
Co-ordinator, IDU, Mona

Dr. Jamillah Grant
Curriculum Specialist, IDU, Cave Hill

ACKNOWLEDGEMENTS

The handbook writing team wishes to express gratitude for the invaluable comments received from Prof. Stafford Griffith and Ms. Jacqueline Moniquette.

TABLE OF CONTENTS

APPENDICES

INTRODUCTION

A Philosophy on Assessment

(The UWI Examination System Review: Implementation Task Force)

Education is what survives when what has been learned has been forgotten
(Skinner, 1956)

Assessment methods and requirements probably have a greater influence on how and what students learn than any other single factor
(Boud 1988)

The Strategic Plan II (2002-2007) of The University of the West Indies outlines as one of its key initiatives a student-centered learning environment and a major quality initiative to "institutionalise a system of regular departmental reviews of the curriculum, analysis of examination results and teaching, learning and assessment methods". This statement is intended to further support these initiatives and in particular, to ensure that the institution as a whole adopts a common view concerning what constitutes appropriate assessment practices.

Increasingly, schools, colleges and universities are moving away from the well-known, traditional learning and assessment avenues and are moving along new paths of learning and less traditional modes of assessment. This movement is underpinned by well-documented research which demonstrates that traditional assessment does not necessarily provide the kind of support that students need to be successful at their studies in this information rich and technology driven world. More and more it is being agreed that alternative or innovative assessment is based on the premise that students can get the most out of their learning experience through seeing knowledge with new eyes and understanding the reasons why assessment decisions are made.

Successful functioning in society demands that students must be able to do more than perform specific tasks. The information age of continual change requires individuals to acquire and hone cognitive, meta-cognitive and social competencies to respond to many different pressures, such as the drive to use more multimedia, the need for lifelong learning and the changing labour market (Harvey & Green, 1994).

> If we want our students to develop and demonstrate higher-order skills such as the application of theoretical knowledge to a given context, analysis and synthesis of new components of their learning and sensitive evaluation of how they themselves and their peers have achieved, then we need to look at new ways of assessment (Brown & Glasner, 1999). Within this context, the University takes the position that assessment of student learning is a vital process for providing systematic indicators of the quality of the students' intellectual engagement. The assessment process should therefore facilitate the maintenance of standards and motivate students throughout their studies so that they become graduates who are highly-skilled and well-prepared for postgraduate studies, professional training and the workplace.

It is therefore imperative that the University commits itself to produce graduates who are equipped with the knowledge, skills, competencies and attitudes necessary for strong leadership for not only the Caribbean region, but indeed for wherever they may find themselves in the global marketplace.

The introduction of innovative modes of assessment has to be aligned with a curriculum that places students at the centre. In so doing, the University will ensure that:

1. its curriculum is relevant, current, achieves expected quality and responds to the demands of a changing world;
2. learning objectives and outcomes in courses and programmes address competencies such as problem-solving, critical and creative thinking, critical analysis and evaluation, independent learning and team work;

3. teaching/learning approaches stimulate and motivate students to think critically, analytically and creatively, encourage independent thought and the expression of new ideas;

4. a variety of innovative and flexible assessment methods are used to measure the skills, competencies and attitudes required of an undergraduate student i.e. those outlined in 2 and 3 above;

5. criterion-referenced tests are designed to determine student performance to agreed standards – where test scores are on an absolute basis not relative to the performance of other students;

6. a mix of formative and summative evaluations will be used to inform both students and lecturers.

The assessment of learning should reflect accurately a combination of a student's abilities, achievements, skills and potential. Consequently, all forms of assessment should be valid, reliable and fair. In its efforts to continue to provide the intellectual space in which teaching and research can find a common place, the University is creating a new path where teaching and research feed the same set of processes that lead to successful outcomes; where curricula are designed to co-ordinate the teaching, learning and assessment strategy, and where assessment processes do not drive curricula, but facilitate learning.

Overview of the handbook

This handbook addresses some aspects of assessment that deal principally with procedures used *to estimate student learning.* This is done with the understanding that estimating student learning is but one of the major goals of assessment, the other being *student development. These concepts and all other topics included in this booklet are explored more fully in various IDU assessment workshops.*

The two primary purposes of assessment can be discerned in the following ways in which the results of assessment may be used:

- To pass or fail a student
- To grade or rank a student

- To diagnose a student's strengths or weaknesses
- To evaluate a course's strengths or weaknesses
- To provide feedback to students
- To provide feedback to lecturers
- To motivate students
- To motivate lecturers
- To provide a profile of what the student has learnt
- To predict success in future courses
- To predict success in employment
- To select for future courses
- To select for future employment
- To present the course as credit worthy to other institutions and employers

Too often an assessment is used for many different purposes without considering the underlying assumptions of the assessment. What is appropriate for student feedback may not be appropriate for determining competence. Hence it is important to ask yourself these three questions before you draw up an assessment plan:

1. What is this assessment for?
2. Who is it for?
3. What is the context?

Some Common Assessment Terms

The following are forms of assessment related to specific purposes:

Formative – to provide feedback which is used to improve teaching and learning

Summative – to determine achievement at the end of a course, unit of study or instruction.

Criterion-referenced – an approach to assessment in which performance is interpreted according to the level of mastery of a well defined assessment domain.

Norm-referenced – an approach to assessment in which performance is interpreted according to how well the student's performance compares with that of others who were assessed.

Assessment and Course Development

An integral part of course design is the alignment of assessment strategies with your predefined learning outcomes or objectives. Learning outcomes drive both the *WHAT* and *HOW* of course assessment.

Table 1 illustrates the link between learning outcomes or objectives and the type of measures that can be employed in assessing those objectives. In planning your course signify your intentions about the learning outcomes (levels of competence) using Bloom's taxonomy. Match the outcome measures with your intentions or objectives.

Your teaching strategies or the learning opportunities that you provide must be congruent with the objectives and the measures of your learning outcomes.

