

CONTENTS

CHAPTER 1: ASSESSMENT IN LEARNING & TRAINING.....	2
CHAPTER 2: METHODS OF MARKING ASSESSMENT.....	4
CHAPTER 3: PRINCIPLES OF ASSESSMENT AND APPLICATION.....	9
CHAPTER 4: FORMULATION OF ASSESSMENT PLANS.....	16
CHAPTER 5: BLOOM'S TAXONOMY OF INSTRUCTIONAL OF OBJECTIVES.....	25
CHAPTER 6: DEVELOPMENT OF THEORY ASSESSMENT SPECIFICATIONS.....	28
CHAPTER 7: TYPES OF TEST ITEMS.....	32
CHAPTER 8: GUIDELINES FOR CONSTRUCTION OF THEORY QUESTIONS..	44
CHAPTER 9: GUIDELINES ON GRADING/MARKING OF TEST ITEMS.....	60
CHAPTER 10: DEVELOP A THEORY TEST PAPER.....	64
CHAPTER 11: CONDUCT OF THEORY ASSESSMENT.....	71
CHAPTER 12: ANALYSIS OF TEST ITEMS.....	79
CHAPTER 13: DESIGN PRACTICAL ASSESSMENT SPECIFICATIONS.....	82
CHAPTER 14: DAVE'S PSYCHOMOTOR CLASSIFICATION.....	89
CHAPTER 15: TYPES OF ASSESSMENT CRITERIA AND TOOLS.....	92
CHAPTER 16: CONSTRUCT A PRACTICAL TEST PAPER.....	96
CHAPTER 17: CONDUCT OF PRACTICAL ASSESSMENT.....	109

CHAPTER 1: ASSESSMENT IN LEARNING & TRAINING

Introduction

- 1 Assessment is not only an integral but also a key part of learning. The role of assessment in learning and training is an important one for without assessment it will not be easy for an instructor to find out whether the learning outcomes have been achieved or for a learner to know about his learning achievement. In a learning environment, the type and number of assessments will depend on a variety of factors.

Purpose of Assessment

- 2 Assessments come in various forms and serve different purposes. Examples of these purposes are:

Award of Qualification: In an institution of learning, it is common for assessments to be carried out by the institution to assess students' abilities as part of the learning process. These assessments can be in the form of tests, assignments, project work and examinations. At the end of the learning period, the assessment results of the students are compiled and those who meet the certification requirements will be awarded the appropriate qualification.

Recognition of Prior Learning: This refers to an assessment carried out to determine whether a learner already possesses the pre-requisite skills and knowledge to follow a particular course of study. The assessment could take the form of a placement test or a quiz. Depending on the result of the test, the learner can be admitted to the course of his choice. Alternatively, he could be assigned an appropriate course or programme level based on the knowledge and skills that he already possesses, as determined following the assessment.

Recruitment: It is not uncommon for an employer to put a job applicant through a job placement test before recruiting him. The test could be a written assessment to measure his job knowledge or a practical assessment to determine his skills competency or both. Based on the results of the placement test as well as job interview performance, the employer would decide if he could be offered employment.

Job Advancement or Deployment: Employees at work do have the opportunity to advance in their careers within the company they work for. Enlightened companies would likely consider their current employees to take on the next-level appointments (for job advancement) or new appointments (to assume new responsibilities in a deployment exercise) when vacancies arise. Assessment of candidates with potentials for job advancement or deployment appointments can take the form of special assignments or project work coupled with work performance evaluation by the candidates' supervisors.

Identification of Competency Gaps: Just like candidates with potentials, who can be assessed to allow their abilities to be gauged to take on higher-level or new positions within their companies, employees who are found wanting could be assessed to determine their need for training and upgrading. Assessment of these under-performing employees could take the form of assignment(s) covering specific duties or tasks, which will allow their supervisors to gauge their deficiencies in terms of their skills, technical knowledge and attitude. Companies do, sometimes, complement their identification of employees' competency gaps with an exercise called *Training Needs Analysis*. Thereafter, training programmes could be mounted to address the competency gaps of the affected employees.

Types of Assessment

- 3 Assessments carried out in a learning environment can be broadly classified into 2 categories, as follows:

Formative Tests

- 4 Formative tests are tests conducted to a group of learners periodically for the purpose of assessing their learning progress. Because the assessment is done periodically, the learners' progress is monitored during the learning process and this will allow for early intervention to address any underlining learning difficulties of learners or deficiencies in instruction. In short, formative tests are used to check the progress of learners. They also allow the instructor to obtain feedback on the effectiveness of his instruction and take appropriate steps, where necessary, to make improvement.

Summative Tests

- 5 Summative tests are those often referred to as end-of-course or end-of-module tests or examinations. Hence, a summative test is one that is conducted after a learner has completed a phase of learning (be it a module, course or programme) and used for the purpose of ascertaining his learning achievement. When a summative test is conducted with no other forms of assessment prescribed, the result of the test will be used to determine the level of performance of his learning achievement.
- 6 Summative testing can yield useful information which can be used to review the instructional process and course contents.

CHAPTER 2: METHODS OF MARKING ASSESSMENT

Introduction

- 1 As in any form of assessment, it is normal for marks to be given to dimensions that are being marked or assessed. There are different methods or approaches which can be used to award marks in assessment. Some of these include:
 - Rating scales
 - Assessment rubrics
 - Marking schemes
- 2 **Rating Scales:** A rating scale is an instrument designed to elicit information about a quantitative or qualitative attribute. To assess the attribute, an assessor or marker will need to select a rating (a number) from a list of numerals assigned to it. The number he picks will be one that he considers to be the likely option which reflects the perceived quality of the attribute. A common example is the **Likert scale** or **Likert item**.
- 3 Rensis Likert is the developer of the Likert scale. Although the Likert scale is the sum of responses on several Likert items, the Likert scale is often regarded to mean the same as the Likert item.
- 4 In any case, the Likert item is simply a statement which the respondent is asked to evaluate according to some subjective or objective criteria. Usually, five (5) ordered-response levels (although 7 is desirable) for the criteria are used in the measurement. The Likert item is appropriate for use where the purpose is for the respondent to select an ordered-response level from a list based on his perceived opinion of the statement. For this reason, the Likert scale is much used for course evaluation purposes where participants are asked for their perceptions on course organisation, lecturer effectiveness and resource provision, etc. An example of a Likert item is given below.

Example of a Likert item

(Tick the appropriate box to record your response)

Attribute Statement	1	2	3	4	5
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
The lecturer speaks with a clear voice and good pronunciation of words.					
The lecture is well presented with good use of power-point slides.					
The training environment is conducive to learning.					

- 5 The **Guttman scale** (named after Louis Guttman who developed it in 1944) is sometimes known as cumulative scaling or scalogram analysis. The purpose of the Guttman scale is to establish a one-dimensional continuum for a concept to be measured.
- 6 Essentially, a Guttman scale consists of a number of items drawn up in the form of statements to which a respondent is requested to consider and indicate whether he agrees or disagrees with each statement. This is normally done in a 'Yes/No' format. Sometimes, it is possible (although not commonly used) for the Guttman scale to be done up using the Likert scale format ie 5-point rating format.
- 7 When using the Guttman scale, the statements (and sometimes questions) used are supposed to gradually increase in specificity ie start from the general to the more specific statements. The intention is for the respondent to agree with the statements on the list up to a point; thereafter, he will stop agreeing as the statements become more specific and difficult to agree.
- 8 The scale is, thus, appropriate for use to determine the extent to which a respondent holds his views on a particular topic or subject (see example of the Guttman scale).

Example of a Guttman scale

Put a 'Yes' or 'No' in the column of the right after reading each statement.	Yes/No
1. I like eating out.	
2. I don't mind eating at hawker centres	
3. I don't mind eating at food-courts.	
4. I like going to smaller restaurants.	
5. I like going to larger restaurants.	
6. I like going to larger restaurants where good service is provided.	
7. I like going to restaurants where service is personalised.	
8. I like going to restaurants that enforce a dress code on their patrons.	
9. I only go to high-class restaurants in 5-star hotels.	
10. I only go to high-class restaurants in 5-star hotels that provide a car valet service.	

8 The steps to develop a Guttman scale are:

- Generate a list of possible statements (*see example above*);
- Get a group of people to score the statements with a 'Yes' or 'No' response, depending on whether they agree or disagree with them;
- Next, draw up a table with the respondents in rows and statements in columns as shown in the table below;
- Sort the columns so the statements with the most number of Yes's are on the left
- Sort the rows so the respondents with the most number of Yes's are at the top
- Select a set of statements that have the least set of No's between all the Yes's (*see table below*)

Table shows respondents and items in a cumulative scale after sorting

Respondent	Item 1	Item 2	Item 3	Item 6	Item 7	Item 5	Item 8	Item 9	Item 10
6	Y	Y	Y	Y	Y	Y	-	-	-
12	Y	Y	Y	Y	Y	Y	-	-	-
4	Y	Y	Y	Y	Y	-	Y	-	-
15	Y	Y	Y	Y	Y	Y	-	-	-
23	Y	Y	Y	Y	Y	-	-	-	-
9	Y	Y	Y	Y	-	Y	-	-	-
6	Y	Y	Y	-	-	-	-	-	-
13	Y	Y	Y	-	-	-	-	-	-
17	Y	Y	-	-	-	Y	-	-	-
25	Y	-	-	-	-	-	-	-	-
30	-	-	-	-	-	-	-	-	-

Responses that could be treated as exceptions

[Notice that the scale is very nearly cumulative when you read from left to right across the columns (items). The respondent who agrees with Item 2 will find others agreeing with item 1. And, if someone agrees with Item 3 will find others agreeing with Items 2 and 1. Similarly, if someone agrees with item 5 will find others agreeing with items 7, 6, 3, 2 and 1. However, the matrix also shows that the cumulateness of the scale is not perfect. While a person agreeing with Item 5 tends to agree with 7, 6, 3, 2 and 1 as well, there are some exceptions to the rule as can be seen in the Table.]

- 9 As the Guttman scale is intended for use to establish a one-dimensional continuum for a concept to be measured, it may be appropriate for use to determine the extent or degree respondents agree with a particular concept or principle as in an opinion survey, for example. However, its use in the assessment of learning and teaching may be limited in view of the cumbersome nature of its application and difficulty in interpreting the cumulateness of the scale.

Assessment Rubrics

- 10 A *rubric* usually contains a set of criteria specially drawn up for use to assess a particular type of work or performance. Hence, it is appropriate for use in the assessment of assignments and project work undertaken by learners.

- 11 In drawing up a rubric, the first step is to determine the criteria that will need to be assessed. The levels of desired achievement are then assigned for each criterion. For example, If there are 8 criteria and each criterion is assigned 4 levels of desired achievement, the rubric will have a total of 32 (8 x 4) 'boxes' with each box having a description of the desired outcome. Next, a maximum possible score is assigned to each criterion based on the highest level of achievement possible.
- 12 Where there are criteria identified as being critical and difficult to achieve, a weighting factor could be assigned to these criteria. A criticality factor usually comprises 3 levels, namely Level 1 – least critical; Level 2 – critical; and Level 3 – most critical. The difficult factor, similarly, has 3 levels as well.
- 13 After the use of the rubric for assessment, the actual score for the work or performance is produced by adding the marks of all the scored criteria. An example of a rubric is given below.

Rubric for assessment of machined parts

Criteria	Level 1	Level 2	Level 3	Level 4	Score	
	Completely Unacceptable	Need improvement	Acceptable	Excellent	Actual (R x C x D)	Max Marks Possible
Part 1 Dimension Accuracy	More than 30% of dimensions are out of tolerance	More than 20% but less than 30% of dimensions are out of tolerance	Less than 20% of dimensions are out of tolerance	All dimensions machined to within tolerance	_ x 3 x 2 = _	4 x 3 x 2 = 24
Part 2 Dimension Accuracy	More than 30% of dimensions are out of tolerance	More than 20% but less than 30% of dimensions are out of tolerance	Less than 20% of dimensions are out of tolerance	All dimensions machined to within tolerance	_ x 3 x 2 = _	4 x 3 x 2 = 24
Part 3 Dimension Accuracy	More than 30% of dimensions are out of tolerance	More than 20% but less than 30% of dimensions are out of tolerance	Less than 20% of dimensions are out of tolerance	All dimensions machined to within tolerance	_ x 3 x 2 = _	4 x 3 x 2 = 24
Assembly Function of the 3 machined parts	All the 3 parts cannot be assembled.	Only 2 out of the 3 parts can be assembled	All 3 parts can be assembled but cannot function well in the assembled form	All 3 parts can be assembled and functional in the assembled form	_ x 3 x 3 = _	4 x 3 x 3 = 36
Quality of Finish	None of the parts has a well polished surface finish	Only 1 or 3 parts have a well polished surface finish	Only 2 of 3 parts have a well polished surface finish	All 3 parts have a well polished surface finish	_ x 2 x 2 = _	4 x 2 x 2 = 16
Completion Time	Candidate completed machining task on time	Candidate completed machining task within 90% of allocated time	Candidate completed machining task within 80% of allocated time	Candidate completed machining task within 70% of allocated time	_ x 1 x 2 = _	4 x 1 x 2 = 8
Score					____(____%)	132 (100%)

R = Rating according to level achieved

C = Criticality comprising 1 (less critical), 2 (critical), 3 (very critical)

D = Difficulty comprising 1 (less difficult), 2 (difficult), 3 (very difficult)

Marking Scheme

- 14 A marking scheme is generally referred to as a documented plan devised to guide an assessor or marker to allocate marks for an assessment. It is widely used in the assessment of learning such as for the marking of answer scripts of assessment or examination papers, and course assignments and projects.
- 15 As assessments vary in design and contents, different marking schemes are needed for different types of assessment. For example, a marking scheme for a theory assessment will be distinct from that of a performance-based or practical work assessment. A theory test may require a learner to provide answers to questions (such as multiple-choice, true-false and short-answer questions) in a written form. For such a test, the marking scheme would likely comprise a list of correct or suggested answers (sometimes called model answers) with suggested allocation of marks.
- 16 A performance-based marking scheme, on the other hand, would tend to focus on assessing a learner on his work process (ie the way he performs his task operation and/or work outcome). Such a scheme used to assess work operation and/or end product would generally specify the performance criteria requirements and include an allocation of marks for each criterion. Where necessary, weighting are added to accord more marks for criteria that are critical and difficult to achieve. This will help to discriminate between the high-performers and the low-performers
- 17 The application of marking schemes for the theory assessment and the performance-based assessment will be discussed in more depth in the later chapters.

CHAPTER 3: PRINCIPLES OF ASSESSMENT AND APPLICATION

Introduction

- 1 In almost all cases, the outcome of learning is determined through assessment. This is so as the assessment result forms the basis that determines whether a learner has attained the achievement following a period of learning. Where a test is set or developed for the assessment, the relevance and quality of the test is governed by a set of assessment principles referred to as:
 - a) Validity;
 - b) Reliability;
 - c) Sufficiency;
 - d) Authenticity; and
 - e) Practicality.

Validity

- 2 The validity of a test refers to the accuracy of an assessment ie the ability of the test to measure what it is supposed to measure. Validity is derived from the results of a test instrument for a given group of learners involved in the test. And the validity being established is based on the interpretation of the results (not the validity of the test itself).
- 3 The degree of validity is the single most important aspect that is derived from the test results. And this degree can be termed as 'high validity', 'moderate validity' and low validity'.
- 4 The validity derived from a test instrument should not be regarded as a general quality (for broad-based use) but specific to a particular use. For example, the results of an arithmetic test may have a high degree of validity for indicating computational skill, a moderate degree of validity for predicting success in future maths courses, and low validity for predicting success in art.

