GUIDELINES FOR PAPER SETTERS

EDUCATIONAL ASSESSMENT UNIT

GUIDELINES FOR PAPER SETTERS

A) General Guidelines to the Paper Setter(s)

1. Make sure you have the latest version of the syllabus and you are familiar with the assessment criteria.

2. Work on a Specification Grid. Before and after setting the paper, check that all the test items are based on the respective syllabus and that the items are graded in difficulty.

3. Do not use material reflecting race, ethnic or sex bias.

4. Develop a Marking Scheme alongside the Specification Grid.

5. Check that the duration of the examination is entered correctly on the paper and that the time allotted is sufficient to enable the students complete the paper and revise their work.

6. Proof read the text.

7. Pass on the *finalized* draft of the paper to an external reviser who has to proof read the text again, ensure that no test item is out of syllabus, check that all set tasks are workable (particularly in mathematics and science subjects) and that the paper can be completed in the set time.

8. Make the necessary changes in the examination paper and the marking scheme as advised by the reviser. Proof read the text once again and pass on the paper to the Reviser for the final proof reading.

9. Hand in the Marking Scheme together with the Examination Paper for printing.

10. Examine printed papers for printing defects (e.g. unclear diagrams or pictures) and for any Errata Corrige that may be required.

B) Layout

1. The layout of the paper should be as clear as possible to make it as student friendly as possible. For write-on papers enough space for working or writing must be provided.

2. Instructions to candidates should be clear and unambiguous. They should be presented in bold type.

3. Wherever possible, use a straightforward and consistent format with regular line lengths.

4. Use typesetting features such as bold, italics, indentation or boxes effectively to help candidates focus their attention on the task. (Note: Print in italics may present reading difficulty for young children.)

5. Long complex questions are best split up by the use of subsidiary numbering systems.

6. Structured questions should follow a graded and logical sequence.

7. The information contained on a page should be well structured through the appropriate use of headings and sub-headings. This would help candidates organise text in advance of reading.

8. Check that the diagrams, pictures or photographs used are necessary, helpful and of high quality.

9. Place the text close to the relevant diagrams or pictures to enable the candidates relate the two effectively. Comprehension text and questions should be set on the same page or on adjacent pages.

10. Ensure that marks assigned for each item / exercise / section are clearly indicated on the paper.

C) Sentence Construction

1. Use the simplest language and structure possible to convey clearly and unambiguously the meaning of the question.

2. Split down even relatively short sentences if they contain a lot of condensed information.

3. Do not use the passive if it can be avoided because it can make a sentence impersonal and complex. Avoid also using the conditional form (sentences starting with “if”) and the double negative.

4. Eliminate superfluous words and any abstract and metaphorical language which is not necessary.

5. Make sure that introductory statements in questions contain only the information which is required for answering those questions relevantly.

D) Specification Grids

1. The writing of test items should be guided by a carefully prepared set of test specifications.

2. The specifications describe the achievement domain being measured and provide guidelines for obtaining a representative sample of test tasks.

3. The specification grid (a two-way table) provides assurance that the test will measure a representative sample of the learning outcomes and the subject matter topics to be measured.

4. The specification grid relates outcomes to content and indicates the relative weight to be given to each of the various areas.

5. A specification grid indicates:
   (i) the learning outcomes to be tested
   (ii) the subject matter or content area
   (iii) the assigned weighting to the learning outcomes and content areas in terms of their relative importance

6. The learning outcomes to be tested include (a) recall of knowledge, (b) intellectual abilities or skills (understanding, application, etc) (c) general skills (eg practical, performance, communication), (d) attitudes, interests, appreciations.

7. The following factors are to be considered when assigning relative weights to each learning outcome and each content area.
   (i) the importance of each area in the total learning experience
   (ii) the time devoted to each area during the learning experience
   (iii) which outcomes have the greater retention and transfer value

E) Constructing Relevant Test Items

The items used could be either ***selection-type*** or ***supply-type*** items. The selection-type items present the students with a set of possible responses from which they are to select the most appropriate answer. The supply-type item requires students to create and supply their own answers.

Selection-type items include:
Multiple Choice, True-False, Matching, Interpretative exercises.

The preparation of good selection-type items is difficult and students can get a proportion of answers correct by guessing.

Supply-type items include:
Short answers, Essays (restricted responses, unrestricted responses)

Supply-type items are easier to construct but more difficult to score.

1. Use the item types that provide the most direct measures of student performance specified by the learning outcome.

2. Avoid verbal associations that give away the answer.

3. Avoid grammatical inconsistencies that eliminate wrong answers.

4. Avoid specific determiners that make certain answers probable (e.g. sometimes) and others impossible (e.g. always).

5. Avoid stereotyped or textbook phrasing of correct answers.

6. Avoid material in an item that aids in answering another item.

7. Avoid trick questions that might cause a knowledgeable student to focus on the wrong aspect of the task.

8. Ensure that the difficulty level matches the intent of the learning outcome and the age group to be tested.

9. Ensure that there is no disagreement concerning the answer. Typically the answer should be one that experts would agree on the correct or best answer.

10. Write the test items far enough in advance that they can be later reviewed and modified as needed.

11. Write more test items than called for by the test plan. This will enable you to discard weak or inappropriate items during the item review and make it easier to match the final set of items to the test specifications.

12. The number of test items depends on the age of the students tested, the time available for testing, type of test items used and on the type of interpretation to be made. Experience in testing is frequently the only dependable guide for determining proper test length.

13. Give due consideration to the best arrangement of the test items. Where possible, all items of the same type should be grouped together. The items should be arranged in terms of increasing difficulty.