TABLE 1: BLOOM'S LEARNING OUTCOMES

Adapted from Bloom et al, 1956 – *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain*

Learning Outcomes	Evidence of Outcome	Terms for Measuring Outcome in Test Question
Knowledge	Knows common terms Knows specific facts Knows methods and procedures Knows basic concepts in course Knows principles	*Define*, *describe*, *identify*, *label*, *list*, *match*, *name*, *outline*, *reproduce*, *select*, *state*
Comprehension	Understands facts and principles Interprets verbal material Interprets graphs and charts Translates verbal material to mathematical formulae Estimates future consequences implied in data Justifies methods and procedures	*Convert*, *defend*, *distinguish*, *estimate*, *explain*, *extend*, *generalise*, *give examples*, *infer*, *paraphrase*, *predict*, *rewrite*, *summarise*
Application	Applies concepts and principles to new situations Applies laws and theories to practical situations Solves mathematical problems Constructs graphs and charts Demonstrates correct use of a method or procedure	*Change*, *compute*, *demonstrate*, *discover*, *manipulate*, *modify*, *operate*, *predict*, *prepare*, *produce*, *relate*, *show*, *solve*, *use*
Analysis	Recognises unstated assumptions Recognises logical fallacies in reasoning Distinguishes between facts and inferences Evaluates the relevance of data Analyses the organisational structure of a work (art, music, writing)	*Break down*, *diagram*, *differentiate*, *discriminate*, *distinguish*, *identify*, *illustrate*, *infer*, *outline*, *point out*, *relate*, *select*, *separate*, *subdivide*
Synthesis	Writes a well-organised theme Writes a creative short story (or poem, or piece of music) Proposes a plan for an experiment Integrates learning from different areas into a plan for solving a problem Formulates a new scheme for classifying objects (or events or ideas)	*Categorise*, *combine*, *compile*, *compose*, *create*, *devise*, *design*, *explain*, generate, *modify*, *organise*, *plan*, *rearrange*, *reconstruct*, *relate*, *rewrite*, *summarise*, *tell*, *write*
Evaluation	Judges the logical consistency of a written passage Judges the adequacy with which conclusions are supported by data Judges the value of a work (art, music, writing) by use of internal criteria Judges the value of a work (art, music, writing) by use of external criteria	*Appraise*, *compare*, *conclude*, *contrast*, *criticise*, *describe*, *discriminate*, *explain*, *justify*, *interpret*, *relate*, *summarise*, *support*

VALIDITY AND RELIABILITY

Two important technical qualities of an assessment instrument are **validity** and **reliability**. These factors must be considered in using any form of assessment.

Validity

Validity is concerned with the proper interpretation and use of test scores.

Content validity is the first priority of any assessment. It is a measure of the degree to which the assessment contains a representative sample of the material taught in the course. Content validity must be judged according to the objectives of the assessment. Therefore in approaching any assessment the first question you must ask is: **What are the objectives of the course?**

Reliability

Reliability is concerned with the consistency of the results produced by the assessment instrument. It is a measure of the extent to which the test scores are free from errors of measurement.

Theoretically, a reliable test should produce the same result if administered to the same student on two separate occasions, provided the conditions are the same and there is adjustment for prior learning and growth.

The other key component in determining the reliability of a test is the **consistency of the marking**. The absence of consistency is responsible for the unacceptable levels of reliability in most forms of direct assessment (performance, authentic, practical) and of written essay type tests.

Improving Validity and Reliability

Validity can be improved by:

- carefully matching a test with the course objectives and teaching methods;

- increasing the selection of objectives and content areas included in any given test;
- using methods that are appropriate for the objectives specified;
- employing a range of test methods;
- ensuring adequate security and supervision to avoid cheating;
- improving the reliability of the test

Reliability can be improved by:

- ensuring that questions are clear and suited to the level of the students;
- checking to make sure that test time limits are realistic;
- writing test instructions that are simple, clear and unambiguous;
- developing a marking scheme of high quality (e.g. explicit and agreed criteria,
- checking of marks, using several skilled markers);
- keeping choices within a test paper to a minimum;
- increasing the number of questions, question type, observations or examination time, when using highly subjective test methods

PLANNING FOR FAIR ASSESSMENT: TABLE OF SPECIFICATION [TOS]

The table of specification or test blueprint is a strategy used to enhance the validity of the assessment process in a course. Essentially it is a two-dimensional table which lists the topic/themes (content of a course) along with the objectives levels targeted by each topic/theme. Using the layout, the kinds of assessment tasks, questions or strategies congruent with the course intent can be better decided upon. Appropriate weightings based on factors such as time spent, difficulty level and nature of competencies to be developed are applied to the topics/themes. The visual presentation of the table gives the course lecturer a clear view of the coverage and strategies afforded by the assessment.

The following hierarchy is based on Grasha's taxonomy of educational objectives, which although reflective of Bloom's, is thought to be more appropriate for university level teaching.

The hierarchical educational levels are:

- Basic knowledge and understanding
- Applications
- Critical thinking
- Problem-solving

Using this taxonomy Grasha defined a table of specification for a hypothetical course which is illustrated in Table 2. The table illustrates that there is a spread of competence development by topic and levels required of different topics. A mosaic of competence development and assessment is targeted based on the distinction between topics that focus on lower competences – knowledge and comprehension, and higher competences – critical thinking and problem solving.

Table 3 provides examples of the types of questions which address the various levels of competence in Grasha's taxonomy. Appendix III provides a template for a table of specification based on this taxonomy.

TABLE 2: ELEMENTS OF A TEST BLUEPRINT OR TOS FOR A HYPOTHETICAL COURSE [Grasha, 1996]

COURSE CONTENT	OBJECTIVES & QUESTIONS			
	Knowledge/ Understanding	**Applications**	**Critical Thinking**	**Problem Solving/ Decision Making**
Id, Ego, Superego	Define each term		What are the assumptions about the mind that each term signifies?	Think of a sociopath. What part is dominant? Weakest?
Oxygen		Describe its molecular structure		Appraise the role it plays in aging
Friction		Predict what will happen when oil is applied to a rough surface		Develop two alternative solutions for reducing friction in a car engine
American Revolution	List the major events leading up to the *shot heard around the world*		Prioritise events in terms of their importance for starting the war and cite evidence to justify your ideas	
Shakespeare	What are three characteristics of his style of writing?		Critique the conclusions of some scholars that he did not write his works.	Identify criteria that would be needed in order to decide whether Shakespeare wrote Hamlet

TABLE 3: COURSE OBJECTIVES AND EXAMINATION QUESTIONS/ITEMS
[Grasha, 1996]

Basic Knowledge & Understanding

The ability to identify and recall content and organise and select facts.

Exam questions within this category would ask students to:

Categorise, convert; compare; contrast, defuse; describe the significance of …; differentiate, explain, generalise; give examples of …; identify, interpret the meaning of …; list; name; organise; outline; repeat; summarise the major points in …

Applications

The ability to use various facts, ideas, concepts, and principles to discuss and/or produce a specific outcome.

Exam questions within this category would ask students to:

Apply; demonstrate; design; develop; illustrate how; model; modify; reconstruct; schedule; use information to estimate or predict what will happen when …; and prepare a [chart, outline, program] using the content

Critical Thinking

The ability to analyse situations, synthesise information, identify assumptions, form valid interpretations and conclusions, and evaluate the adequacy of evidence to support positions.