Types of Validity

- 5 Three types of validity are commonly used in educational and psychological measurement. These are:
 - Content-related validity;
 - Criterion related validity; and
 - Construct-related validity

- 6 The descriptions of the 3 types of validity are shown in the table below.

Type of Validity	Descriptions
Content-related validity	It describes the extent the sample of the tasks used in the test represents the domain of tasks to be measured.
Criterion-related validity	It describes how well the test performance is able to predict future performance.
Construct-related validity	It describes how well the test performance is able to determine specific individual characteristics* such as reading comprehension, communication ability, and understanding of scientific principles. [*refers to abilities (or skills), psychological traits, and personal qualities.]

Validity of Assessment

- 7 A valid assessment should meet learning objectives. For a test to be valid, it needs to be set in tandem with the learning outcomes ie the questions asked are in line with and should test the stated learning objectives. For example, a question should not be designed to test a learner's ability to recall knowledge when the intention, in accordance with the learning objective, is to test the learner's ability to apply knowledge.

Factors affecting Validity

- 8 Numerous factors tend to make test results invalid for their intended use. The factors below can prevent the test items from functioning as intended and thereby lower the validity of the interpretations from the test results.

Unclear instructions to candidate:

- 9 Validity can be affected or lowered if instructions on test papers given to candidates are not clear eg candidate are in doubt on how to respond to the items; whether candidates are permitted to guess their responses; and what to do to record their responses.

Reading vocabulary and sentence structure too difficult:

- 10 If candidates are not able to understand what they read in the test items, the test may end up measuring reading comprehension rather than the aspects of performance required.

Inappropriate level of difficulty of the test items:

- 11 Items that are too easy or too difficult do not provide reliable discriminations among candidates. Validity may be lowered as a result.

Poorly constructed test items:

- 12 Test items which unintentionally provide clues to the answer tend to measure the candidate's alertness in detecting clues, in addition to the aspects of performance that the test is intended to measure.

Ambiguous test items:

- 13 Ambiguous statements in the test can cause misinterpretation and confusion. Ambiguity can, sometimes, confuse the better candidates more than the weaker candidates, thus causing the items to discriminate in a negative direction.

Test items inappropriate for the outcomes being measured:

- 14 Having wrong items in the test (eg items to measure understanding, thinking skills and other complex types of achievement) which is intended for measuring factual knowledge, the test results can become invalid.

Test too short:

- 15 If a test is supposed to be of a long duration and requires many items to be tested, a shortened test may not provide a representative sample of the performance of the candidates. In this case, the validity may be lowered.

Improper arrangement of items:

- 16 Test items are usually arranged in order of difficulty, starting with the easiest items. If difficult items are placed early in the test, candidates may spend too much time on these, and fail to reach the items they could easily answer. Improper arrangement may also influence validity by having an adverse effect on candidate motivation.

Identifiable pattern of answers:

- 17 A systematic pattern of correct answers (e.g. T, T, F, F for True/False questions or D, C, B, A, then D, C, B, A for multiple-choice questions) may enable candidates to guess their answers. When this happens, validity may be affected.

Reliability

- 18 Reliability refers to the results obtained with an evaluation instrument, and not to the instrument itself. It is more appropriate, thus, to speak of the reliability of 'the test scores' or of 'the measurement', rather than of "the test" or of "the instrument".
- 19 In assessment, reliability refers to the consistency of test measurement i.e. the ability of a test to produce the same results if administered to the same candidate later. For example, if a learner has sat for a theory test a few months' ago and sit the same test today, the scores of the 2 tests should be similar if the results are to be treated as reliable scores.

- 20 Reliability is correlated to the confidence level one has on the test scores. The higher the confidence level, the higher is the reliability. Hence, fewer and smaller the errors of measurement will increase the reliability of the test scores.

Improving Reliability of Test Scores

- 21 Reliability increases with more test items – If test scores are based on a larger number of test items, the scores will be more reliable. More items provide for more adequate item sampling. There will be greater stability less affected by chance factors such as familiarity with some items and ambiguity in instructions. Reliability tends to decrease when test scores are based on few test items.
- 22 Reliability increases with objective marking – If test scores are based on an objective marking system (such as use of a rubric or marking scheme), the scores will be more reliable. Consistency in the allocation of scores for test items will further increase the reliability of the test scores. Likewise, subjective marking of test items tends to lower the reliability of test scores.
- 23 Reliability increases with conducive test environment – If test scores are based on a test conducted in a non-conducive physical environment (such as a location that causes distractions and discomfort), the scores will tend to be less reliable. Candidates could be affected by the low comfort level and, hence, perform lower than expected. An uncomfortable environment might do more harm than good to candidates who are not in the best of physical condition. On the contrary, a favourable environment will tend to increase the reliability of test scores since candidates will not be affected by any environmental issues.

Relationship between Reliability and Validity

- 24 Reliability is a necessary condition to support the validity of test scores, but not sufficient enough for use to confirm test score validity. A test that provides totally inconsistent (unreliable) results will not be able to provide valid information about the performance being measured. On the other hand, highly consistent (or reliable) test results may be measuring a test that is invalid or used in ways that are inappropriate.
- 25 Thus, low reliability can be expected to restrict the degree of validity that is obtained, but high reliability does not provide any assurance that a satisfactory degree of validity is present. In short, reliability merely provides the consistency that makes validity possible.
- 26 The relationship is illustrated in Annex A.

Sufficiency

- 27 Sufficiency in the amount of evidence collected in an assessment is key to ensuring that all aspects of the competency standard have been captured to satisfy the need for repeated performance. For an assessment to be regarded as meeting the 'sufficiency requirement', the test items should have adequate coverage to ascertain a candidate's competency or sample a sufficient range of the performance of the candidate.
- 28 In short, the test items in the assessment should comprise items that adequately cover the contents or topics taught. This means that all learning outcomes should be appropriately assessed (ie, if 10 topics are taught, there should be questions asked on all the 10 topics). In addition, more questions should be asked on topics that are accorded higher relative importance (ie more items on topics allocated more training time). If need be, topics that are inter-related could be tested as integrated items.

Authenticity

- 29 For an assessment to be regarded as 'authentic', the test items must be set within the context of the real world. In an assessment, a learner must be able to see the consistency between the assessment and the real-world application for which the learner is being prepared for. In other words, the contents of the questions such as examples and problems posed are directly related to the course and not outside it.
- 30 Authenticity of assessment is particularly important when dealing with the assessment of competence as it involves determining whether the candidate is sufficiently competent, specific to the content of work or domain of practice. For example, it may not be appropriate to use a multiple-choice question examination to assess a candidate's communication skills.

Practicality

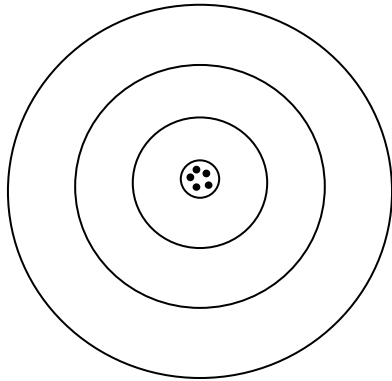
- 31 Practical considerations will need to be taken into account in the selection of a test instrument. These considerations include:
- Ease of administration – tests should be designed and set such that they are easy to administer. If sub-tests are required, they should be kept to a manageable number for ease as well as efficient administration. When tests become cumbersome to administer, thus, resulting in the occurrence of error and problem, the validity and reliability of the test scores could be affected
 - Time duration – reliability is directly related to the time duration (or length) of a test. If a test is too short, it could reduce the reliability of the test scores. Similarly, a lengthy test may not necessarily increase the reliability. A test should be allotted with just the right amount of time for the sole purpose of obtaining valid and reliable results, and no more.

- Ease of marking of test – the way test responses are marked has a direct effect on reliability. Since subjective marking tends to make test scores less reliable, it is desirable for marking to be carried in an objective way and guided by a marking scheme. The use objective-type questions such as true/false and multiple choice questions allow for objective marking to be carried out. And objective-type questions can be marked with relative ease. The use of a separate answer sheet for multiple choice questions and answer booklet for short-answer questions could further facilitate the marking process.
- Ease of interpretation and evaluation – Analysis done to test scores could impact on decisions taken based on the outcomes of the analysis. Hence the quality of analysis depends to a great extent to how well the test scores are interpreted and evaluated. Misinterpretation could be detrimental not only to the organization but also to the candidates. It is, therefore, essential that the test scores are compiled in such a way that will allow for ease of interpretation and evaluation.
- Cost to administer test – Although cost is a consideration item, it is not considered a major item per se. Nonetheless, a test should not incur high costs to administer. Neither should it be administered at a very low cost which could contribute to lowering the reliability of the test scores.

Relationship between Validity and Reliability

The concept is to test the centre of target

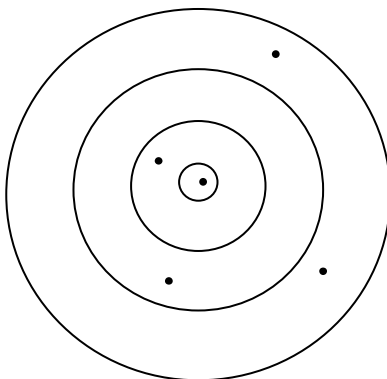
Scores are Valid and Reliable



Scores are valid as all the shots hit the centre of target.

Scores are reliable as every shot hits the centre, thus, obtaining maximum points.

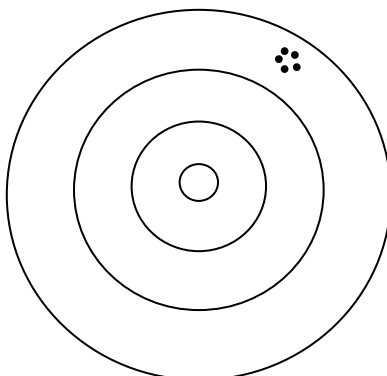
Scores are Not Valid and Not Reliable



Scores are not valid as only one shot hits the centre and the rest are outside the centre of target.

Scores are not reliable as every shot is not able to secure the same score or same number of points.

Scores are not Valid but Reliable



Scores are not valid as all the shots have landed outside the centre of target.

Scores can be said to be reliable as every shot is able to secure the same number of points although each point is low (ie reliable in getting low scores).

CHAPTER 4: FORMULATION OF ASSESSMENT PLANS

Introduction

- 1 An assessment plan (or test plan) is a document which specifies the scope of and requirements for a formal assessment or test of a module or course. The plan is meant for use by the trainer who is assigned to set the test paper. He is commonly referred to as the setter.
- 2 There are generally 2 types of test plan ie theory test plan and practical test plan. Some of the modes for testing candidates under the 2 plans are:
 - a) Theory test mode (to test knowledge acquisition)
 - Pen and paper test – candidate to provide written answers on test paper
 - Oral questioning – candidate to provide answers orally from questions asked. This mode is used only when it is not possible for evidence to be obtained of a performance criterion. If used, the allocation of marks for oral questions should not exceed 10% of the total marks.
 - On-line assessment – candidate attempt test via a intranet system
 - b) Practical test mode (to test skills abilities or competence)
 - Practical performance test – candidate to perform an activity or task
 - Role play – candidate to act out a role as activity for assessment
- 3 In setting the test, the setter will need to comply with all the instructions and specifications as prescribed in the test plan.
- 4 Various personnel linked to the training process could be involved in formulating and developing the assessment plan. These personnel, sometimes known as stakeholders, could include the Training Manager, Supervisors, Curriculum developer, Subject Matter Expert (Specialist) and Assessor.

Role of Stakeholders in the development of test plan

- 5 The development of a test plan may involve some or all of the stakeholders. Their involvement may range from providing inputs to carrying the actual developmental work to giving approval to the developed test plan.

6 The stakeholders' direct and indirect roles are tabulated in the table below.

Stakeholder	Key Job Function	Direct Role in Development of Test Plan	Indirect Role
Training Manager	Manage the training function in the training department or organisation.	- Oversee and maintain standard and framework established for test. - Approve test plan.	- Appoint or assign setter and marker/assessor <u>After completion of test and marking</u> - Consider and approve test results. - Interpret and evaluate test score analysis.
Curriculum Developer	Design and develop curriculum within the training department or organisation. He is could be a content expert as well.	Develop test plan (if he is a content or subject matter expert as well) for module or course to align with instructional objectives as prescribed in the skills standard of the course.	
Subject Matter Experts	Provide instructions on a particular subject or skills area, and is well-versed with the subject's contents. He is a content specialist.	- Provide guidelines for proper construction of test items eg true/false, multiple choice, short-answer and extended response questions. The guidelines are to serve as reference materials to the setter of test items. - Provide inputs to the Curriculum Developer (CD) on topics covered for module or course. - Develop test plan, if required	<u>After completion of test plan development</u> - Set theory and/or practical papers if assigned. - Mark theory test scripts for theory test. - Mark practical projects for practical test. <u>After completion of test</u> - Mark theory test scripts for theory test. - Mark practical projects for practical test.
Assessor	Carry out assessment of theory or practical (plus on-site) test. The assessment work will include marking of scripts or practical work performance.	(May not have direct role in test plan development. However, may provide Curriculum Developer with feedback on candidates' performance from previous tests.)	<u>After completion of test plan development</u> - Set theory and/or practical papers if assigned. - Develop assessment checklist or marking scheme for process assessment on-site. - Carry out process assessment on-site and allocate marks based on checklist or marking scheme. - Mark theory test scripts for theory test. - Mark practical projects for practical test.

Supervisor	Oversee and guide workers or trainees in performance of work or training	(May not have direct role in test plan development. However, may provide information on availability of facilities and equipment, including tools and materials.)	<u>After completion of test plan development</u> • Prepare facilities for the conduct of both theory and practical tests • Prepare tools, equipment and materials for use by trainees for the practical test • Facilitate the smooth conduct of test during test period.
------------	--	---	--

Scope and Design of a Theory Assessment Plan

7 A theory test plan developed for formal assessment is a confidential document. After development, it should be handled and kept by an authorised person who, in turn, is to pass it to the appointed setter of the test paper. For the purpose of accountability, it is normal for the setter to acknowledge receipt of the test plan. A setter is usually issued with a copy on 'Instructions for Setter' to guide him in setting the test paper.

8 In developing the test plan, the developer should make reference to the skills standard of the course. This is because the number and type of questions to be prescribed in the Table of Specifications will need to be aligned to the instructional objectives as specified in the Skills Standard.

9 An appropriate theory test plan should contain the following information:

- Course Title

It is important to specify the course or module title in the test plan so that the setter will not set anything outside the contents of the course or module. This will allow him to make reference to the appropriate skills standard. In addition, the number of pages of the test plan should be indicated. This is to inform the setter that he has the right number of pages of information.

- Scope of Test

Under the scope of test, the following information must be provided to the setter:

- a) The number of paper to be set and the title of each paper;
- b) The number of sections for each paper (eg Section A and Section B);

- c) The type and number of test items (questions) to be set for each section (eg multiple choice questions may be suitable for a pen and paper test but not for test involving oral questioning);
- d) The marks to be allocated to each section and the total marks for each paper;
- e) Whether candidates are required to attempt all questions or they are given a choice of questions to answer;
- f) Guidelines, if any, to be used by the setter to set his questions to the expected quality (eg reference should be made to the 'Guidelines for the construction of test items' issued by the Training Centre); and
- g) A **Table of Specifications**, which provides information on the distribution of test items among the test syllabus topics. (This will be covered in detail in Chapter 6.)

- Duration of Test

The duration of a theory test is normally stated in terms of number of hours (eg 2 hours) or hours and minutes (eg 1 hour 30 minutes). The setter is to take cognation of the test duration to set questions up to a difficulty level such that candidates will be able to complete the test within the given time.