14. For **True-False** items make sure that:

- each statement is unequivocally judged true or false
- the statement is brief and stated in simple, clear language
- negative statements are used sparingly and double negatives are avoided
- the statements are free of clues to the answer (e.g. verbal clues, length)
- there is approximately an equal number of true and false statements
- the true and false items are arranged in random order.

15. For **Matching** items ensure that:

- the items are based on homogeneous material
- the instructions clearly state the basis for matching and that each response can be used once, more than once, or not at all
- the items appear on the same page
- an uneven match is provided by making the list of responses longer or shorter than the list of premises.

16. For **Multiple-Choice** items make certain that:

- the stem of the item present a single, clearly formulated problem
- the stem is stated in simple, clear language
- the stem is worded so that there is no repetition of material in the alternatives
- the stem is stated in positive form wherever possible
- if negative wording is used in the stem, it is emphasized in bold or by underlining
- the intended answer is correct or clearly best
- all alternatives are grammatically consistent with the stem and parallel in form
- the alternatives are free from verbal clues to the correct answer
- the distracters are plausible and attractive to the uninformed
- to eliminate length as a clue, the relative length of the correct answer is varied
- the alternative “all of the above” or “none of the above” are used only when appropriate

17. For **Short-Answer** items ensure that:

- the item calls for a single, brief answer
- the item has been written as a direct question or a well-stated incomplete sentence
- the desired response is related to the main point of the item
- clues to the answer have been avoided (e.g. “a” or “an”, length of the blank)
- the units and degree of precision is indicated for numerical answers.

18. For **Essay** questions make sure that:

- questions starting questions with “who”, “what”, “when”, “where”, “name”, “list” are avoided as these terms limit the response
- questions demanding higher order skills, such as those indicated in the following table (Gronlund: 2006, p.120), are used

Outcome	Sample Terms
Comparing	Compare, classify, describe, distinguish between, explain, outline, summarize
Interpreting	Convert, draw, estimate, illustrate, interpret, restate, summarize, translate
Inferring	Derive, draw, estimate, extend, extrapolate, predict, propose, relate
Applying	Arrange, compute, describe, demonstrate, illustrate, rearrange, relate, summarize
Analyzing	Break down, describe, diagram, differentiate, divide, list, outline separate
Creating	Compose, design, devise, draw, formulate, make up, present, propose
Synthesizing	Arrange, combine, construct, design, rearrange, regroup, relate, write
Generalizing	Construct, develop, explain, formulate, generate, make, propose, state
Evaluating	Appraise, criticize, defend, describe, evaluate, explain, judge, write

19. Use this table (Gronlund: 2006, p 63) to help you decide between selection-type and supply-type items:

		Supply-Type Items	
Characteristic	**Selection-Type Items**	*Short Answer*	*Essay*
Measures factual information	Yes	Yes	Yes*
Measures understanding	Yes	No**	Yes
Measures synthesis	No**	No**	Yes
Easy to construct	No	Yes	Yes
Samples broadly	Yes	Yes	No
Eliminates bluffing	Yes	No	No
Eliminates writing skill	Yes	No	No
Eliminates blind guessing	No	Yes	Yes
Easy to score	Yes	No	No
Scoring is objective	Yes	No	No
Pinpoints learning errors	Yes	Yes	No
Encourages originality	No	No	Yes

* The essay test can measure knowledge of facts, but because of scoring and sampling problems it probably should not be used for this purpose.

** These items can be designed to measure limited aspects of these characteristics.

F) Marking Schemes

Marking schemes should:

1. be clear and designed to be easily and consistently applied;

2. allocate marks in proportion with the demands of questions;

3. include the mark allocation for each question and parts of a question, with a more detailed breakdown where necessary;

4. include an indication of the nature and range of responses likely to be worthy of credit and likely responses which would be unacceptable;

5. state the acceptable responses to each question, or parts thereof, in sufficient detail to enable marking to be undertaken in a standardised manner;

6. provide guidance to help markers make judgements on alternative answers;

7. allow credit to be allocated for what candidates know, understand and can do;

8. include marking instructions for assessing quality of written communication, where applicable.

Some important principles with respect to marking schemes:

1. The total number of marks available for each question and each part of a question should be shown in the mark scheme and must tally with the marks shown on the question paper.

2. Each mark should reflect equal demand.

3. All marking should be positive, and as far as possible candidates should gain credit for valid answers and not lose credit for incorrect or irrelevant answers.

Marking schemes must encourage the examiner to use the full range of marks available. Full marks should be available for a level of achievement appropriate to able candidates of the relevant age rather than for a theoretical perfect answer.

G) Checklist (Gronlund: 2006, p 69) for Reviser for Evaluating the Test Items in the Assembled Paper

		YES	NO
Balance	The items measure a representative sample of the learning outcomes The allocation of marks to each item reflects the item difficulty		
Relevance	The items present relevant tasks which reflect the current syllabus		
Conciseness	The items and tasks are stated in simple, clear language		
Soundness	The items are of the proper difficulty, free of defects and have answers that are defensible Questions do not contain gender, cultural or religious bias		
Independence	The items are free from overlapping, so that one item does not aid in answering another		
Arrangement	The items measuring the same outcome are grouped together and are in order of increasing difficulty		
Numbering	The items are numbered in order throughout the test paper		
Instructions	There are clear, concise instructions for each part in the whole test paper There are directions for how to record answers The time limit is specified		
Spacing	The spacing on the page contributes to ease of reading and responding		
Typing	The final copy is free of typographical errors. The marks for the whole paper add up to the total number of marks specified in the syllabus		

References and Additional Reading:

Gronlund, N. E. (2006). *Assessment of Student Achievement*. Pearson Education, Inc.

Linn, R. L., & Miller D. M. (2005). *Measurement and Assessment in Teaching*. Pearson Educational Inc.

Nitko, A. J. (2004). *Educational Assessment of Students*. Pearson Educational Inc.