Exam questions within this category would ask students to:

Analyse; appraise; assess the validity of …; conclude; critique; deduce; develop support for; evaluate the evidence for …; examine the other side of …; identify assumptions; identify arguments made by …; infer, integrate; interpret; justify; paraphrase …; prioritise; rate the appropriateness of …; synthesise

Problem, Solving & Decision Making

The ability to analyse and define problems, generate alternative solutions, and use criteria in order to select appropriate solutions or to make decisions.

Exam questions within this category would ask students to:

Brainstorm ideas for, choose, compute, define the problem in …; develop alternative solutions for….; develop an appropriate representation of the elements in the problem of ….; identify the critical elements in the problem of…; identify relevant criteria for selecting ….; plan; solve; use criteria in order to select….; use appropriate heuristics/formal rules to ….

FAIR MARKING

The issues of validity and reliability underscore the need to not only make assessments public, but also fair to the students. Further, the standards for developing test items and tests refute the notion that assessments and other aspects of the teaching/learning are *ad hoc* and arbitrary. The teaching/learning discipline is replete with strategies and techniques that provide structures and systems by which those who engage in teaching ought to practise. The recommended approach to assessment is one that focuses on fairness, validity and reliability.

The following strategies are key to satisfying the demand for high validity and reliability.

Task analysis

This is to determine the content and ability requirements of the task.

- Identify **all** areas of response required of the question
- Highlight areas that are essential and complex/challenging

Criteria definition

- Select areas of significance (those areas that indicate achievement of tasks)
- Decide on the appropriateness of using either a mark scheme or a set of rubrics

Marking Scheme (See page 11)

- Specify content areas/concepts/processes expected or anticipated in students' responses
- Allot marks to content specified/consider weighting
- Consider factors for reducing error (increasing reliability e.g. spread of marks, standardisation process)
- Consider how marking will proceed, for example, use table/cottage marking where large numbers of candidates are involved

Scoring rubric (See page 13)

- Select areas of significance/criteria (those areas that indicate achievement of tasks)
- Describe levels of performance
- Assign marks along a continuum
- Consider weighting
- Obtain consensus
- Review scoring scheme
- Consider how marking will proceed including standardisation processes

Threats to Fair Assessment Practices

Threats to fair assessment practices come from a variety of sources which are generally underpinned by insufficient knowledge of theories and principles involved in assessment. Threats to some aspects of assessment are highlighted below. Immediate attention to these can significantly improve the quality of current assessment practices.

Threats to Fair Marking

1. Relevance of mark scheme
2. Unfair allotment of marks or weightings to questions and question sections
3. 'Halo' effects and the potential unreliability of impression marking. Note that the halo effect refers to awarding higher marks to students of whom the lecturer thinks highly

   NB: This is one reason why it is important for the mark scheme to be devised prior to marking and be used to mark each script.
4. Absence of marking guidelines
5. Inadequate skill in marking. Training and the use of high quality mark schemes are useful in developing skill in marking.

Threats to Fair Tests

1. Choices on question papers reduce reliability since each student is writing a different examination from which inferences about their competence are made.

   Where choices are necessary, for example, in a voluminous syllabus, strategies to improve validity and reliability include:

   a. Keeping the number of choices to a minimum

   b. Ensuring that question options are equivalent in terms of skill, level and complexity

Threats to Grade Interpretation

1. Failing to recognise that in testing there is **no true zero point** and that the intervals between grade points are **not equal units**
2. Using norm-referenced interpretations exclusively and in place of criterion-referenced interpretations
3. Not recognising that tests only represent samplings of the domain of the discipline.

MARK SCHEMES

Mark schemes are guides to allocating marks to student responses to questions. Mark schemes are a strategy aimed at reducing *intra-rate*r and *inter-rater* errors in assessment. Mark schemes thus allow the same individual to allocate marks to successive student responses in a similar manner (intra-rater) or that successive markers (for example, first and second examiners) allocate marks to the same response in the same way. This strategy helps to reduce error in assigning marks and thus improve the reliability of the examination.

Mark schemes are not to be confused with model answers. They indicate the components in a response to which marks have been assigned. The clearer the requirements for the award of marks the greater the likelihood of reducing error. If a large number of marks are assigned to a part of a question, there should be some guidelines as to how these marks are to be distributed.

Mark schemes can allow for general evidence of competence, such as grammar and spelling and organisation of material, but this must be made known to the students before preparation of the examination.

For open-ended responses, mark schemes should be composed to accommodate unexpected and alternative responses. This is important to facilitate creative and critical thinking.

When to construct a mark scheme?

Mark schemes should be composed at the time of construction of the question. They help clarify the intent of the question, minimise ambiguity and allow for a fairer allocation of marks. It is advisable that the mark allocations based on the mark scheme be used to guide the total marks allocated to each question on the examination paper, rather than, for example, use the traditional but arbitrary method of assigning equal marks to questions on essay papers when the questions on a paper are often quite variable in terms of knowledge and competencies required of the students.

Good mark schemes reduce the arbitrariness of impression marking and 'halo' effects.

The example used in Figure 1 shows a mark scheme that can be applied to a simulated lecture and objectives for part of an Ecology course.

THE LECTURE

THE GREENHOUSE EFFECTS

That 'greenhouses' trap heat radiating from the earth's surface is a natural process, much like a greenhouse traps sunlight. Usually this would not be a problem, but because people are adding more gases to the atmosphere than occur naturally, some scientists think we may be significantly changing the atmosphere's chemistry. This is the 'Greenhouse effect' or simply put this is the effect of the excessive accumulation of 'greenhouse gases'.

Carbon dioxide (CO_2) is by far the worst offending greenhouse gas – mainly because of the sheer quantities of it that human activities generate. Fossil fuel combustion and deforestation are steadily increasing the percentage of CO_2 in the atmosphere. There are disagreements over what the effects of greenhouse gases will be - but many scientists acknowledge that there is a good probability that the increasing level of CO_2 and other greenhouse gases may be causing the earth's climate change.

While scientists disagree on the exact outcome of global climate change many theories have widespread support. One theory says that since CO_2 traps heat, earth's temperature may increase, causing global warming. Some scientists say that there could be up to 10°F (5.6°C) increase overall in world temperatures within the next 50 years. If this happened, whole agricultural regions, such as the Mid West, United States, could become too hot and dry for many crops. An increase in temperature could also cause the sea levels to rise enough to flood coastal cities. For many wildlife species already stressed by habitat loss and pollution, global climate change could be a death sentence.

Climate change due to 'Greenhouse effect' can have a significant deleterious reaction on the composition, resilience or productivity of natural and managed ecosystems or on the operation of socio-economic systems or on human welfare.