- Venue for test

This refers to the location where the theory test is to be conducted. Usually, a room with desks and chairs are provided is used to the test. Whoever is in-charge of the candidates (eg trainees' supervisor) or a person assigned to do make the preparation arrangement is to make sure that adequate facilities are provided before the test.

- Marking scheme

The test plan should specify to the setter to provide a marking scheme as part of the test setting process. The marking scheme is to provide the correct answers of the questions (such true/false and multiple choice questions) or model answers to the short answer or extended answer questions). In addition, the marking scheme should provide instructions to the marker on the allocation of marks especially to incomplete answers to short or extended questions provided by the candidate. If there is no penalty for any wrong answer given (eg in the case of true/false or multiple choice questions), it should be clearly stated in the test plan as well so that the setter will not make provision for any deduction of marks in the marking scheme.

- Pass mark and grading system

This should be made known in the test plan so that the setter can incorporate it into the marking scheme. The maximum possible percentage of marks attainable is normally 100%. Information that could be included would be:

- Pass mark percentage (eg 50%); and
- Grading system for marks (percentage) scored (see example below)

Grade	Score	Grade Interpretation
A - Excellent	80% and above	Attained highest level objectives
B - Good	70% to 79%	Attained high level objectives
C - Credit	60% to 69%	Attained mid-level objectives only
D - Pass	50% to 59%	Attained adequate level of competence
F - Fail	Below 50%	Unable to attain minimum level of competence

Scope and Design of a Practical Assessment Plan

- 10 Unlike a theory test, a practical test assesses the practical skills or skills competence of a trainee or candidate. The test measures the candidate's skills performance against a set of performance criteria as specified in the skills standard.
- 11 The test could take the form of a single project or task. It could also comprise a set of sub-tasks or one or more assignments which, together, form a complete test.
- 12 As with a theory test plan, a practical test plan formulated for formal assessment is a confidential document. After development, it should be handled and kept by an authorised person who, in turn, is to pass it to the appointed setter of the test paper. For the purpose of accountability, it is normal for the setter to acknowledge receipt of the test plan. A copy on 'Instructions for Setter' is usually given to the setter to guide him in setting the test paper.
- 13 In developing the practical test plan, the developer should make reference to the skills standard of the course, in particular the performance criteria.
- 14 A practical test plan should contain the following information:
 - Course Title
It is important to specify the course or module title in the test plan so that the setter will not set anything outside the contents of the course or module. This will allow him to make reference to the appropriate skills standard. In addition, the number of pages of the test plan should be

indicated. This is to inform the setter that he has the right number of pages of information.

- Design Specifications

Under design specifications, the following information must be provided to the setter:

- a) Description of test task (or project) - The test plan should provide a simple and clear description of the task which is required of the candidate to perform.
- b) Scope of Test – The scope of test refers to the range of skills or competencies which the candidate will be tested on. The skills areas must be those found in the skills standards. Where appropriate, the competencies should include abilities pertaining to work safety ie test candidate's ability to carry out task activities in a safe manner. [For the test to be valid, it should cover at least 70% of the competencies stipulated in the skills standard, and all critical competencies should be included in the scope of test.]
- c) Design features – Under this section, the test plan developer is required to specify how the test is to be conducted.
- d) Duration of test – The test duration will need to be specified by the test plan develop. The setter is expected to set the practical test paper based on the duration stipulated and in a way that the candidate will be able to complete the task within the time given. If necessary, the setter could carry out a dry run of the test paper before finalising it. But, he is not allowed to change the duration. [As a general rule, setter should make sure that 90% of the candidates are able to complete the test task within the given time set.]

Design specifications will be covered in greater detail in Chapter 13.
--

- Tools, equipment and Materials

The test plan should include a list of tools, equipment and materials that should be used in the test. Safety equipment and gears should also be included in the list.

- Test venue

This refers to the location where the practical test is to be conducted. This information is needed by the person in-charge the workshop or work place to make the necessary physical preparations for the test, including making provision for the tools, equipment and materials to be ready for the test.

- Marking scheme

It is necessary to require the setter to provide a marking scheme for the practical test. The items to be measured are governed by the performance criteria stated in the skills standard. The marking scheme should set out how the performance is to be scored ie how marks are to be awarded to the candidate's performance in an objective way. Generally, the marking scheme should address the following aspects:

- a) Accuracy, quality and finish of work (product marking)
- b) Process and method of working observed (process marking)
- c) Correct use of tools and equipment observed (process marking)
- d) Observance of safety precautions by candidate (process marking)
- e) Economic use of materials by candidate (process marking)

A well-structured and objective marking scheme should take into consideration the level of criticality and difficulty of each item to be measured. Such a marking scheme is needed for performance-based assessments (refer to Chapter 13 for details).

Sometimes, a marking scheme can be as simple as a checklist for tasks or sub-tasks that are simple and require a very short time to perform. The checklist comprises a list of items to be measured and the assessor ticks the appropriate boxes based on his assessment. The drawback in using a checklist for assessment is that scores are not assigned and the criticality and difficulty levels are taken into consideration.

An example of a simple checklist is shown below.

No	Item to be measured	Yes	No
1	Did the trainee wear safety shoes?		
2	The length of the machined shaft is within 20 ± 0.05 cm?		

- Pass mark and grading system

This should be made known in the test plan so that the setter can incorporate it into the marking scheme. The maximum possible percentage of marks attainable is normally 100%. Information that could be included would be:

- Pass mark percentage for practical work (eg 60%)
- Grading system for marks (percentage) scored
(see example).

Grade	Score	Grade Interpretation
Distinction	80% and above	Attained high level of competence.
Pass with credit	70% to 79%	Attain mid to high level of competence
Pass	60% to 69%	Meet minimum level of competence
Fail	Below 60%	Unable to meet minimum level of competence

Duties of a Setter of Test Paper

- 15 The role of a setter of test papers is of critical importance. This is because the test scores could ultimately determine whether or not candidates will qualify for the award of certificates.
- 16 Since the quality (both validity and reliability) of the each test or assessment depends significantly on the professional competence of the setter, it is therefore Important that the appointment of a setter must be carried out through a proper selection exercise to ensure only those who are qualified and with experience in conducting training in the relevant field are considered (such as trainers, subject matter experts and assessors).
- 17 A setter, when appointed, is normally required to sign a 'confidentiality declaration' to require him to ensure the confidentiality of the test that he sets. Such requirement is to safeguard the integrity of the test administration system as well to require the setter to take steps to prevent leakage of the contents of the test paper. The requirement is necessary as pilot testing may be carried out for some tests.
- 18 The main duties of a setter are to:
 - Set questions according to the given skills standard and test plan;
 - Provide model answers and marking scheme;
 - Indicate tools, equipment and materials required for the practical test;
 - State instructions for candidates to ensure a smooth conduct of the test; and
 - Maintain confidentiality of test paper set.
- 19 A setter should ensure that the questions are set to give the right emphasis to the topics to be tested and that they meet the requirements of the table of specifications as stipulated in the test plan. It is expected that each test should be set such that:
 - The questions asked are appropriate to the standard of the course level;
 - The general and specific instructions contained in the test paper are clear and easy-to- understand without any typographical error;

- (For a theory test) The questions are well-phrased, grammatically correct and to elicit definitive responses for a theory test;
- (For a practical test) The assignments to be undertaken by candidates and their assessment criteria are specific and clearly stated; and
- The marking scheme contains specific instructions and information to guide the assessor to perform his assessment objectively.

CHAPTER 5: BLOOM'S TAXONOMY OF INSTRUCTIONAL OBJECTIVES

Introduction

- 1 Bloom's Taxonomy (named after educational psychologist Benjamin Bloom) is a classification of education objectives. The taxonomy provides a useful structure in which to categorize learning objectives and test questions.
- 2 Various terms used such as Instructional objectives, learning objectives, behavioural objectives and learning outcomes are words commonly used to refer to a description of observable trainee behaviour on learning performance in education and training. Instructional objectives or IOs specify what a trainee should be able to acquire following each learning session or programme. These IOs can be classified into different levels requiring different abilities or skills to achieve.
- 3 Bloom has classified learning into 3 domains (or classifications) as follows:
 - Cognitive domain – development of mental abilities
 - Affective domain – development of behavioural –related abilities
 - Psychomotor domain – development of practical skills
- 4 As achievement testing is to assess mental abilities (thinking and intellectual skills), the focus of this chapter will, therefore, discuss the different levels of intellectual abilities within the Cognitive domain.

Cognitive Domain

- 5 There are 6 levels of abilities in the Cognitive domain. These levels are Knowledge (level 1), Comprehension (level 2), Application (level 3), Analysis (level 4), Synthesis (level 5), and Evaluation (level 6). The table below shows the different abilities that can be demonstrated or displayed by a learner at different stages learning achievement.

(Cognitive domain of Bloom's Taxonomy)

Domain Level (level of competence)	Intellectual Ability (ability displayed in test)	Words for Use to Test Ability
L1 - Knowledge	• observation and recall of information • knowledge of dates, events, places • knowledge of major ideas • mastery of subject matter	list, define, tell, identify, show, label, collect, examine, tabulate, quote, name, who, when, where.

L2 - Comprehension	• understanding information • grasp meaning • translate knowledge into new context • interpret facts, compare, contrast • order, group, infer causes • predict consequences	summarize, describe, explain, interpret, contrast, predict, associate, distinguish, estimate, differentiate, discuss.
L3 - Application	• use information • use methods, concepts, theories in new situations • solve problems using required skills or knowledge	apply, demonstrate, calculate, complete, illustrate, show, solve, examine, modify, relate, change, classify, experiment, discover.
L4 - Analysis	• seeing patterns • organisation of parts • recognition of hidden meanings • identification of components	analyse, separate, order, connect, classify, arrange, divide, compare, select, infer.
L5 – Synthesis	• use old ideas to create new ones • generalise from given facts • relate knowledge from several areas • predict, draw conclusions	combine, integrate, modify, rearrange, substitute, plan, create, design, invent, what if, compose, formulate, prepare, generalize, rewrite.
L6 - Evaluation	• compare and discriminate between ideas • assess value of theories, presentations • make choices based on reasoned argument • verify value of evidence • recognise subjectivity	assess, decide, rank, grade, test, measure, recommend, convince, select, judge, discriminate, support, conclude, compare.

6 Examples of test items for the 6 taxonomic levels are given at Annex A.

7 With good understanding of Bloom's cognitive domain's taxonomic levels, test items can be assigned different levels for the purpose of assessing cognitive abilities. The levels should be indicated in the test plan to direct the setter to the desired questions.

Cognitive Taxonomy Examples

A Knowledge

- 1 Name a personal protection equipment (PPE) used at a construction site.
- 2 State the formula for calculating the area of a circle.

B Comprehension

- 3 Give two examples of an electrical conductor.
- 4 Interpret the data from a graph.
- 5 Distinguish between outbound and inbound travel.

C Application

- 6 Given the set of data, calculate the mean and standard deviation.
- 7 Convert local time to Greenwich Mean Time (GMT).

D Analysis

- 8 Outline the costs and benefits of the business proposal.
- 9 Given a case study, point out the causes of the problem.

E Synthesis

- 10 Create an experiment to show air expands when its temperature increases.
- 11 Formulate a procedure to improve safety awareness in the plant.

F Evaluation

- 12 Evaluate the quality of the new food product.
- 13 Assess the course project in accordance with the prescribed criteria.

CHAPTER 6: DEVELOPMENT OF THEORY ASSESSMENT SPECIFICATIONS

Introduction

- 1 One of the key aspects of the theory test plan (as mentioned in Chapter 4) is the development of assessment specifications, normally referred to as a **Table of Specifications (TOS)**.
- 2 In drawing up the **TOS**, two (2) types of specifications will need to be considered, as follows:
 - Test Domain Specification – This specifies to the setter the theory or technical knowledge items in the skills standard to be tested.
 - Item Specification – Since a test tends to have items asked on a representative sample of test topics (syllabus topics), the type of test items selected becomes significant. Questions set should be those that will allow for more test items to be used (higher sampling) and those that can be used best to assess candidates' performance. Hence, objective-type questions (such as true/false, multiple choice and short-answer questions) tend to meet the requirements. Thus, in the **TOS**, the test items should be distributed among the questions selected for the test.

In addition, the cognitive taxonomy level* for each test item should also be specified. (*Refer to Bloom's Taxonomy of Instructional Objectives in Chapter 5)

- 3 An illustration of **TOS** is given below.

Table of Specification

Item Specification + Taxonomic Specification Test Item/ Skills Standard Code	Section A Multiple choice Questions			Section B Short Answer Questions		
	I	II	III	I	II	III
Competence 1 (TK item 1.1 to 1.6)	1	1	1	-	1	
Competence 2 (TK item 2.1 to 2.4)	-	1	1	-		1
Competence 3 (TK item 3.1 to 3.10)	1	2		-	1	1
Competence 4 (TK item 4.1 to 4.5)	1	1	-	-	1	
Sub-total	3	5	2		3	2
Total	10			5		
	15					

Taxonomic Level

- I – Level 1: Knowledge
- II – Level 2: Comprehension
- III – Level 3: Application

Distribution of test items by question type and taxonomic level.

Topics to be tested (based on the Skills Standard) – For example, Competence 4 (TK item 4.1 to 4.5) refers to Competence No 4 in the Skills Standard and the Technical Knowledge items (No 1 to 5) which the candidate is expected to have learned. The 3 test items to be set for Competence 4 should be based on the 5 Technical Knowledge items as reflected in the Skills Standard.

4 The steps in drawing up the **TOS** are listed in the table below.

Step 1	Create a blank table as given in the illustration.
Step 2	From the skills standard of the module or course, identify the competencies and technical knowledge topics that need to be tested. Thereafter, list the topics under the 'Test Item/Skills Standard Code' column.
Step 3	Look for the question types and the item distribution as indicated in the test plan eg 10 multiple choice questions and 5 short-answer questions. Thereafter, decide on the allocation of test items for each competence based on 2 criteria, as follows: a) More knowledge topics taught, more items b) More time allocated to topics, more items.
Step 4	Within each question type (eg Multiple choice questions in Section A), distribute the items under taxonomic level I (knowledge), level II (comprehension) and level III (application) based on the distribution proportion pre-determined by the training centre or organisation. One example of a distribution proportion is as follows: • Level I – maximum 30% of total questions • Level II – minimum 35% of total questions • Level III – minimum 30% of total questions.
Step 5	Enter the total number of questions in the sub-total and total boxes.

5 An example of a test plan complete with TOS is given at [Annex A](#)

THEORY TEST PLAN

Certification Title	Electronics Competency Certificate
Module	Fundamental Electronics
Date of Approval	XXXXXX
No. of Pages	3

THIS TEST PLAN MUST BE USED IN CONJUNCTION WITH THE SKILLS STANDARD DATED XXXXX.

Scope of Tests

1. The test comprises one paper - Fundamental Electronics.
2. The paper consists of two (2) sections of 23 questions in total in the following format:

 Section A – 15 Multiple choice questions.
 Section B – 8 Short-answer questions.
3. The marks for each section are as follow

 Section A – 60 marks (4 marks for each question)
 Section B – 40 marks (5 marks for each question)
4. Candidates must attempt all questions.
5. The table of specification for the test is given at **Appendix 1**.
6. Reference must be made to the guidelines for constructing test items issued by the Training Centre.

Duration of Test

7. The duration for the paper is 2 hours.

Marking Scheme

8. Answer for each question must be provided. For short-answer questions, Instructions must be provided to guide the marker to allocate marks for incomplete answers. There should be no penalty for wrong answers.
9. The score of the test accounts for 70% of the final score for the theory component, while coursework accounts for the remaining 30%.