OBJECTIVES

The student will be able to:

1. explain the 'greenhouse effect'
2. outline how the 'greenhouse effect' affects life on planet earth.

ESSAY QUESTION

Explain the process of the 'greenhouse effect' and discuss how this process is alleged to affect climates and sea levels. (20 marks)

MARK SCHEME

Gases regulate earth temperature **(2)**
Greenhouse effect – gases retained, accumulated – trapped **(3)**
Increase in gases because of man and industrialisation **(3)**
Types of gases **(2)**
Temperature rise – estimated increase **(1)**
Sea-level rise **(1)**
Melting of glaciers and polar caps **(1)**
Grammar, spelling **(2)**
Organisation (sequence of paragraphs, argument) **(3)**
Additional, correct researched information **(2)**

Adapted from: Brown, University of Sheffield

Figure 1. Example of a mark scheme as applied to an essay in the context of a particular course topic

RUBRICS

What are rubrics?

Rubrics are guides to assessing levels of performance of significant tasks. They are student-centred and standards-driven and thus describe benchmarks or standards, promote self assessment and allow feedback ('feedforward'). Rubrics address the issue of subjectivity in performance assessment. They assure assessments are objective, fair, and reliable.

Performance-based assessments have a *clear purpose*, *observable aspects*, *means of eliciting the performance* and clearly articulated scoring procedures. They are authentic, that is, grounded in the real world of the student.

Performance–based assessment include a wide range of student work that target specific skills, for example,

- communication – demonstrated through reading, writing, speaking, listening
- psychomotor skills requiring physical abilities
- complex behaviours such as:
  - problem-solving
  - logical thinking
  - critical thinking
  - argumentation
  - decision-making
  - interpretation
  - laboratory/field work
  - collaborative working
  - communication
  - time management
  - information technology

Rubric Development

The rubric development process consists of the following:

1. Task analysis – deciding what students are required to do
2. Criteria definition – identifying significant tasks
3. Identification of performance levels for each criterion – benchmarking or determination of standards
4. Definition of descriptors – providing comparative descriptions of levels of performance defined for each criterion
5. Assignment of marks to criterion levels along a continuum

Tables 4, 5 and 6 provide examples of rubrics. Table 4 shows a detailed rubric for assessing essay writing. Table 5 shows a rubric in which only the performance standard for each criterion has been defined. Table 6 shows a rubric for assessing undergraduate project work. These rubrics are generic for wider application. Rubrics can have more specific criteria usually determined by consensus of experts in the discipline. Consensus is a critical element in building appropriate and effective rubrics.

TABLE 4: GENERIC RUBRIC FOR ASSESSING ESSAY WRITING

CRITERIA	PERFORMANCE LEVELS					Total 30
	5	4	3	2	1	
Introduction & Interpretation	Introduction shows a sound grasp of the assignment brief and provides a clear outline of the scope of the essay; key issues stated; provides a framework for answering the question	Introduction shows a reasonably good grasp of the key issues; clear framework for answering the question	Perfunctory or overlong introduction; framework is present but unclear	The topic is not well introduced; scope of the essay poorly defined	Launched straight into the essay with little or no attempt to introduce and define topic; purpose of assignment misunderstood	
Development/ Quality of Argument	Develops logical argument and clearly marshals the relevant ideas and evidence and their strengths and weaknesses	Ideas are logically and clearly developed	Some aspects of argument are weak	Poor or inappropriate organisation and sequencing of material; clear theme or line of argument not developed	Major weaknesses in argument or argument not discernible	
Use of Evidence/ Sources	Critical and wide ranging use of relevant literature backed up by good bibliography; evidence well evaluated; connections have been made between evidence and the frame work of the question	Critical and wide range use of relevant literature backed up by a good bibliography	Some of the sources and material are relevant to the theme/topic	Inclusion of some irrelevant or unlinked material	Little evidence of supporting reading that shows inadequate preparation	

TABLE 4: GENERIC RUBRIC FOR ASSESSING ESSAY WRITING (Cont'd.)

CRITERIA	PERFORMANCE LEVELS					Total 30
	5	4	3	3	3	
Understanding of Topic	Well argued exposition of the topic; all main issues have been explored and evaluated	Reasonable grasp of the topic and most of the main issues have been explored; evidence of critical evaluation	Most of the main issues in relation to the topic have been explored; evidence of some analysis	Few issues have been explored; little evidence of analysis and critical evaluation	The work is rather descriptive; tendency to accept views with little analysis/criticism	
Presentation	Correct and effective use of English; accurate referencing	No problems with grammar; minor spelling errors; reasonably good use of English	Few spelling errors; some minor errors in grammar, syntax	Several intrusive spelling errors and many errors in grammar, syntax	So many intrusive errors in spelling grammar and syntax that reading is difficult	
Conclusion	The conclusion draws together the strands of the argument; it creates a coherent perspective on the question and indicates links to related deep issues; it is based on the argument and evidence presented	A good clear concluding section that draws together the various important points made	There is a rather brief and formalised conclusion to the essay	Conclusion merely rephrases the introduction	Little or no conclusion or conclusion is not based on the argument and evidence in the essay	

TABLE 5: SAMPLE ASSESSMENT/FEEDBACK SHEETS

ESSAY ASSESSMENT SHEET						
Name:				**Marker:**		
Date:		**Date back:**			**Mark:**	
Marker's general comment on the essay						
ASSESSMENT CRITERIA	Interpretation of criterion	**EXCELLENT 5**	**GOOD 4**	**SATISFACTORY 3**	**NEEDS MORE WORK 2**	**NEEDS MUCH MORE WORK 1**
INTRODUCTION TO THE ESSAY	Grasp of the topic					
	Identification of the issues					
DEVELOPMENT OF QUALITY OF ARGUMENT	Logical argument					
	Insight and originality					
	Constructive critical analysis					
USE OF EVIDENCE/ SOURCES	Relevant ideas and evidence					
	Strengths and weaknesses of evidence					
	Connection between evidence and question					

TABLE 5: SAMPLE ASSESSMENT/FEEDBACK SHEETS (Cont'd.)

ESSAY ASSESSMENT SHEET						
Name:			**Marker:**			
Date:		**Date back:**			**Mark:**	
Marker's general comment on the essay						
ASSESSMENT CRITERIA	Interpretation of criterion	**EXCELLENT 5**	**GOOD 4**	**SATISFACTORY 3**	**NEEDS MORE WORK 2**	**NEEDS MUCH MORE WORK 1**
UNDERSTANDING TOPIC	Main issues identified					
	Main issues explored and evaluated					
PRESENTATION	Correct use of grammar and syntax					
	Correct spelling					
	Overall presentation					
CONCLUSION	Draws strands of argument					
	Coherent perspectives on the question					
	Links to related deep issues					
	Based on argument and evidence					

TABLE 5: SAMPLE ASSESSMENT/FEEDBACK SHEETS (Cont'd.)