Pass Mark and Grading System

10. The maximum possible percentage of marks attainable is 100%. The pass mark for this theory paper is 50%. The grading system is as follows:

<u>Grade</u>	<u>Score</u>
A – Excellent	80% and above
B – Good	70% to 79%
C – Credit	60% to 69%
D – Pass	50% to 59%
F – Fail	Below 50%

**TABLE OF SPECIFICATIONS
(Fundamental Electronics)**

Test Item/Skills Standard code	Section A Multiple choice Questions			Section B Short-answer Questions		
	I	II	III	I	II	III
Fundamentals						
1 Atomic Theory 1.1 –1.2	1					
2 Ohm's Law 2.1 – 2.4	1					1
3 Series & Parallel Circuits 3.1 – 3.6		1	1			
4 Voltage Dividers 4.1 – 4.3		1				
5 Power 5.1 – 5.3		1	1			
Electronic Components						
6 Resistors 6.1 – 6.3	1					1
7 Capacitors 7.1 – 7.3	1					1
8 Inductors 8.1 – 8.2		1			1	
9 Switches & Fuses 9.1 – 9.3		1			1	
AC Theory						
10 AC Circuits 10.1 – 10.5		1	2			
Semiconductors						
11 Diodes 11.1 – 11.3					1	
12 Transistors 12.1 – 12.6		1		1		1
Sub-total	4	7	4	1	3	4
Total	15			8		
	23					

Taxonomic Level	Expected Distribution (determined by Orgn)	Actual Distribution of Test Items	Actual Distribution (%)
I - Knowledge	Max 30%	6	24%
II - Comprehension	Min 35%	9	36%
III - Application	Min 30%	8	32%

CHAPTER 7: TYPES OF TEST ITEMS

Introduction

- 1 There are many types of test items that can be used to assess knowledge learning and achievement. These will include:

Selection-type Questions

- True-False (T/F);
- Matching;
- Multiple choice (MCQ);

Supply-type Questions

- Short-answer;
- Restricted response essay; and
- Extended response essay.

- 2 Each type of test item has its own characteristics, uses, advantages, limitations and rules of construction. Hence, the selection of the types of questions to use for an assessment could have an impact on the achievement of test objectives. It is important at the test plan development stage to select the types of questions that will be most effective in assessing learning outcome.
- 3 As seen in Para 1, types of test questions can be grouped under 2 categories ie Selection-type and Supply-type. Selection-type refers to questions that have responses (answers) contained in them. All the candidate needs to do is to pick an answer (as in the case of True/False questions) or select an option from a list as the answer (as in the case of Multiple choice questions).
- 4 On the other, supply-type will have questions that require the candidate to provide the responses, which can range from providing a few words (short-answer) to many sentences as expected of an essay-type question.
- 5 Items may be classified as Objective or Subjective questions, as well. Objective-type questions are those that are highly structured and they limit the response a candidate can make. Each objective question has only one correct or best answer. Scoring can be consistent and done fast and with ease. Selection type of questions including short-answer questions belong to the group of objective questions.
- 6 Subjective questions are those essay questions classified under the supply-type of questions viz restricted-response and extended response questions. Scoring of subjective questions tends to judgemental and may not be consistent.

True/False Questions

- 7 A true-false item consists of a statement and the candidate is asked to pick one of 2 answers provided. The response could be: true or false; right or wrong; correct or incorrect; yes or no; fact or opinion; agree or disagree. Of the 2 possible answers provided, only one is the best or correct answer.

- 8 An example of a True-False item is as follows;

<i>Read the statement. Circle T if it is true. Circle F if it is false.</i>	
1. Water is used to fight fires in electrical installations.	T / F

- 9 True-False items are versatile and useful for assessing thinking abilities at the lower taxonomic levels of knowledge, comprehension and application (see examples below):

- Testing for factual knowledge.

The chemical formula for water is H ₂ O	T / F
--	-------

- Testing for comprehension.

Stroke occurs when blood supply to the heart is blocked.	T / F
--	-------

- Testing for application

If heat is supplied at a constant rate to vapourise a substance, the temperature of the substance will increase at a constant rate.	T / F
---	-------

- Testing for problem-solving ability

Given the general gas law $PV/T = k$, Where $k=2$, $P=5$, $T=50$ C, then $V = \underline{20}$	T / F
--	-------

Advantages and Disadvantages of True-False Questions

- 10 Some of the advantages and disadvantages of True-False questions are tabulated below:

Advantages	Disadvantages
• Good for learners who are not good in reading or have a low understanding of the language • Are adaptable to most content areas • Can cover a larger amount of subject matter in a test (ie more questions can be asked) • Are particularly suitable for testing beliefs in popular misconceptions and superstitions • If well constructed, can be used to assess higher intellectual abilities eg at taxonomic level of application, analysis and synthesis • Can be marked quickly, reliably and objectively by marker or assessor who are not content experts	• Candidates' responses could be provided through guesswork because there only 2 possible answers. May affect reliability of scores. • Can be vague if poorly worded. Candidate may misinterpret or misread the statement. This may result in low reliability of scores. • Candidate may find it easy to cheat during the test. He can work out a system in which a smarter candidate can signal the correct answers to him while the test is in progress. • May encourage candidates to use a particular pattern in providing their responses without giving much thought in providing the answers. If they do well as a result of the pattern, the reliability of the scores may be affected. • Difficult for setter to write truly 'true' or 'false' items. For example, the statement "Water boils at 100 C" is true only at sea level. The boiling point depends on air pressure. The lower the pressure, the lower the boiling point.

Matching Questions

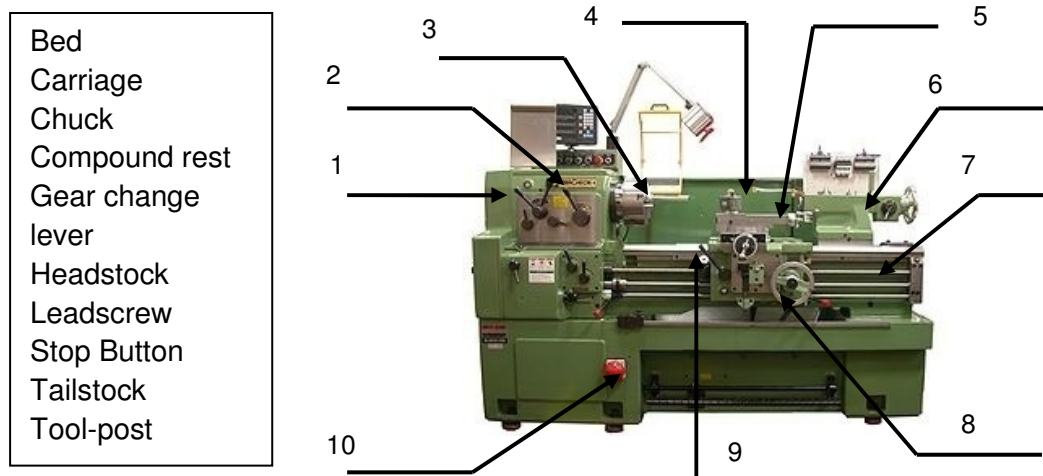
- 11 The traditional matching form consists of two parallel columns with each word, number or symbol in one column being matched to a word, phrase or sentence in the other column. The items must have a relationship e.g. parts of a machine and their corresponding functions.
- 12 There are 4 ways of presenting the matching exercise. They are:
- The traditional format;
 - The diagram format;
 - The classification format; and
 - The sequence format

- 13 The traditional format consists of 2 columns. One column consists of problems (premises) and the other of answers (responses). The candidate is required to match each premise with each response. See example below.

Answer	Column 1(premises)	Column 2 (responses)
()	Australia	Seoul
()	People's Republic of China	Berlin
()	Malaysia	Paris
()	United States of America	Kuala Lumpur
()	Germany	Washington DC
()	South Korea	Canberra
()	France	Beijing

- 14 In the diagram format, the candidate is given a diagram of an item of equipment with its parts numbered, and a name list of the parts. The candidate is required to match the name of a part with the number on the diagram.

The picture shows a Lathe machine. Label the parts of the machine using the words from the list provided.



- 15 In the classification format, the candidate is presented with a list of homogeneous items and a classification scheme. The candidate is required to classify each homogeneous item according to the scheme.

A list of 12 words is given in Column A. Place each word under its correct classification of Adjective or Abstract Noun.

Column A – List of Words		Adjective	Abstract Noun
Accuracy	Hardship		
Ambitious	Lonesome		
Bravery	Occupancy		
Fantastic	Patience		
Generous	Persistent		
Happiness	Strong		

- 16 In the sequence format, the steps for doing a certain job are presented in a jumbled order. The candidate is required to arrange them in the correct sequence.

A list of actions to prepare a cup of coffee is provided below. Re-arrange the actions, according to the company's Standard Operating Procedure (SOP) and enter the appropriate letters in the space provided.

List of Action Items	SOP Sequence Order	
A. Scoop one teaspoon of instant coffee and put it into a cup.	Step 1	
B. Stir well and remove spoon.	Step 2	
C. Place a clean spoon onto the saucer and serve customer.	Step 3	
D. Boil water using kettle.	Step 4	
E. When boiled, pour hot water into cup up to $\frac{3}{4}$ full.	Step 5	
F. Add 2 cubes of sugar and 1 teaspoon of condensed milk.	Step 6	
G. Prepare coffee cup and sauce	Step 7	

Possible Relationships For Matching

- 17 The matching exercise can be used to measure relationships in a variety of fields. Some of these relationships are given below:

Persons	Achievements
Dates	Historical events
Terms	Definitions
Rules	Examples
Symbols	Concepts
Authors	Titles of books
Machines	Uses
Plants/Animals	Classifications
Principles	Illustrations
Objects	Names
Parts	Functions

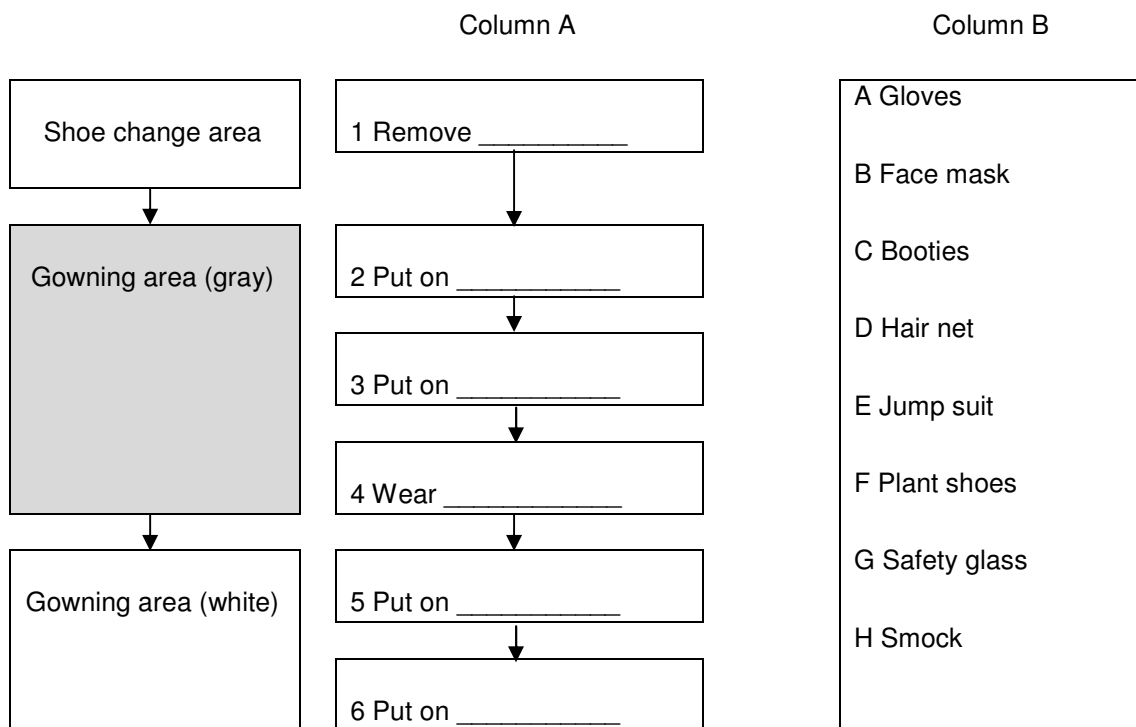
Example 1

Column A contains a list of characteristics of metals. In the bracket to the left of the statement, write the letter of the metal in Column B which fits the statement. You may use a letter in column B once, more than once, or not at all

Column A	Column B
() A reddish brown colour metal which is a high conductor of electricity.	A stainless steel
() Has high tensile strength and low cost for making machine components	B bronze
() Light in weight and used in making aircraft bodies.	C mild steel
() A ferrous metal which withstands corrosion.	D aluminium
() It is a mixture mainly of two metals one of which is tin.	E copper
	F silver

Example 2

Column A shows the flowchart of a Gowning Procedure. Write the letter of the safety wear in Column B into the blank space in Column A which fits the procedure.



Advantages and Disadvantages of Matching Questions

- 18 Some of the advantages and disadvantages of Matching Questions are tabulated below.

Advantages	Disadvantages
• Test items require relatively little reading time; hence many questions can be asked and this allows for a larger sampling of content. • Test items can be marked easily. The marker need not be an expert in the subject. • The matching exercise is the most compact and efficient way of measuring the learner's ability to identify the relationship between two things.	• If not carefully prepared, the matching lists may encourage serial memorization rather than association. • It is difficult to get clusters of questions that are sufficiently alike that a common set of responses can be used. • The matching exercise encourages rote learning as it can be used to measure only factual information.

Multiple choice Questions

- 19 A multiple choice item consists of 2 parts: a stem which contains the problem, and a list of suggested answers (responses or options). The candidate is required to read the stem and then select the best or correct answer from the list of options. The correct answer is called the key. The incorrect responses are called distracters.
- 20 The stem may be written as a direct question or an incomplete statement. The direct question format is generally preferred because it is easier to write and reduces the possibility of giving grammatical clues. The examples below show the different formats of a multiple choice test item.

Example 1

What is the national flower of Singapore?

- (A) Hibiscus. (distracter)
- (B) Lily. (distracter)
- (C) Orchid. (key)
- (D) Rose. (distracter)

Note: When the stem is a question, the options will start with a **Capital** letter.

Example 2

The national flower of Singapore is the

- (A) hibiscus.
- (B) lily.
- (C) orchid.
- (D) rose.

Note: When the stem is part of a sentence, the options will start with a letter in the **lower** case.

Taxonomic Level of Test Item

- 21 The 3 levels of questioning based on the 3 lower taxonomic levels are given below:

- **Level I:** Knowledge / Recall items test the ability to recall facts, principles or methods.

What type of grated cheese should be served with minestrone?

- (A) Bel Paese.
- (B) Cheddar.
- (C) Cheshire.
- (D) Parmesan.

Note: For single-word options, arrange them according to alpha order. A name is treated as a single word as in Bel Paese, which is a name of a kind of cheese.

- **Level II:** Comprehension / Understanding items test the ability to describe or explain a principle or method.

The MAIN reason for using flux when soldering is to

- (A) remove metallic impurities.
- (B) prevent oxidation of the surface.
- (C) enable the solder to melt more quickly.
- (D) increase the mechanical strength of the joint.

Note: For options with 2 or more words, the options should be arranged from short to long.

- **Level III:** Application items test the ability to apply facts, principles or methods to solve problems.

If engine power is increased during steady level flight and without any other changes being made to the aircraft, the aircraft will

- (A) accelerate and climb.
- (B) accelerate and lose altitude.
- (C) remain at constant speed and climb.
- (D) accelerate and remain at constant height.

Advantages and Disadvantages of Multiple choice Questions

- 22 Some of the advantages and disadvantages of multiple choice questions are tabulated below:

Advantages	Disadvantages
• Is versatile. It can be used to measure the candidate's ability to recall facts, to reason, and to exercise judgement. • Can be marked quickly and accurately. The marker does not have to be a content expert. • Is relatively more efficient than the essay question in terms of the number of questions that can be asked in a prescribed time, and the space needed to present the questions. • The degree of difficulty of the test can be controlled by changing the degree of homogeneity of the responses. In other words, the test item can be made easy or difficult by making the responses less or more homogeneous.	• Is difficult to construct. It is not easy to think of plausible distracters. • Requires time to respond, especially when very fine discriminations have to be made. • Cannot measure a candidate's ability to organize and present ideas.

• Is less susceptible to score variations than true-false items because the chances of guessing are reduced. • Is free from response sets. In other words, the candidate is unlikely to give a different answer when the same content is presented in a different form.	
--	--

Short Answer Questions

23 A short-answer item is classified as a supply-type objective question. The candidate is required to write his answer to the question with a word, phrase, mathematical symbol, formula, and so on.

24 The 2 common varieties of the short-answer form are:

- The Question variety

What is the current GST rate levied on most goods and services in Singapore?

- The Completion variety

The current GST rate levied on most goods and services is

Taxonomic Level of Test Item

25 Examples of questions asked at the 3 lower taxonomic levels are given below:

- **Level I**: Knowledge / Recall items test the ability to recall facts, principles or methods

What are the 2 main metals which form bronze?

- **Level II**: Comprehension / Understanding items test the ability to describe or explain a principle or method

Give TWO examples of fixed costs in a small enterprise.

- **Level III**: Application items test the ability to apply facts, principles or method to solve problems

How much liquid can a spherical tank hold if its diameter is 2 metres?
(Volume of a sphere is $\frac{4}{3} \pi r^3$)

- 26 Some of the advantages and disadvantages of short-answer questions are tabulated below:

Advantages	Disadvantages
- Short-answer items are useful in tests, where specific bits of information are usually tested. They are useful for testing the knowledge of definitions and technical terms. - Short-answer items are especially useful in mathematics, science and technology, where a formula or equation is to be written. - These items can also test understanding, application, computational and problem-solving skills in mathematics and technology, as illustrated below: The sum of twice a number and 7 is 31. What is the number? To solve this problem, the candidate would first have to introduce an unknown X to stand for the number he is looking for. He would then have to translate the statement into an equation such that $2x + 7 = 31$. Then he would solve for x. - Short-answer items are suitable for measuring highly specific facts, such as names of tools, equipment, materials, devices, components and their functions, safety measures.	- Short-answer items are limited to questions that can be answered by a word, phrase, symbol, or number. There are very few instances when an abstraction, generalisation or interpretation can be adequately presented by the candidate in one or 2 words. As a result, most short item items measure factual recall. - Because short-answer items are best for measuring highly specific facts (dates, names, places, etc), excessive use may encourage rote memory and poor study recall. - Marking may be tedious, inaccurate and somewhat subjective because of the variety of acceptable answers. Frequently the marker must decide whether a candidate's answer is right, wrong, or partially right. This can lead to bias in marking and lower the reliability and validity of the test. - Furthermore, because of the variety of acceptable answers, the marker must be a content expert who knows the subject thoroughly.

Essay Questions

- 27 Essay questions require the candidates to supply the answers. They are subdivided into 2 major types, as follows:

- Extended response
- Restricted response

- 28 The extended response question places almost no restrictions on the candidate as to the points he can discuss and the way he organises his answer. An example of an extended response essay question is as follows:

Discuss the importance of banking and finance sector to the development of the Singapore economy.

(Your answer should not exceed two pages. It will be evaluated on its organisation, comprehensiveness, relevance and clarity)

- 29 In answering this question, the candidate has to select those aspects of banking and financial sector and services that he thinks are most important and relevant to his argument. He may organise the material in whatever way he wishes. In short, the extended response question permits the candidate to demonstrate his ability to:
- Call upon factual knowledge;
 - Evaluate his factual knowledge;
 - Organise his ideas; and
 - Present his ideas in a logical and coherent manner.
- 30 It is at the levels of synthesis and evaluation of writing skills (style and quality) that this type of question makes the greatest contribution.
- 31 The restricted response question limits the candidate in the form and scope of his answer because he is told specifically the context that his answer is to take. An example of a restricted response essay question is:
- A 4.5 V electric bell is fed from a 240-V supply.

(A) Make a labelled sketch of the electric bell circuit.

(B) Describe the operation of the bell circuit.
- 32 This question is restrictive in that the candidate need only address himself to specified areas.
- 33 By aiming the candidate at the desired response, we minimise the problems of reliable marking and make marking easier. But by restricting the candidate's response, we give up one of the major advantages of the essay question - a measure of the candidate's ability to synthesise his ideas and express them in a logical, coherent fashion. Because of this, the restricted response question is of greatest value for measuring learning outcomes at the comprehension, application and analysis levels.

Advantages and Disadvantages of Essay Questions

- 34 Some of the advantages and disadvantages of essay questions are tabulated below:

Advantages	Disadvantages
• The essay test is relatively easier to prepare than an objective test. • The essay test is the only means available to assess a candidate's ability to compose an answer and present it in writing. • The essay question tests the candidate's ability to supply rather than select the correct answer. • The essay test provides a measure of complex learning outcomes which cannot be so well measured by other means.	• The essay test offers limited content sampling, i.e. you can test only a limited number of topics covered in the course content. • The essay test offers low reliability. In other words, a candidate's score may depend on who marks the paper, rather than on the quality of the written answer. • The candidate does not always understand the essay question and therefore is not sure how to answer it. • The essay test is time-consuming and laborious to read and mark the answer. • Also, the essay test has to be marked by a content expert who knows the subject thoroughly.

CHAPTER 8: GUIDELINES FOR CONSTRUCTION OF THEORY QUESTIONS

Introduction

- 1 While there are different types of test items such as multiple choice, short answer and essay questions that can be used appropriately to assess candidates' thinking abilities at different taxonomic levels, the way questions are constructed can have an effect on the reliability of the scores. There are guidelines developed to guide setters of theory papers in the construction of high quality questions.
- 2 The guidelines for the various types of test items are provided in the ensuing paragraphs.

Guidelines for Writing True-False Items

- 3 Avoid using ambiguous words and statements. Ambiguity is often present when an item deals with qualitative terms such as MORE, FEW, HEAVIER and LIGHTER. Since these words are somewhat vague, indefinite and imprecise, they may be interpreted differently by different people. Wherever possible, use quantitative rather than qualitative terminology.
- 4 Base true-false items on statements that are clearly true or clearly false. If the item does not satisfy this condition and if an element of ambiguity or trickiness is injected into the item, it will confuse the candidate.
- 5 Avoid using negative statements, and especially double negatives. Use simple and clear language.
- 6 Avoid having "trick" questions. Remember that candidates are under pressure when they take tests. They tend to read rapidly and pay little attention to "tricky" wording. One of the objectives in testing is to see what the candidate has learned, not whether he can be tricked. "Trick" questions affect the validity of a test.
- 7 When testing for cause-and-effect relationship, always make the first proposition in the statement true and the second part (reason or explanation) either true or false.
- 8 Avoid using specific determiners which provide irrelevant clues. Qualifiers such as USUALLY, SOME, GENERALLY, SHOULD, SOMETIMES and MAY tend to make the statement true. On the other hand, qualifiers such as ALWAYS, NEVER, NONE and ALL tend to make an item false.
- 9 Avoid lifting statements verbatim (word for word) from textbooks and notes. Rephrase all textual material or put it in a new context to discourage rote memory.

- 10 Avoid making true statements consistently longer than false statements. Setters, in their attempt to ensure absolute truth in a statement, sometimes add qualifying statements that make the item longer. Candidates readily learn to recognise this irrelevant clue. To minimise this effect, vary the length of both true and false statements, or make all statements of about the same length.
- 11 Avoid having a disproportionate number of either true or false items. Some candidates tend to mark items they are unsure of as TRUE, while other candidates tend to mark the items, FALSE. By having approximately an equal number of true and false statements in a test, the setter limits the response of this tendency on the validity or the test score.
- 12 Write clear instructions on how the student should indicate his choice of answer e.g.

‘Read the statement. Circle T if it is true. Circle F if it is false.’

Variations

- 13 True-False items may be made more challenging as shown by the examples below:

Example 1

Circle T if the statement is true and F if it is false. Change the underlined word to make the statement true. Place the new word in the blank space next to F.

T ☒ F <u>electrons</u>	1 Particles of negatively charged electricity are called <u>neutrons</u>
T ☒ F _____	2 Mechanical energy is turned into electrical energy by means of the <u>generator</u>
T F <u>store</u>	3 An electric condenser is used to <u>generate</u> electricity

Example 2

A precision machinist is asked to machine a stainless steel component to a fine surface finish. Which of the following information items are essential for him to achieve the task?

Circle Yes if it is essential. Circle No if it is not essential.

Yes	No	1	Cutting speed of stainless steel
Yes	No	2	Diameter of milling cutter
Yes	No	3	Size of the machine
Yes	No	4	Type of coolant
Yes	No	5	Weight of the component
Yes	No	6	Rate of cutting feed

Example 3:

Which of these statements about energy and energy conversion are true?
(Check the box with a tick if the statement is true.)

- ☐ Thermal energy may be defined as light energy.
- ☐ Chemical to thermal energy conversion occurs when petrol burns.
- ☐ Mechanical energy is defined as the ability to do work.
- ☐ Mechanical to electrical energy conversion occurs when the engine drives the generator.

Guidelines for Writing Matching Test Items

- 14 Ensure that each matching exercise consists of homogeneous items, i.e. items dealing with only a single concept, classification, or area.
- 15 Put the problems in the premise column on the left-hand side of the page and the answers in the response column on the right-hand side of the page, to facilitate easy reading from left to right.
- 16 Keep each list of items in a single matching exercise relatively short. The number of premises and responses in each list should ordinarily range from 5 to 12. It is generally difficult to maintain the homogeneity of material if the list is too long. Also, it is very tedious for the candidate to go through an excessively long list.
- 17 Avoid having an equal number of premises and responses. To reduce the chances for successful guessing, have 2 or 3 more responses than premises. Alternatively, provide fewer responses than premises, and instruct the candidate that the responses may be used once, more than once, or not at all.

- 18 Arrange the responses in some systematic fashion. Words should be listed in alphabetical order; dates and numbers in either ascending or descending order.
- 19 Check that every response is a plausible answer to every premise in a single matching exercise. This is to minimise the opportunity for successful guessing.
- 20 Maintain grammatical consistency. Use all proper names or all common nouns; all singulars or all plurals. Mixing the list may provide unintended clues,
- 21 Give clear instructions on how each match is to be made, e.g.

In the bracket to the left of each function in Column A, write the letter of the component in Column B which fits the statement. You may use a letter in Column B once, more than once, or not at all.

- 22 Place all of the items for one matching exercise on the same page so that the candidate can scan the responses with ease.

Guidelines for Writing Multiple choice Items

- 23 Ensure that the stem of the item is meaningful by itself and presents a definite problem. The candidate should be able to understand what is expected of him when he reads the stem before he begins reading the responses. Compare the 2 items as follows:

- Farmers
 - (A) rotate their crops to conserve the soil.
 - (B) set up co-operatives to buy farm equipment cheaply.
 - (C) need uniform working conditions throughout the year.
 - (D) need to mechanise to increase farm productivity.
- Farmers rotate their crops to
 - (A) conserve the soil.
 - (B) make marketing easier.
 - (C) improve the soil texture.
 - (D) provide for strip cropping.

Note: For options with 2 or more words, the options should be arranged from short to long.

24 Avoid repetition of words in the options. The stem should be written to incorporate key words. This saves the candidate reading time and helps to focus his attention on the problem. Look at the examples as follows:

- According to Engel's Law,
 - (A) family expenditures for food rise in proportion to income.
 - (B) family expenditures for food decrease as income increases.
 - (C) family expenditures for food increase as family size increases.
 - (D) family expenditures for food require a smaller percentage of an increasing income.
- According to Engel's Law, family expenditures for food
 - (A) rise in proportion to income.
 - (B) decrease as income increases.
 - (C) increase as family size increases.
 - (D) require a smaller percentage of an increasing income.

25 Use clear and concise words. The stem should be easy to read and understand, and free of elaborate explanations. Compare these examples:

- Test standardization is the process whereby a common set of administrative and scoring directions is established. It is usually the major distinguishing factor between the teacher-made and commercially-published test. Many steps are involved in test standardization. Some of these steps are rigorously adhered to, while others are not. Of the following steps in test standardization, the one usually neglected is
 - (A) finding the group mean.
 - (B) trying out each item in a pretest.
 - (C) furnishing adequate validity information.
 - (D) reporting the standard deviation of the sample.
- Some steps in test standardization are listed below. The step most often neglected is
 - (A) finding the group mean.
 - (B) trying out each item in a pretest.
 - (C) furnishing adequate validity information.
 - (D) reporting the standard deviation of the sample.

26 Avoid using a negatively-stated item stem as far as possible. A negatively-stated item stem tends to measure insignificant learning outcomes and confuse the candidate. Compare these items:

- Which one of the following items is not a fruit?

- (A) Apple.
- (B) Cherry.
- (C) Papaya.
- (D) Potato.

Note: if 'not' is to be used, highlight it. In other words, use ' NOT ' instead of 'not'

- Which one of the following items is a vegetable?

- (A) Apple.
- (B) Cherry.
- (C) Papaya.
- (D) Potato.

27 Ensure that all of the options are grammatically consistent with the stem of the item. This is to prevent giving unintended clues to the candidate. Compare the items as follows:

- The main purpose of the United Nations is

- (A) forming new governments.
- (B) to maintain peace in the world.
- (C) concerned with military control.
- (D) for establishing international law.

- The main purpose of the United Nations is to

- (A) provide military control.
- (B) form new governments.
- (C) establish international law.
- (D) maintain peace in the world.

28 Ensure that all options are plausible and homogeneous with regard to the knowledge being measured. The candidate is then forced to read and consider all the options. Compare these examples:

- Who invented the telephone?

(A) Adolf Hitler.
(B) Alexander Bell.
(C) Michael Jackson.
(D) Winston Churchill.

- Who invented the telephone?

(A) Alexander Bell.
(B) Guglielmo Marconi.
(C) Samuel Morse.
(D) Thomas Edison.

29 Provide only one best or correct answer to every item. As the best answer item is more susceptible to ambiguity, the correct answer item is preferred. Compare these items:

- The Republic of Singapore borders on

(A) Brunei.
(B) Indonesia.
(C) Malaysia.
(D) Philippines.

- The immediate neighbour of Singapore is

(A) Brunei.
(B) Indonesia.
(C) Malaysia.
(D) Philippines.

30 Avoid overlapping options. These create confusion because if one option is correct, another option also has to be correct since one may incorporate the other. Look at these items:

- The average weight of the adult Singapore female is less than

- (A) 45 kg.
- (B) 50 kg.
- (C) 55 kg.
- (D) 60 kg.

- The average weight of the adult Singapore female is

- (A) less than 45 kg.
- (B) between 45 to 49 kg.
- (C) between 50 and 54 kg.
- (D) more than 54 kg.

Note: in this case, the arrangement of short to long options does not apply. It is treated as an exception to the rule.

31 Place options at the end of the statement when the incomplete statement format is used. The problem should be presented in the stem as early as possible. Compare these examples:

- The _____ is used for measuring temperature.

- (A) barometer.
- (B) pyrometer.
- (C) thermometer.
- (D) vaporimeter.

- An instrument used to measure temperature is the

- (A) barometer.
- (B) pyrometer.
- (C) thermometer.
- (D) vaporimeter.