ESSAY ASSESSMENT SHEET		
Name:	**Marker:**	
Date:	**Date back:**	**Mark:**
Marker's general comment on the essay		
SPECIFIC ASPECTS OF YOUR ESSAY that the marker likes	**SPECIFIC ASPECTS OF YOUR ESSAY that need more work**	
1.	1.	
2.	2.	
3.	3.	
4.	4.	
5.	5.	

TABLE 6: RUBRIC FOR ASSESSING AN UNDERGRADUATE RESEARCH PROJECT
Source: IDU, St. Augustine workshops

EXCELLENT (A)	GOOD (B)	SATISFACTORY (C)	WEAK (D)
CRITERIA			
Presentation & formatting **(5 points)** Adheres to requirements: cover page, margins, line spacing, font, pagination, paragraphing, appendices, illustrations, graphs/charts, tables, labelling and numbering.	**(4-3 points)** Adheres to most requirements	**(2-1 points)** Adheres to few requirements	**(<1 point)** Does not adhere to requirements
Abstract **(10-8 points)** Concise summary: background, significance of research/study, objectives, methodology, major findings, and conclusions; all key words; within word limit	**(7-6 points)** Generally well-written summary except for shortcomings in one or two areas	**(5-4 points)** Vague/imprecise summary; some findings and keywords omitted; insufficient attention to word limit	**(<1 points)** Very inadequate/imprecise summary; no keywords; text beyond the word limit
Introduction and objectives **(10-8 points)** *Introduction:* Clearly defines scope of study; shows significance of research; provides clear justification for study *Objectives:* Clearly stated; relate to the hypothesis/problem statement; appropriately placed after the introduction	**(7-6 points)** *Introduction:* Defines scope of study; shows significance of research; provides limited justification for study *Objectives:* Well-stated and shows some relationship to the hypothesis/ problem statement; appropriately placed after the introduction	**(5-4 points)** *Introduction:* Some material not relevant/unconvincing in showing the significance of the work; background material sketchy *Objectives:* Some indication of the objectives given but unconvincing as to relationship with hypothesis/problem statement	**(<4 points)** *Introduction:* Much of the material irrelevant; significance of research/ study not shown; vague description of background e.g. no connection of literature to research *Objectives:* Not clearly stated and not related to the hypothesis/ problem statement; inappropriate placement of objectives e.g. not after introduction

TABLE 6: RUBRIC FOR ASSESSING AN UNDERGRADUATE RESEARCH PROJECT (Cont'd.)
Source: IDU, St. Augustine workshops

EXCELLENT (A)	GOOD (B)	SATISFACTORY (C)	WEAK (D)
CRITERIA			
Literature Review/Theory **(10-8 points)** Relevant and current; logical development of supporting argument; wide range of sources; relevant, appropriately-cited quotes.	**(7-6 points)** Relevant and current; supporting argument not very well developed; range of sources; appropriately cited quotes.	**(5-4 points)** Relevant material but not effectively organised to support argument; many references omitted; some theory included, but application not clearly specified	**(<4 points)** Much of literature irrelevant; material insufficient to support argument; verbose, but shows little evidence of critical analysis; relevant theory not included/ developed.
Methodology/Process Description **(10-8 points)** Accurate, clear, easy to follow; repeatable where appropriate; describes equipment and materials where relevant.	**(7-6 points)** Generally clear, and easy to follow; repeatable where appropriate; describes equipment and materials where relevant.	**(5-4 points)** Unclear in key areas/some steps are missing, making it difficult to repeat the procedure/follow process.	**(<4 points)** Imprecisely/inaccurately written; difficult to follow or repeat, no/inadequate description of equipment and materials/process.
Results **(20-16 points)** Accurate and defensible; logically displayed; appropriate use and description of graphs and tables; strong evidence of data analysis including statistics where appropriate.	**(15-12 points)** Generally accurate and defensible; logically displayed; most graphs and tables appropriately used and described; solid evidence of data analysis including statistics where appropriate.	**(11-9 points)** Results are at times carelessly/ inappropriately displayed; description sometimes difficult to follow; data analysis presented but incompletely/inadequately analysed.	**(<9 points)** Many inaccuracies/inconsistencies; not logically presented; no description of tables/graphs or inappropriate use; no data analysis; missing/unaccounted for information.

TABLE 6: RUBRIC FOR ASSESSING AN UNDERGRADUATE RESEARCH PROJECT (Cont'd.)
Source: IDU, St. Augustine workshops

EXCELLENT (A)	GOOD (B)	SATISFACTORY (C)	WEAK (D)
CRITERIA			
Discussion/Conclusion/Recommendation **(20-16 points)** Logical organisation of thoughts; adequate discussion of all aspects; strong evidence of critical analysis; new insights, strong comparison of results with existing literature; limitations of study; points to future research	**(15-12 points)** Good organisation of thoughts; adequate discussion of almost all aspects; evidence of some critical analysis; strong comparison of results with existing literature; limitations of study; points to future research	**(11-9 points)** Fairly good organisation; some evidence of analysis of work; results are insufficiently referenced; some effort made to describe the significance; some attempt made to relate to literature.	**(<9 points)** Not very well organised; analysis bears little relevance to results; some/most interpretations flawed, little/no reference to existing literature.
Referencing/Documentation **(5 points)** Accurate and consistent use of recommended style guide; appropriate and complete in-text citations; complete compilation of literature cited/references/ bibliography	*(4-3 points)* Very minor inconsistencies/ inaccuracies in use of style guide; few in-text citations omitted; almost complete compilation of literature cited/references/ bibliography	(2 points) Some inconsistencies/inaccuracies in use of style guide; few in-text citations omitted/ unconventional; some omissions of literature cited/references/ bibliography	**(<2 points)** Major inconsistencies/inaccuracies in use of style guide/style guide ignored; unconventional in-text citations; numerous omissions of literature cited/references/ bibliography
Style **(5 points)** Paragraph structure; correct use of terms/ diction; grammar, spelling, punctuation; sentence variety.	**(4 points)** Errors in one area only	**(3-2 points)** Error in two/three areas only	**(<2 points)** Errors in more than 3 areas

TABLE 6: RUBRIC FOR ASSESSING AN UNDERGRADUATE RESEARCH PROJECT (Cont'd.)
Source: IDU, St. Augustine workshops

EXCELLENT (A)	GOOD (B)	SATISFACTORY (C)	WEAK (D)
CRITERIA			
Overall presentation (5 points) Unity. coherence, emphasis, completeness; clarity	(4 points) Shortcomings in one area only	(3-2 points) Shortcomings in two/three areas only	(<2 points) Shortcomings in more than three areas.