32 Avoid verbal associations between the stem and the correct answer. Verbal associations provide irrelevant clues and do not test learning outcomes. Compare the following examples:

- Which one of the following tools is used for tapping?
 - (A) Die.
 - (B) Drill.
 - (C) Reamer.
 - (D) Tap.

- Which one of the following tools is used for cutting an internal thread?
 - (A) Die.
 - (B) Drill.
 - (C) Reamer.
 - (D) Tap.

33 Cast the item in a novel situation to measure higher mental processes such as understanding and application. Test items should not contain problems that were used in class, otherwise they only measure recall. Look at the 2 items as follows:

- Area equals
 - (A) length + width.
 - (B) length – width.
 - (C) length x width.
 - (D) length / width.

- Assume a litre of paint covers 4 square metres. How many litres of paint would be needed to paint the walls of a room that measures 4 metres x 4 metres?
 - (A) 8.
 - (B) 12.
 - (C) 16.
 - (D) 20.

- 34 Arrange the options as simply as possible. If single words are used, arrange them in alphabetical order; if numerals or dates are used, arrange them in either ascending or descending order. Compare these examples:
- The square root of 16 is
 - (A) 4.
 - (B) 2.
 - (C) 8.
 - (D) 6.

 - The square root of 16 is
 - (A) 2.
 - (B) 4.
 - (C) 6.
 - (D) 8.
- 35 Avoid making the correct option longer than the incorrect options. Because of the need for qualifying words, the correct answer is often longer than the distracters. In such a case, lengthen the distracters with qualifiers so that they also look more plausible.
- 36 Use special options such as "none of the above" or "all of the above" sparingly. These options do not test learning outcomes as they encourage guessing. For example, if the candidate is unable to "recognise" the correct answer, he is likely to choose "none of the above" option as his answer. On the other hand, if the candidate is able to "recognise" that at least 2 answers are correct, he can automatically choose "all of the above" without having to consider the remaining options.
- 37 Use 3 to 5 options, taking into consideration the age of the candidates and the nature of the material being tested. The number of options in a test item does not affect the reliability of the test. The most common number of options used is 4.
- 38 Construct the test items such that the correct answer appears in each of the alternative positions approximately an equal number of times, but in random order. If the correct answer appears in the middle position more often than it does in the first or last position, it provides an unintended clue to an alert candidate.
- 39 Write clear instructions, for example:
- 'Read the question. Choose the correct answer and shade the appropriate circle on the separate answer sheet which is provided.'*

Other Formats

Combination of Answers

- 40 This type of questioning requires more mental processing and should be avoided for lower level candidates. This format should not be over-used in one test as the time taken to derive the answer is usually longer than other type of multiple choice questions.

Example

Which of the following are properties of a rhombus?

- (I) All the sides are equal.
 - (II) Opposite sides are parallel.
 - (III) Diagonals bisect each other at right angles.
- (A) I only.
 - (B) I and II only.
 - (C) II and III only.
 - (D) I, II and III.

Group items

- 41 An item group is a series of multiple choice questions, which refer to a given situation or a diagram presented at the beginning. The items should all be independent in that no item should suggest the key to another item nor rely on candidates being able to answer another item. Also there should not be so many items on one syllabus reference that the item group could not be accommodated in a test without distorting the test specification.

Example

Questions 1 to 4 refer to the situation described below.

A client has tinted hair and requires a permanent wave which can be dressed into soft wave movements and loose curls.

- 1 The hair should be cut by
- (A) bevel cutting.
 - (B) cutting in layers.
 - (C) thinning and tapering.
 - (D) clubbing to one length.

- 2 A special reagent should be chosen because tinted hair is likely to
- (A) take a loose curl.
 - (B) process more slowly.
 - (C) process more quickly.
 - (D) develop a darker colour.
- 3 Where should the first application of reagent be made?
- (A) At the roots.
 - (B) At the points.
 - (C) At the mid-lengths.
 - (D) 10 mm from the roots.
- 4 The hair should be wound by taking
- (A) deep meshes on large diameter curlers.
 - (B) deep meshes on small diameter curlers.
 - (C) shallow meshes on large diameter curlers.
 - (D) shallow meshes on small diameter curlers.

Guidelines for Writing Short-Answer Items

- 42 For computational problems, specify the degree of precision and the units of expression expected in the answer. Examine these questions:

- What is the product of 2.3 and 6.94 to 2 decimal places?
- If each standard tennis ball weighs 142 g, how much would 12 weigh?
- If each standard tennis ball weighs 142 g, how much would 12 weigh?
Give your answer in kg.

- 43 Omit important words only. In short-answer items, the words that are to be omitted in a sentence must be key or important words. The candidate should be asked to supply an important fact. Avoid omitting verbs, pronouns or adverbs, otherwise you may receive a wide variety of plausible answers, and the item may no longer assess the instructional objective. Look at these examples:

- Raffles _____ Singapore in 1819.
- In 1819, Raffles founded _____ .

44 Word the item so that the required answer is both brief and definite. Compare these items:

- An animal that eats meat is _____.
- An animal that eats meat is classified as _____.
- The ITE was established in _____.
- The ITE was established in the year _____.

45 Avoid excessive blanks in a single item. Excessive omission of key words may cause the item to lose its specific meaning. Omitting too many key words can result in ambiguity and confusion and may turn a test into a guessing game. Study these examples:

- In geometry, the ratio of the _____ to the _____ is called the _____.
- In geometry, the ratio of the opposite side to the adjacent side is called the _____.

46 Place the blank near the end of the sentence. When the blank is at the beginning or middle of the sentence, the candidate may overlook or forget the essential point of the question by the time he reads the item. Compare these items:

- The _____ is a non-profit organization which protects the interests and welfare of consumers in Singapore.
- The non-profit organization which protects the interests and welfare of consumers in Singapore is the _____.

47 Make sure all blanks are of equal length.

- What instrument is used to measure temperature?
_____.
- What is the machine tool used to turn a piece of metal?
_____.
- What is the instrument used to measure the internal diameter of a pipe?
_____.

48 It is generally better to use the direct question rather than the incomplete statement. The direct question is easier to phrase so that it is less ambiguous. If you want to test a candidate's knowledge of definitions or his comprehension of technical terms, use a direct question asking for a definition. Compare these items:

- The physical components of a computer are referred to as the _____ of the system.
- What is term used to refer to the physical components of a computer system?

Guidelines for Writing Essay Questions

49 Give adequate time and thought to the preparation of essay questions. Satisfy yourself that:

- Each question measures its intended objective;
- The wording is simple and clear to the candidates; and
- Every question is reasonable and can be answered by the candidates.

50 Formulate each question so that it will elicit the behaviour you want to measure.

51 Delimit the area covered by the question. Be specific. Candidates should know exactly what they have to do. The lack of specificity often reduces validity.

52 Use descriptive words and clear, simple vocabulary. Phrase each question so that the candidate's task is clearly indicated.

53 Indicate the value of each question (i.e. the number of marks it carries) and the time to be spent in answering it. The candidate can then judge for himself where he should place his emphasis in answering the questions.

54 Decide in advance what factors will be considered in evaluating an essay answer. If you intend to consider spelling, punctuation, composition, grammar, handwriting and clarity of expression, then indicate this clearly in the instructions to candidates.

55 Avoid optional questions. There are 3 main reasons for this:

- It is difficult to construct questions of equal difficulty;
- Candidates are unable to select those questions which they will be able to do best in; and
- The good candidate may be penalised because he is challenged by the more difficult and complex questions.

- 56 Furthermore, it is not possible to make valid comparisons of candidates' achievement unless they have all answered the same questions.
- 57 Use a relatively large number of questions requiring brief answers (up to one page per answer) rather than just a few questions involving long answers (2-3 pages per answer). By setting many questions, you can get a broader sampling of content, thereby increasing reliability.
- 58 Do not start essay questions with such words as LIST, WHO and WHAT. These words tend to elicit responses that require only a regurgitation of factual information. If this is the only way you can begin a question, it is probably a short-answer recall question and not an essay question.
- 59 Adapt the length of the response and the complexity of the question and answer to the maturity level of the candidates.
- 60 Use the novel type of question whenever feasible. To answer a question discussed in class or in a textbook requires little more than a good memory. But, to apply that same principle to a new situation requires a higher level of learning.
- 61 Prepare the model answers and marking scheme. Marking can be made more reliable and more objective if the setter prepares in advance the "ideal" answer and decides how the marks are to be distributed to different parts of each question.

Conduct of Pilot Test for Test Items

- 62 In the process or following the setting of test items for a theory test, it is always advisable if the test items can be tested before actual use on candidates. A test run (or sometimes referred to as pilot testing) will enable the setter to find out the quality and content relevance of the test items. The evaluation of the test run can provide useful feedback to the setter on the usefulness of the test items. With the input, it will allow the setter to assess the need of making changes to the test items, if necessary.
- 63 Changes to test items after a pilot run could take the form of:
- Review of factual contents – to ensure that information is factually correct
 - Review the choice of words of the questions – to remove ambiguity
 - Review question structure – to make it easy-to-understand
 - Review difficulty level – to be consistent with taxonomic levels prescribed
 - Check for close adherence to guidelines in constructing test items

- 64 The challenge to a setter, however, is be able to conduct a pilot test without causing the test items to lose their confidentiality. One possible way is not to carry out a test run for every test paper set, but to construct test items over a period of time and consciously pilot testing the items a few at a time to candidates of similar level. Over time, a pool of quality questions could be gradually compiled as the pilot tested questions are reviewed and improved upon.
- 65 For multiple choice questions (MCQs), a process called Item Analysis can be carried to test the suitability of the 'key' and 'distracters' in each MCQ. Through this process, quality test items (ie items that have been analysed, revised and shown to be of good quality) could be placed into a 'Item bank' from which the items can be selected randomly for repeated use.
- 66 Details on the Item Analysis process will be covered in another chapter.

CHAPTER 9: GUIDELINES ON GRADING/MARKING OF TEST ITEMS

Introduction

- 1 Marking of theory test items is an easy task to some but not easy to others. Much depends on the types of test items contained in the test paper and the content knowledge of the marker. For selection-type questions (true/false, multiple choice and matching questions), where there can only be one specific answer or response to be picked from each test item, marking will be less difficult. Furthermore, the marker will not need to be a content expert to carry out the marking task.
- 2 On the other hand, marking of supply-type questions (such as short-answer and essay questions) may not be a breeze for the marker. Test items under this type of questions call for the candidate to provide the response to each item and the marker to read and understand the response before assigning a mark to it. Objectivity of marking often depends on the content knowledge of the marker and his interpretation of the candidate's response. To improve marking objectivity, the setter normally provides a marking scheme to guide the marker in scoring the items.

Marking of Objective (Selection-type) Questions

- 3 For objective questions, the marking scheme is a simple document which contains the correct answers or responses. The marking scheme is normally prepared by the setter. Often, the correct answers are provided in the form of an alphabet, numeral or word depending on the type of test items. An example of a marking scheme for objective questions is given below.

Marking Scheme

Type of Item	Question No	Correct Answer	Mark Allocation
True/False	1	True	2 marks
	2	True	2 marks
	3	False	2 marks
	4	True	2 marks
Matching	5	3	3 marks
	6	5	3 marks
	7	6	3 marks
	8	2	3 marks
MCQ	9	A	5 marks
	10	D	5 marks
	11	C	5 marks
	12	B	5 marks
Total Possible Marks:			40 marks

Marking of Essay Questions

- 4 For the marking of essay questions (such as short answer, restricted and extended-response items), a well-prepared marking scheme will improve objectivity of marking. This, in turn, will increase reliability of scores.
- 5 It is expected of a setter of the theory test paper to provide a marking scheme for essay questions. The marking scheme should include suggested solutions (also known as model answers) to each essay question and allocation of marks.
- 6 Some of the guidelines to the marking of essay questions are as follows:
 - Read and understand the model answers and basis for mark allocation as stated in the marking scheme. The model answers should normally provide information on the scoring key ie key words or phrases to look out for before allocation of marks. An example of a marking scheme of an essay question is given at **Annex A**.
 - Check the scoring key against candidate's actual answers before allocating marks.
 - Be consistent throughout in the marking. Markers may be influenced by the first few papers they read and thereby grade either too leniently or too harshly, depending upon their initial mindset. For this reason, once the marking has started, reference should be made occasionally to the first few papers marked to ensure that the standards are being applied consistently.
 - Randomly shuffle the papers before marking them. A candidate's marks are influenced by the position of his paper.
 - Mark all the answers to one question before going on to the next question. This will enable the marker to concentrate and become thoroughly familiar with just one set of marking criteria, and not be distracted by moving from one question to another.
 - Try to mark all the answers to a particular question without interruption. If you do take a long break half-way through a question, then reread some of the first few answers when you resume marking to re-familiarise yourself with your scoring standards so that you will not change them in midstream.
 - Mark the papers without looking at the candidate's name or trying to discover his identity. (In any case, most tests only asked candidates to write their 'Index Nos' to avoid having their identities being made known.)
 - Decide on provisions for handling factors which are not relevant to the learning outcomes being measured. (In other words, decide on whether marks should be deducted for errors as in spelling, grammar, punctuation, legibility, etc.)

- If possible, appoint 2 independent markers for the test, and use the average as the final score. Double marking by 2 different persons makes the scores more reliable. If it is not possible to have 2 different markers, then consider marking the papers twice, several days apart.
- Provide written comments on the answer scripts if there are errors during the marking process. One value of doing this is that the allocation of marks can be explained easily should an appeal be made. In the event that the test papers are returned to the candidates, the comments and corrections could serve as valuable feedback to the candidates.
- Set reasonable standards during marking; do not be too strict or too lenient.

Bias Affecting Grading/Marking of Essay Answers

7 Bias in marking can adversely affect the marking of essay answers. The more common ones are:

- **Halo effect:** the marker's general impression of the candidate affects the marks given. A good impression could lead to more marks being accorded.
- **Item-to-item carryover effect:** the marker's impression of the candidate's performance on the first question colours the judgment of the second item
- **Test-to-test carryover effect:** the grade given to a paper is affected by the grade given to a preceding paper
- **Order effect:** the order in which a paper appears affects the marks given. Papers which are marked earlier tend to be rated higher than those marked near the end
- **Language mechanics effect:** the candidate's spelling, punctuation, grammar, length of answer and handwriting affect the marks given.

EXAMPLE OF MARKING SCHEME FOR ESSAY QUESTIONS

Question	Outline Solution (suggested answer)	Allocation of Mark
<u>2 (a)</u> It is reported that a d.c. shunt motor in a workshop is running "very hot". Suggest THREE possible causes and in each case explain the procedures necessary to establish the cause of the troubles.	<u>2 (a)</u> Accept any 3 of the following reasons. (1 mark each) (I) Overloading (II) Faulty bearing (III) Lack of adequate ventilation (IV) Faulty windings - short circuit fault Procedures 3 required according to answers given above. (2 marks each) For (I) Check motor current against rating For (II) Examine bearings – usually localised over-heating. Together with rumbling/knocking at bearing. When stopped the "lift" can be checked. For (III) Check flow of ventilating air, examine fan and vent holes. For (IV) Check insulation of, windings and continuity volt drop test on armature.	3 marks 6 marks
<u>2(b)</u> State TWO faults which might cause severe sparking at the brushes of a d.c. motor when running on load.	<u>2(b)</u> Accept answers below. (3 marks each) (I) Badly worn commutator or brushes. (II) Open or short circuited armature windings.	6 marks
Total Possible Marks:		15 marks

CHAPTER 10: DEVELOP A THEORY TEST PAPER

Introduction

- 1 Based on the test plan provided, an appointed setter should be able to proceed to develop the theory test paper. Much of the information needed should have been provided in the test plan. What is needed for the setter to do is set the test items according to the table of specifications given in the test plan and put them in the form of a test paper. In addition, the setter will need to draw up a marking scheme to provide the suggested solutions and allocation of marks.