Adaptations may be made to this rubric to suit specific purposes.
The rubric should be given to the students along with the project assignment

Note: The criteria selected and weightings were agreed to by the rubric developers based on their requirements. Rubrics make public the emphases and requirements of the assessor and provide opportunity for a collective approach to defining standards of performance.

The above rubric was initially produced by staff who attended IDU, St. Augustine's Rubrics workshops in 2004 - Prof. Mellowes and Dr. Neela Badrie.

The rubric was reviewed and updated at the 2005 Rubrics workshop by Ms. R. Walker, Dr. G. Sirju-Charran, Dr. A Lennon, Dr. P. Collins, Dr. R. Pingal, Dr. A. Edwards-Henry

MULTIPLE-CHOICE: GUIDELINES TO ITEM WRITING

Multiple-choice is one form of objective type items or questions. An objective item or question is one for which there is only ONE correct answer. The answer is specific and unambiguous. Other forms of objective type items include: true/false, fill-in-blanks, matching items, and simple questions.

Characteristics of Multiple-choice Items

A multiple-choice item is distinguished by its structure. It usually consists of a given task, stimulus material or question statement which is called the **stem,** followed by a series of **options** or alternative responses from which the student is required to select the correct response or **key.** The incorrect options are called **distractors.** Distractors must be plausible to encourage the student to think carefully in arriving at the key. Distractors that are obviously incorrect are of little value. The following example taken from the discipline of Social Sciences, illustrates these features.

Example:

Individuals who are androgynous

A. are masculine if female, but feminine if male
B. are as masculine as they are feminine
C. display behaviours and abilities consistent with their gender
D. score high on measure of both masculinity and femininity

Note that the example illustrates the following:

Stem – in the form of a statement that seeks to define a term
Options – four alternative responses of which one is the correct response. Each option statement grammatically completes the stem.
Key – the correct answer which is B
Distractors – the incorrect options which are A, C and D.

Types of Multiple-choice Items

There are several formats in which multiple-choice items occur. These are as follows:

- Direct question
- Incomplete statement
- Matching items

- Multiple completion
- Assertion/reason

A high quality test should have a range of item types and should target a range of levels of objectives.

Examples illustrating the formats of these different types of multiple-choice items are provided below.

Direct question

In this format the task is posed in the form of a direct question.

Example:

In the area of physical science, which of the following definitions describes the term "polarisation"?

A. Excitation of electrons by high frequency light
B. Ionisation of atoms by high temperatures
C. Interference of sound waves in a closed chamber
D. Separation of electric charges by friction
E. Vibration of transverse waves in a single plane **Key: E**

Incomplete statement

In the incomplete statement the task is presented in the form of a statement which is correctly completed by the key. All options must grammatically correctly complete the statement.

Example:

Test scores are difficult to interpret because:

A. norm and criterion-referencing are sometimes indistinguishable.
B. the zero-point indicates that there is no achievement at all.
C. of a lack of a zero-point and equal units.
D. the distribution is not always normal. **Key: C**

Matching items

For matching items, a series of questions or statements on the same theme or topic is related to the options provided. The student/candidate is required to match each statement with the correct option for that statement. This format allows for more efficient coverage of the knowledge areas of the syllabus.

Example

Items 1 to 3 refer to the following terms:

A. correlation
B. variance
C. standard error
D. normal distribution

Match the terms listed above with the descriptions given below. You may use each option once, more than once or not at all.

1. Arithmetic average of the squares of the deviation of scores from the mean. **Key: B**
2. Measure of the relationship between two sets of scores. **Key: A**
3. Spread of scores that produces a bell-shaped curve. **Key: D**

Multiple completion

The multiple completion format allows for more complex thinking skills where the student must be more discerning. The format can be used where there is great difficulty finding plausible descriptors or where there are several components to a correct response.

Example:

i. Mean
ii. Median
iii. Standard deviation
iv. Variance

Which of the above are measures of variability?

A. i only
B. i and ii only
C. ii and iii only
D. iii and iv only

Key: D

Assertion/Reason

The assertion/reason is a sophisticated form of multiple choice question in which the stem comprises two statements linked by *because*. The student is required to determine whether the statements are correct or not as statements in their own right and/or whether one statement - the reason – accounts for the other statement – the assertion. This form usually targets assessment of complex and higher order thinking skills.

Example:

Assess the sentence in italics according to the criteria given in the options below:

"The United States took part in the Gulf War against Iraq BECAUSE of the lack of civil liberties imposed on the Kurds by Saddam Hussein's regime."

A. The assertion and the reason are both correct, and the reason is valid.
B. The assertion and the reason are both correct, but the reason is invalid.
C. The assertion is correct but the reason is incorrect.
D. The assertion is incorrect but the reason is correct.
E. Both the assertion and the reason are incorrect.

Key: B

Guidelines for Writing Multiple-choice Items

Writing multiple-choice items is a highly skilled activity which is most effectively accomplished when there is opportunity for items to be critiqued both by the item writer and his or her peers. The guidelines below provide a checklist which individual item writers and colleagues can use as a starting point for critiquing items. As with all other question writing techniques the items must be aligned to course objectives and learning opportunities provided during the course.

The following list provides further specific guidelines to be used when preparing multiple-choice items.

1. The essence of the problem should be in the stem
2. Avoid repetition of words in the options
3. Avoid superfluous wording
4. When the incomplete statement format is used, options should be at the end of the statement
5. Arrange options as simply as possible
6. Arrange options in alphabetical or numeric sequence. Avoid contriving to allow the key to fall in any particular order. Arranging in alphabetical or numerical order will allow successive keys to be at random.
7. All distractors should be plausible and homogenous

8. Avoid making the key consistently longer or more complex than the distractors
9. Avoid irrelevant clues to the correct answer
10. Ensure that there is only **one** correct or best answer to every item
11. Avoid using "all of the above" and "none of the above" as options. "None of the above" should never be used as it is conceptually the antithesis of testing when used as the key. While it identifies the incorrect things that the student recognises, there is no assurance that he/she knows what is correct and can identify it among a range of options.
12. Use four or five options, but be consistent in the number of options used in a particular test.
13. Avoid overlapping options

Advantages and Disadvantages of Multiple-Choice Items/Tests

Like all other forms of assessment there are advantages and disadvantages to the use of multiple-choice items. Note that multiple-choice tests are more easily controlled to show desirable features and to improve reliability. However, mastery in writing items, putting the test together and analysis of test results are also areas in which levels of knowledge and skill are required.

Advantages

1. **Versatility** – Multiple-choice items are adaptable to the measurement of a wide variety of learning outcomes, from knowledge of facts through analysis and interpretation of information to reasoning, making inferences, solving problems, and exercising judgment.
2. **Efficiency** – Because of the large number of items that can be posed in a given length of time multiple-choice items permit wide sampling and broad coverage of the content domain. This increases validity and potential reliability of such tests.

3. **Scoring accuracy and economy** – Expert agreement on the correct answer to multiple-choice items is easy to obtain, and scoring keys can be economically applied by machine or clerical assistants.
4. **Reliability** – Consistency in scoring and wide sampling of content provide test results that are generalisable to the domain of interest.
5. **Diagnosis** – Patterns of incorrect responses can provide diagnostic information about the learning of individual students or groups.
6. **Control of difficulty** – The level of difficulty of a test can be increased or decreased by adjusting the degree of similarity among the options to the items.
7. **Reduction of guessing** – In comparison with two-choice (e.g., true-false) tests, guessing is reduced by multiple choice items.
8. **Freedom from response sets** – Multiple choice items are relatively uninfluenced by response sets, such as a tendency to answer "true."
9. **Amenable to item analysis** – Multiple choice items are amenable to item analysis, by means of which they can be improved.

Disadvantages

1. Multiple-choice tests can be **difficult and time consuming to write.** The construction of plausible distractors can be especially difficult. The quality of the test is therefore dependent on the item-writing skill of the instructor.
2. There is a tendency to write items requiring only factual knowledge rather than those testing higher level skills and understandings.
3. **Performance** on multiple-choice items can be markedly influenced by student characteristics unrelated to the subject of measurement, such as reading ability, testwiseness, and risk-taking
4. **Multiple-choice** items are subject to unintentional 'clueing'.

5. Multiple-choice items do not measure ability to organise and express ideas as in short answer responses or essays.

Multiple-Choice Writing References and Resources

Listed below are some textbooks and/or web sites that will provide additional help with developing multiple-choice items/questions

Text Resources

Airasian, Peter W. (1994). Classroom assessment. New York: McGraw-Hill.

Carey, Lou M. (1988). Measuring and evaluating school learning. Boston: Allyn and Bacon, Inc.

Ebel, Robert L. & Frisbie, David A. (1986). Essentials of educational measurement. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Gallagher, Jo D. (1998). Classroom assessment for teachers. Upper Saddle River, NJ: Prentice-Hall Inc.

Haladyna, Thomas M. (1999). Developing and validating multiple-choice test items. (2nd edition). Mahwah, NJ: Lawrence Erlbaum Associates

Hopkins, Charles D. and Antes, Richard L. (1990). Classroom measurement and evaluation. Itasca, IL: F. E. Peacock

Kubiszyn, Tom and Borich, Gary. (2000). Educational testing and measurement. New York: John Wiley and Sons

Linn, R. L. & Miller, D. M. (2005) Measurement and assessment in teaching. 9th ed. New York. Pearson Prentice Hall.

McMillan, J. H. (1997). Classroom assessment: Principles and practice for effective instruction. Boston: Allyn & Bacon.

Nitko, A. J. (2001). Educational assessment of students. Upper Saddle River, NJ: Merrill Prentice-Hall

Oosterhof, Albert. (2001). Classroom applications for educational measurement. Upper Saddle River, NJ: Merrill Prentice Hall.

Wiersma, William & Jurs, Stephen G. (1985). Educational measurement and testing. Boston: Allyn and Bacon, Inc.

Websites:

Writing multiple-choice test items. ERIC AE Digest:
www.ed.gov/databases/ERIC_Digests/ed398236.html

More multiple-choice item writing do's and don'ts. ERIC AE Digest:
www.ed.gov/databases/ERIC_Digests/ed398238.html

Constructing classroom achievement tests. ERIC Digest:
ericae.net/edo/ed315426.htm

Samples of Multiple-choice questions from APPSAT and other tests produced by College Board: www.collegeboard.org

http://web.uct.ac.za/projects/cbe/mcqman/mcqappb.html

Adapted from:
http://web.utk.edu/~mccay/apdm/mchoice/mc_h.htm and other websites.

ONLINE ASSESSMENT

Online assessment tools are integrated in most course management or learning systems management software. The value of using online assessment is dependent upon the types of knowledge that will be assessed. One of the benefits of using online tools for assessment is that they can be scored by the computer after input of point value per question by the lecturer/instructor. Online assessments work well for tests that have definite right or wrong answers.

Most course management systems (e.g. WebCT, Moodle) provide a space for a question database. Instructors may supply a number of questions for later selection from the database to create quizzes, surveys, and examinations. These questions can be updated as necessary and questions can be continually added to the database. Questions in the database may be assigned different categories and weighted at different values by the instructor. Categories allow easy retrieval of questions on a particular topic or theme. Five types of test are available in most course management software tools. The types of tests that can be administrated online are multiple-choice including matching items, calculations, short answer questions and essays.

Multiple-choice questions

The interface allows lecturers/instructors to create questions with single or multiple answers. The questions may be presented randomly so that each student is looking at a different question each time. Feedback can be provided to students about why certain options may have been a better choice than others.

Matching items

Matching pairs can be inputted into the quiz maker up to five parallel choices. Choices may be given a weight by the instructor so that each match is not equally valued. Here too, most management tools provide a space for feedback to students about why certain choices are better than others.

Calculations

Questions which require calculations may be corrected by the computer. The instructor needs to input the formula where indicated and include an explanation for feedback where appropriate.

Short question

Short answer questions may be more difficult to mark using online tools. However, if the answer can be limited to one or two words the management tool may be of use for grading. The problem is that exact wording only will count as correct. The interface does provide several areas for input so that you may allow for variations of wording.

Essays

An essay response may be downloaded and uploaded within a learning management system. Rather than writing a full essay response within the exam interface, the essay should be written in Word or a similar processing document.

BIBLIOGRAPHY

Airasian, P. W. (1991). Classroom assessment. New York: Mc Graw-Hill Inc.

Astin, A. (1991). Assessment of excellence: The philosophy and practice of assessment and evaluation in higher education. New York: American Council on Education and Macmillan Publishing Co.

Alexander Astin et.al. (1996), American Association for Higher Education (www.aahe.org for AAHE)

Boud, D. (1988) Developing student autonomy in learning. London: Kogan Page.

Bloom, B. S. (Ed.). (1956). Taxonomy of educational objectives, handbook I: The cognitive domain. New York: McKay

Brown, S. & Glasner, A. (1999) Assessment matters in higher education: Choosing and using diverse approaches. Buckingham: The Society for Research into Higher Education & Open University Press.