Development of Theory Test Paper

- 2 An example of a developed theory test paper is given at Annex A.
- 3 In general, the approach in developing a theory test paper is as follows:

3.1 Make Preparation

- a) Collect relevant textbooks, notes and reference materials on the module or course to be tested. Make reference to past test papers if possible.
- b) Have a copy of the skills standard on hand.
- c) Make reference to the test plan containing the **table of specifications**. Key information needed to set the paper is as follows:
 - Number of test items;
 - Type of test item (e.g. MCQs);
 - Taxonomic level of item;
 - Knowledge items from skills standard; and
 - Duration of test.

3.2 Construct Test Items

- a) Set the test items according to the instructions given in the test plan. If there are 2 types of test items, set all the questions to one type first before proceeding to the next.
- b) Refer to the skills standard for the learning objectives and the knowledge materials contained in textbooks and notes used by the candidates.

c) Set questions based on:

- The learning objectives and topics covered in the module or course;
- The taxonomic levels (level I II or III) and types of test items (MCQs or short-answer questions) indicated in the table of specifications. (Do not change or alter the taxonomic levels prescribed for the test).

d) Refer to relevant guidelines in constructing the test items.

e) When finalizing the questions, double check against the required taxonomic levels.

f) Make note of the correct solutions or answers to the questions set. This information will be required later for the preparation of the marking scheme.

3.3 Format and Develop Test Paper

a) After constructing the test items, the next step is to assemble the test items into a test paper. (The test paper should follow a consistent style and format to give it a pleasant feel and look to the candidates. The margins, spacing of the text, font type and size should contribute to the quality of the presentation.)

b) The layout of the paper should be neat and spacious so that the content is easy to read and pleasant to look at. The individual test items must be consistent in visual presentation and must not pose any reading difficulties to the test candidate.

c) Use plain A4-size paper for the test paper. Provide a **cover page** which should contain both general and specific instructions for candidates to note. Refer to the test plan for most the information needed for the cover page.

d) The numbering of questions must be consistent to avoid confusion especially when candidates are asked to record answers in a separate booklet or paper. Use standard scientific symbols and units, abbreviations and conventions.

Note: refer to Annex A for example of a developed theory test paper.

3.4 Prepare Marking Scheme

- a) Prepare the marking scheme for the test paper. The marking scheme should show clearly the suggested solutions (correct answers) and allocation of marks (refer to Chapter 9 on Guidelines to grading/marking of test items).
- b) Next include the marking scheme below the marking scheme so as to enable the marker to grade the candidate based on his score attained (refer to the test plan for the grading system provided)
- c) Submit the marking scheme and grading system together with the developed test paper to the proper authority within the organization.

Confidentiality of Test Paper

- 4 The developed test paper and marking scheme are to be submitted to the proper authority in the organization for moderation (if necessary) and safe keeping. Maintaining confidentiality of test paper information is a responsibility of any setter. In other words, information of the test items and marking scheme must not leak out to anyone and anytime.

An example for illustration purpose

NRIC No./FIN

--	--	--	--	--	--	--	--	--	--

CERTIFICATE in DENTAL ASSISTING**END OF MODULE TEST**

ORAL BIOLOGY

THEORY PAPER

2 HOURS

APRIL 2010

INSTRUCTIONS TO CANDIDATES

- 1 This paper comprises two sections:
Section A – 25 multiple-choice questions.
Section B – 10 short answer questions.
- 2 The marks for each section are as follows:
Section A – 50 marks.
Section B – 50 marks
- 3 Attempt all questions in Sections A and B.
- 4 For Section A, choose the correct answer and shade the appropriate circle on the separate answer sheet provided.
- 5 For Section B, write your answers on the blanks provided in this question paper.
- 6 Write your NRIC number on the top right corner of each answer sheet.
- 7 Return this question paper stapled together with the Section A answer sheet.

This question paper consists of 8 printed pages (including this page)

Section A: Multiple Choice Questions. Answer all 25 questions (2 marks per question)

Choose the correct answer A, B, C or D and shade the appropriate circle on the answer sheet provided. Make sure you write your NRIC No on each answer sheet.

- 1 The bite-wing radiograph is used for detecting
 - (A) maxillary sinus.
 - (B) abscessed teeth.
 - (C) mandibular canal.
 - (D) interproximal decay.

- 2 Sodium hypochlorite is commonly used in
 - (A) implant surgery.
 - (B) dental prophylaxis.
 - (C) denture fabrication.
 - (D) endodontic treatment.

- 3 Which one of the following items will **NOT** affect the setting time of alginate impression materials?
 - (A) Mixing time.
 - (B) Powder/liquid ratio.
 - (C) Temperature of the water used.
 - (D) Uneven loading of the alginate impression material onto the tray.

- 4 Which one of the following statements describes waste management, which is **NOT** appropriate in the dental operator?
 - (A) Saliva coated items must be handled as contaminated waste.
 - (B) Sharp containers must be placed outside the dental operator.
 - (C) Appropriate PPE must be worn when handling contaminated waste.
 - (D) General waste can be disposed into covered bins lined with plastic bags.

- 5 -----

- 24 -----

- 25 -----

Section B: Short Answer Questions. Answer all 10 questions (5 marks per question)

Answer each question with a word, a phrase or a short sentence. Write your answer on the blank space provided below.

26 Name the two types of evacuators used in operative dental procedures?

27 Name the instruments that are found in the basic set-up tray.

28 Define the term 'indirect vision'.

29 What precautions should be taken when children use fluoridated toothpaste?

30

31

35

*** End ***

SOLUTIONS AND MARKING SCHEME

Section A: 25 Multiple Choice Questions (2 marks each)

Question No	Correct Answer	Award of Marks
1	B	2
2	D	2
3	D	2
4	D	2
5		
24		
25		

Continue until Question 25

Section B: 10 Short Answer Questions (5 marks each)

Question No	Outline Solution (suggestion answer)	Award of marks
26	1) saliva ejector 2) high-volume evacuators	2 correct~5 marks. only one correct~1 mark
27	1) dental mirror 2) dental explorer 3) cotton pliers	3 correct~ 5 marks 2 correct~ 3 marks 1 correct~ 1 mark
28	'Indirect vision' means 'viewing an object through a dental mirror'	5 marks if fully correct. award 3 marks if it is just acceptable; 0 for unacceptable answer
29	1) Use under adult supervision 2) remind not to swallow toothpaste 3) use child strength 500 ppm toothpaste	3 correct~ 5 marks 2 correct~ 3 marks 1 correct~ 1 mark
30		

Continue until Question 35

Grading System: Award a grade based the total marks scored

Score	Grade
80% and above	A - Excellent
70% - 79%	B - Good
60% - 69%	C - Credit
50% - 59%	D - Pass
Below 50%	F - Fail

CHAPTER 11: CONDUCT OF THEORY ASSESSMENT

Introduction

1. This chapter will cover areas related to the conduct of theory assessment of learners. The premise is that although assessments or tests may be well designed and constructed, the outcomes of the tests (test scores) may not necessarily reflect a high level of reliability. Hence, the way tests are organised and conducted could have a direct impact on the reliability of the test scores.
2. The likely areas that could pose a concern in the conduct of theory assessment of learners include:
 - Setting up of assessment venue;
 - Rules and regulations for conduct of assessment;
 - Discharge of roles and responsibilities by assessors (including compliance with code of practice);
 - Matching evidence collected to assessment criteria and its impact on assessment outcome;
 - Appropriate forms are used to collect evidence;
 - Overcoming limitations in assessment evidence collection;
 - Provision of constructive feedback to learners; and
 - Improvement to the assessment process.

Setting up of Assessment Venue

3. Venue for assessment should be selected based on its suitability. The chosen venue should satisfy all the physical requirements needed for the conduct of the theory assessment. The supervisor or person in-charge of the test venue is normally the one who will oversee all the necessary preparations to get the venue ready for the theory test.
4. The physical requirements for the conduct of a theory test are generally less demanding than that for a practical test. A theory test normally requires a classroom that is able to accommodate the prescribed number of candidates. It should have, at least, basic furniture such as desks and chairs. The room should be well ventilated with proper lighting. In addition, the room should be located at a place where it is least affected by noise.
5. How well the theory test venue is set up will have a bearing not only on the test itself but also on the candidates as well. For example, if the venue is

stuffy and noisy, this could upset the candidates and affect them psychologically. This will, in turn, affect their ability to perform on the test. Furthermore, if the pre-test preparation work is not carried out according to the specifications prescribed in the test plan (such as use of faulty or unsuitable furniture), it may affect the comfort level of the candidates and, consequently, the quality and effectiveness of the test. An uncondusive environment could affect the reliability of the test scores as well (refer to para 13 in Chapter 3).

Rules and Regulations for conduct of Assessment

6. It is common for a training organisation to draw up Rules and Regulations (R&R) pertaining to the conduct of assessment such as assessment administration processes and procedures. Separate R&R could be formulated for different purposes, say, one set for candidates; others for invigilators; assessors and moderators, where applicable.
7. For test candidates, the rules generally cover aspects such as:
 - Attendance requirement for eligibility to sit for test;
 - Rules pertaining to test reporting time, lateness, absence, plagiarism and cheating, etc;
 - Conditions for deferment of assessment and re-test;
 - Pass mark, grading system and award criteria; and
 - Eligibility for review of results and appeal process.

Roles and Responsibilities of Assessors

8. For assessors, the R&R are usually prescribed in the form of roles and responsibilities. They are formulated to clearly stipulate the functions of an assessor who is appointed to conduct theory assessment of learners. Some of these functions include:

Responsibilities

- Ensure that the assessment is conducted in a fair and objective manner;
- Ensure that the assessment is scored based on evidence sighted or verified;

Roles

- Liaise with supervisor to prepare venue and physical requirements for the assessment;
- Check on venue suitability and readiness prior to commencement of assessment;

- Conduct assessment and score candidate based on prescribed marking scheme;
- Observe a courteous behaviour towards the candidates throughout the assessment;
- Attend to candidates' well-being and welfare as and when needed during assessment (eg, need for attention arising from sickness or anxiety);
- Make submission of candidates' assessment scores and grade to the proper authority in the organisation; and
- Provide feedback to candidates on assessment performance (especially for formative assessment).

Code of Practice for Assessors

- Besides roles and responsibilities, some organisations also formulate a code of practice (or code of conduct) for their assessors, especially for organisations that have a large pool of assessors. The intended purpose of the code is to guide the assessors to conduct themselves in an appropriate manner that is consistent with the expectations, values and culture of the respective organisations. Some examples of statements written for the code of practice are as follows:
 - An assessor shall have the skills and knowledge to carry out each assessment assignment;
 - An assessor shall, at all times, displays exemplary conduct and behaviour consistent with his role;
 - An assessor shall display due diligence and objectivity in carrying out his assessment duties; and
 - An assessor shall maintain confidentiality regarding assessment decisions/outcomes and records of assessment outcomes.

Matching Evidence Collected to Assessment Criteria

9. Although theory assessments may take different modes (eg, pen & paper and oral assessment), the key point to note is that these assessments (whether formative or summative) are meant to test knowledge acquisition of learners. Hence, in a theory assessment, the assessor should focus on the evidence of knowledge acquisition given by test candidates. In the case of a pen & paper test, the evidence will come in the form of written answers from the candidates.
10. To assess knowledge acquisition objectively, it is important for an assessor to evaluate candidates based on the evidence collected and match it to the

assessment criteria. This will entail checking the answers provided by the candidates against the model or suggested answers from a prescribed marking scheme, which is normally drawn up based on the assessment criteria. A well-prepared marking scheme would include adequate instructions to guide an assessor to award marks to the answers. The guidelines would normally include the provision for award of marks to totally correct as well as partially correct answers.

Outcome of Assessment

11. The outcome of an assessment, based on the results reflected in the marking scheme, can determine whether a candidate makes the grade or otherwise. Besides looking at the overall score and final percentage of marks, the entries in the marking scheme can be used to identify the strengths and weaknesses of the candidate. While the strengths speak well of the candidate, the weaknesses will help to identify the gaps that the candidate might wish to address.
12. With analysis, the gaps can be the result of one or a combination of skills incompetence, knowledge and attitude (or SKA). If the candidate concerned is given the appropriate feedback on his gaps, he will know the areas for improvements and be able to take steps to address them.

Appropriate Forms Are Used to Collect Evidence

13. With advancements in technology, it is now possible to conduct a theory test not only by the traditional pen-and-paper method but also online.
14. Printed assessment papers and all online tests should be according to the format as prescribed by the organisation. For example, a standardised format could be used for all the summative assessments. An assessment template could be prepared for use by setters of test papers or online tests.
15. For an online test, the format should be easy to use and the evidence collected easily processed for analysis. During a written theory assessment, it is expected of candidates to provide their answers in a written form. The answers can be provided in the following documents:
 - Question paper itself – whereby provision is made (blank space) for candidates to write the answers on;
 - Answer booklet – which contain blank pages with lines for answers to be written on; and
 - Answer sheet – which is mainly used for candidates to provide answers to objective-type questions (eg pre-printed blank answer sheets which can be used by candidates to shade their answers on bullet holes, and

the completed sheets can be marking using a optical mark reader or OMR). The use of OMR answer sheets could eliminate errors in marking.

16. An assessor conducting a theory assessment should always use the appropriate and approved documents to collect evidence (answers) from candidates. This will help to prevent any issue of misunderstanding that may arise from the use on non-approved documents.
17. Other forms for collecting evidence may include:
 - Questionnaire (sometimes referred to as a checklist) – for oral assessment;
 - Assessor/Invigilator's report form – for recording evidence observed during the conduct of the assessment eg, discrepancy, cheating, disruption, etc, which the assessment administrator may wish to take note of; and
 - Appeal form – for candidates who may wish to request the assessment administrator to review the assessment results.

Overcoming Limitations in Evidence Collection

18. In the conduct of an assessment, it is not uncommon for problems to arise no matter how much preparations go into it. One of the problems is the issue relating to the failure to collect adequate evidence by the assessor. This issue has been explained in the earlier paragraphs.
19. There are, nonetheless, other problems or limitations which may be encountered and some of these are:
 - **Faults in the design of the assessment tool** – This refers to errors or omissions in the test papers which could cause confusion and anxiety among the candidates. Also, such faults could affect the performance of the candidates thereby affecting the reliability of the test scores. Hence, it is common for training organisations to require all test papers to be subjected to the process of moderation. And moderation is normally done by a relevant content expert or Subject Matter Expert (SME) who will check the contents of the test paper against the test plan and table of specifications. This is done to ensure that the test paper meets the test requirements before use.
 - **Assessment criteria not explicit enough** – This refers to candidates not being informed clearly of the criteria used in the theory assessment ie, what the candidates will need to do in terms of the questions they need to attempt, marks allocation to each questions, pass mark and

grading system. If candidates are not absolutely certain of the assessment criteria, they could make the wrong decisions in terms of selecting the sections and questions to attempt. Consequently, this could affect their grades. To prevent this from happening, the assessment paper must be thoroughly checked to ensure that the assessment criteria are clearly spelt out. Additionally, the moderation of assessment papers could be incorporated into the assessment system.