Gallagher, Jo D. (1998). Classroom assessment for teachers. New York Prentice-Hall.

Grasha, A. F. (1996). Teaching with style. Pittsburg, PA: Alliance publishers

Harvey, L., & Green, D. (1994). Quality in higher education project: employer satisfaction summary report. Birmingham: University Central England.

Linn, R. L & Gronlund, N. E. (1995). Measurement and assessment in teaching. 5th ed. New Jersey: Prentice-Hall

Linn, R. L. & Miller, D. M. (2005). Measurement and assessment in teaching. 9th ed. New York. Pearson Prentice Hall.

Mehrens, W. A. and Lehmann, I. J. (1973). Measurement and evaluation in education and psychology. 4th edition. Florida: Harcourt Brace Jovanovich.

Messick, S. (1989). Validity. In R. Linn (Ed.). Educational measurement (3rd ed.). NY: American Council on Education.

Popham, W. J. (2005a). Classroom assessment: What teachers need to know (4th ed.). Needham Heights, MA: Allyn & Bacon.

University of Minnesota. (1995). Author.

APPENDICES

APPENDIX I

PRINCIPLES OF GOOD ASSESSMENT PRACTICE

The following are nine principles of good assessment practice as identified By Alexander Astin and his associates at the American Association for Higher Education.

1. Assessment of student learning begins with educational values. Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time
2. Assessment works best when the programmes it seeks to improve have **clear**, **explicitly stated purposes.**
3. Assessment requires attention to outcomes but also and equally to the experiences that lead to those outcomes.
4. Assessment works best when it is ongoing and not episodic.
5. Assessment fosters wider improvement when representatives from across the educational community are involved.
6. Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about.
7. Assessment is most likely to lead to improvement when it is a larger set of conditions that promote change.
8. Through assessment, educators meet responsibilities to students and the public.

Alexander Astin et.al. (1996), **American Association for Higher Education** (www.aahe.org).

APPENDIX II

SOME ASSESSSMENT STRATEGIES: ADVANTAGES AND DISADVANTAGES

Assessment Form	Description	Advantages	Disadvantages
Traditional exams	Unseen written exams most often using a free response essay style.	• Relatively economical • Equality of opportunity – all students have to do the same tasks in the same way within the same timeframe • Author of the work is known • Teaching staff are familiar with the process • Motivate students to engage with the subject matter	• Students get little or no feedback • Badly set exams encourage surface learning • Exam-taking technique is too important • Exams represent only a snapshot of student performance, rather than a realistic indicator of it
Open-book exams	Similar to traditional exams, except that students are allowed to take into the examination room their chosen sources of reference material.	• Less stress on memory • Measure retrieval skills	• Insufficient books – not all students equally equipped • More desk space required – fewer students can be accommodated at a time
Multiple-choice	Select-type objective tests where students are required to choose the correct answer from among options provided	• Greater syllabus coverage • Saves staff time and energy to administer • Easier to mark – can be computer aided • Provides sound test statistics	• Guess factor • Designing multiple-choice questions take time and skill • Risk of impersonators
Essays	Free-response and open ended often used in traditional and open-book exams	• Allow individuality of expression • Reflect depth of student learning • Familiar to students	• Essay writing is an art in itself • Essays take a long time to mark objectively • Syllabus coverage necessarily limited • Great potential for unreliable marking

Assessment Form	Description	Advantages	Disadvantages
Report writing	Extended writing requiring adherence to specific structure according to discipline	• Report writing skills relevant to many jobs • Form end-product of useful learning activities • Great opportunity for students to display their talents	• Undesirable collaboration may occur • Time-consuming for students • Time-consuming for staff
Practical work	Hands-on activity related to particular aspects of theory	Practical work may be critical to the discipline Practical work is learning-by-doing and thus effective	• Often difficult to assess in its own right • Difficult to obtain agreement on assessment criteria • Students often discomfited when observed
Portfolios	Containers of evidence in support of a particular purpose	• Tell much more about a student • Can reflect development/growth • Can reflect attitudes and values as well as knowledge and skill • Students learn to select what best represents their strengths	• Take a long time to assess • Generally more difficult to mark objectively • Sometimes ownership of material included is in doubt.
Oral presentations	Demonstration of communication on a subject matter through speaking	• There is no doubt about who is being assessed • Key skills are assessed • Opportunity for effective collaboration • Students take the role and responsibility seriously	• Difficult to administer in large classes • Could be traumatic for some students • Evidence is transient unless recorded.
Vivas	Viva-voce exams are primarily oral exams used as supplementary evidence of knowledge and competence	• Valuable check on ownership of the material • Help decision making in borderline cases • Can be used to probe student understanding	• Like oral presentations, some candidates never show well in vivas • Actual agenda covered by a viva is usually limited.

Adapted from Phil Race and Sally Brown (1998). The lecturer's toolkit. London: Kogan Page

APPENDIX III

TABLE OF SPECIFICATION (TEST BLUEPRINT) TEMPLATE

COURSE CONTENT	OBJECTIVES & QUESTIONS			
	Knowledge/ Understanding	Applications	Critical Thinking	Problem Solving/ Decision Making
Topic 1				
Topic 2				
Topic 3				
Topic 4				
Topic 5				

[After Grasha, 1996]

APPENDIX IV

ESSAY QUESTION-WRITING GUIDELINES

1. Restrict the use of essay questions to those learning outcomes that cannot be satisfactorily measured by objective items.
2. Formulate questions that will elicit the behaviour specified in the learning outcomes
3. Phrase the questions/items to allow the students to demonstrate the cognitive process to be measured
4. Phrase each question so that the students' task is clearly indicated
5. State specifically what you expect the answer to include
6. Indicate an approximate time limit for each question.
7. Avoid the use of optional questions
8. Incorporate novel material or novel organisation of material
9. Use precise, unambiguous language

APPENDIX V

EXAMPLES OF INNOVATIVE ASSESSMENT STRATEGIES FOR STUDENT DEVELOPMENT

Try these

Include in your course, innovative formative assessment strategies aimed at developing higher order skills, for example:

- *Plans and Drafts* - Review and discuss essay plans and drafts with tutors/other students before formal assessment. This provides a built-in feedback mechanism. (Assign a limited amount of marks to improve efforts).
- *Question writing by students* – Allow students to prepare questions to help them more fully engage with the subject material. This strategy provides an estimate of what students see as important in a course. It also helps clarify criteria and lecturer's intentions. 'Good' questions can be included in examinations.
- *Self-assessment* – This strategy helps clarify criteria for assessment. Allows for the development of reflective skills. (Training of lecturers required in order to make this an effective strategy).