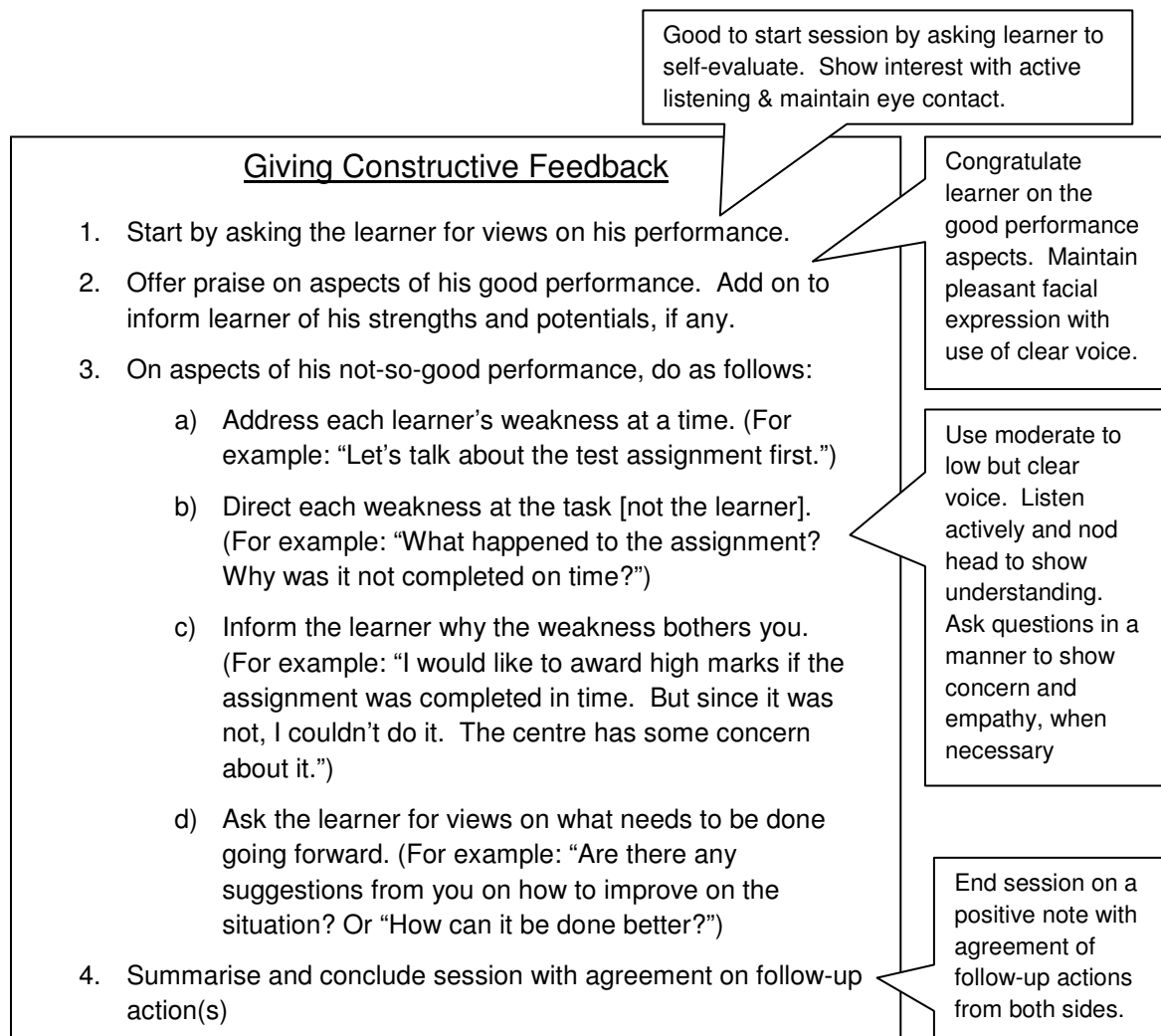
- **Unclear instructions in assessment papers** – For a theory test paper, instructions will include both general and specific instructions given to candidates on the paper itself. If instructions are not clear, it could lead to misinterpretation by the candidates during the test and this could affect the way they approach the test. Hence, setters will need to ensure that all instructions on test papers are clear, specific and easy to understand. Short sentences are preferred over long sentences.
- **Insufficient time for assessment** – This is a common grouse of any candidate sitting for a theory test. Insufficient time could lead to candidates rushing to complete answering their questions and, therefore, take short-cuts or provide answers in a hurry. This could then lead to the assessor not being able to collect evidence that accurately reflects the knowledge acquired by the candidates. Hence, the setter must ensure that the candidates will have enough time to complete the test. The conduct of a pilot run could provide useful feedback to the setter to make necessary changes to the questions construct or duration of the theory test paper. In general, if 90% of the pilot run candidates are able to complete the test within the time set, there may not be a need to make changes to the draft test paper.
- **Assessment activities not integrated** – Where candidates are required to attempt more than one test paper for a particular module and the papers are to be taken in sequence, these will need to be well integrated and co-ordinated to allow candidates to attempt the test papers in a progressive manner. If the flow of the activities is disjointed, candidates may encounter difficulties in performance. For example, the summative assessment for a module may comprise 2 written papers (ie, Paper 1 and 2) and an oral test. If the 2 papers and oral test are scheduled far apart and on different days, the prolonged testing could pose problems not only to the candidates (who have to prepare for the module assessment 3 times) but also to the assessor as evidence needs to be separately collected, leaving room for error. On the other hand, if the activities are well integrated and co-ordinated (by the assessment administrator) such that the assessment activities are arranged with a logical flow, it will certainly enable the candidates'

continuity in test performance, where the knowledge learnt is not forgotten and has to be re-learnt. This also allows the assessor to collect evidence in a smooth and logical manner.

- **Model answers are limited** – Preparing model answers to essay or short-answer items is an excellent way to check the quality of the items. It is also important that where these are provided, they should be complete (not partial). If the model answers are limited, the test is flawed or unfair. The allocation of marks for each answer will then be left to the discretion of the marker, who will be required to make use of his content knowledge to judge the answers and allocate marks. This may not be desirable as the test scores could become unreliable because of the lack of objectivity. For short-answer and essay questions, the setters are expected to provide instructions on allocation of marks for partial answers as well.

Provision of Constructive Feedback to Learners

20. All learners need feedback. This is one form of learning. The provision of feedback will enable the learner not only to know the positive outcome of learning but also the areas he could improve upon. So, in a performance feedback session, the learner should be told of his strengths as well as his weaknesses.
21. When providing feedback to the learner, whether done orally or through a written form, it is important to take note of the way the message is conveyed. If the feedback is provided orally, the tone of voice, pace, eye contact and active listening, etc should be monitored. For written feedback, it is important to pay attention to the choice and phrasing of words in order to convey the message in a constructive manner.
22. The feedback session should normally conclude with some agreement made between the learner and the feedback provider, be it his trainer, coach or assessor. Generally, the agreement is in the form of possible follow-up actions to be taken by both the learner and feedback provider.
23. When conveying feedback on the learner's weaknesses, the trainer should provide it constructively. One suggested approach in conducting a constructive feedback session is as follows:



Improvement to Assessment Process

24. As with the implementation of any process or system, review is a necessary activity for the purpose of making continual improvement. The assessment process should not be any exception.
25. Inputs to the review process could be in the form of end-of-module or course evaluation by learners. In addition, learner satisfaction surveys and staff satisfaction surveys could be conducted on an annual basis. The module evaluations and surveys could be analysed and the outcomes could be used to future improvement to the assessment process.

CHAPTER 12: ANALYSIS OF TEST ITEMS

Introduction

- 1 The aim of Item Analysis is to evaluate the effectiveness of the test item. This is done by examining the responses of trainees to the test item. Item Analysis is usually done on multiple choice questions (MCQs). The item analysis is to find out the:
 - 1.1 Difficulty Index (DI) of the MCQ item; and
 - 1.2 Effectiveness of each option in the item.
- 2 The conduct of Item analysis will enable the trainer or organisation to know whether the item is too easy or too difficult (DI), how well it is able to discriminate between the high and low scorers who have sat the test (DP) and whether all the options are able to functioned as intended.

Procedure

- 3 In carrying out item analysis, first is to arrange the test papers in ranking order based on the scores ie from the highest to the lowest scores of the entire group of candidates. If there are 30 candidates, for example, there should be 32 test papers with scores. Select 1/3 of the papers with the highest scores (ie 10 papers) and label them as the UPPER GROUP. Next, select 1/3 of the papers with the lowest scores (another 10 papers) and label them as the LOWER GROUP. (Set aside the middle group of papers.)
- 4 Look at the MCQ responses from the candidates of the 2 groups item by item. For the UPPER GROUP, record the responses for each MCQ as shown in the table below. Do the same for the lower group.

Item Analysis Table					
MCQ Test Item No 5	No of candidates	A	B*	C	D
UPPER GROUP (UG)	10	1	6	3	0
LOWER GROUP (LG)	10	6	2	2	0

*Key or correct answer

Difficulty Index (DI)

- 5 The **Difficulty Index (DI)** of an item is defined as the number of candidates who have answered the question correctly divided by the total number of trainees taking the test. The result should be indicated as a percentage. **The interpretation of the percentage is that the smaller the percentage, the more difficult is the item.** If the result is 100%, it means that everyone has picked the correct answer. In this case, the item is deemed to be very easy. Zero percent (0%) means that the item is very difficult as none has obtained the right answer.

- 6 It is impractical to extract raw data from all the candidates' test papers manually to compute the DI if the number is very large. Hence, the practice in such a situation is to use the data from the top 1/3 of the group (Upper Group) and the bottom 1/3 group (Lower Group).

- 7 The difficulty Index for item 5 is obtained using the formula, as follows:

$$\begin{aligned}\text{Difficulty Index (DI) in percentage (\%)} &= \frac{\text{Number in UG and LG with correct answer (B)}}{\text{Total number of candidates in UG and LG}} \\ &= (6 + 2) \div (10 + 10) \\ &= 8 \div 20 \\ &= 0.40 \text{ or } \underline{\underline{40\%}}\end{aligned}$$

- 8 As shown, the DI for item 5 is 40% and this refers to the percentage of candidates that have answered the item correctly. Since 40% lies in the middle of the band of 0% to 100%, it could mean that the correct answer (ie the key B) is neither too easy nor too difficult to the candidates. In other words, the key to this item might not need to be changed or replaced.

Effectiveness of Distracters

- 9 The **effectiveness of the distracters** is determined by comparing the number of candidates in the upper and lower groups who have selected each incorrect option. A good distracter should attract more trainees from the lower group.
- 10 For item 5, distracter A is functioning well while C is not. D is completely ineffective. If any action were to be taken, it would be good to retain A and revise C. As for D, it should be replaced with a completely new distracter.
- 11 The practice of item analysis is important when maintaining a question item bank. Only valid, reliable and effective items should be included in the bank.

Follow-up Actions based on Item Analysis results

- 12 Recommended follow-up action(s) must be appropriate and consistent with the results of the item analysis. Measures must be in place to check both quantitatively and qualitatively that the candidate's scores are accurate and reasonably free from measurement error.
- 13 Statistical methods and legitimate data interpretations would serve to investigate suitable inferences to appropriate recommendations to be made.

14 Recommendations may include:

- Improving the effectiveness of distracters or the frequency of such distracters;
- Attractiveness of the distracters; and
- Changing the difficulty of the assessments, etc.

CHAPTER 13: DESIGN PRACTICAL ASSESSMENT SPECIFICATIONS

Introduction

- 1 One of the key aspects of the practical test plan is the preparation of design specifications for the intended test (as mentioned in Chapter 4).
- 2 The design specifications are to be prepared by the test plan developer, who will determine the appropriate test to suit the nature and contents of the practical training.
- 3 The components of practical assessment specifications should include:
 - Assessment methods;
 - Competencies to be tested;
 - Tools & equipment;
 - Organisational & industry requirements;
 - Task assessment checklist;
 - Duration; and
 - Venue.

Table of Specifications

- 4 The table of specifications will comprise the following aspects:

- a) Competency to be tested (task assignment)

The description of a task assignment (or competency) should start with an active verb (or action word). Some examples of a task description are:

- Dress a patient's wound;
- Service a car engine;
- Machine a steel shaft; and
- Assemble a desktop computer.

- b) Scope of Test

The scope of test defines the range of competencies to be assessed in the task assignment to be performed by the candidate. The competencies are found in the skills standard of the relevant module.

Reference should be made to the skills standard of the module to identify the competencies to be tested. The test plan developer should make a list of the competencies and include them under the scope of test. It is important that only those competencies that need to be tested are selected. Each competency should start

with an active verb. Safety aspects to be observed during the test performance by the candidate should be included, if required.

c) Design Features

Design features provide the scenario under which the task assignment is to be performed or describe the actions needed to be carried out by the candidate for the assessment.

The design features will, in effect, guide the setter in setting the practical test paper. It is necessary for the features to be stated as clearly as possible to enable the setter to provide clear instructions to the candidate. Reference should be made to the performance criteria of those competencies selected for inclusion in the task assignment.

d) Duration of Test

The duration of test refers to the amount allocated for the test. It is important that the duration allowed (eg 30 minutes, 4 hours, etc) should be adequate to enable the candidate to complete the task or assignment.

- 5 Two (2) examples of the Table of Specifications are given below ie one for a manufacturing trade and the other for a service trade.

Example of Design Specifications for a Manufacturing Trade

Task Assignment	Scope of Test	Design Features
Machine a steel component on a CNC milling machine to specifications of a given drawing.	▪ Operate CNC Milling Machine ▪ Perform CNC programming ▪ Perform setting-up of work piece ▪ Perform setting-up of cutting tools ▪ Perform flat milling ▪ Perform angular milling ▪ Perform drilling ▪ Perform precision measurements ▪ Observe safety precautions	Candidate to perform machining operation in accordance with blueprint provided by setter. Candidate is to: ▪ Use CNC milling machine ▪ Use mild steel of raw size 50x75x100 mm ▪ Perform programming on CNC controller ▪ Use standard cutting tools ▪ Machine slider with 45⁰ ways ▪ Drill of 12 mm holes
Duration of Test: 3 hours		

Example of Design Specifications for a Service Trade

Task Assignment	Scope of Test	Design Features
Record the height and weight of a patient	▪ Prepare patient for a service task ▪ Prepare equipment ▪ Take height measurement ▪ Take weight measurement ▪ Recording measurements ▪ Perform customer service	Candidate to perform a service task in accordance with scenario provided. Candidate is to: ▪ Prepare patient who is an elderly person, who comes in a wheelchair but is able to stand when assisted. ▪ Take height/weight measurements of elderly person on a machine with digital readout. ▪ Perform the service task with appropriate service techniques ▪ Record the measurements taken in the appropriate record sheet provided.
Duration of Test: 15 minutes		

Scheduling the Assessment

6 It is important to ensure the availability of the following aspects before the practical assessment is scheduled:

- Candidates;
- Trainers and assessors; and
- Resources required as stipulated in the assessment plan.

It is also important to note the time required to complete the training so that the assessment can be scheduled at an appropriated juncture.

Marking Scheme

- 7 An example of a marking scheme (incorporating the rubric) for a practical test is given below:

Obtained from the
Performance Criteria
in the skills standard

Raw Score (R) refers to a range of initial marks assigned to each item being measured. The assignment of marks, whether [0 or 4] or [0, 1, 2, 3, 4], depends on the requirement as specified in the rubric. Zero (0) mark is for non-performance of item by candidate or if item performed does not meet specification.

$(\text{max R}) \times (\text{max C}) \times (\text{max D})$
= Max score possible

MARKING SCHEME

Performance Criteria	Items to be Measured	Raw Score (R)				Score	
		Award '0' score for non-performance of item or if item performed does not meet specification; otherwise, award score accordingly.				Actual R x C x D	Max Marks Possible
		1 Unacceptable	2 Satisfactory	3 Good	4 Excellent		
1 Safety precautions are observed when handling chemicals	1.1 Use of Safety attire	NA	NA	NA	4	$4 \times 3 \times 2$ = 24	$4 \times 3 \times 2$ = 24
	1.2 Reading of safety labels	1	2	3	4	$3 \times 2 \times 1$ = 6	$4 \times 2 \times 1$ = 8
	1.3 Application of chemical-handling procedure	1	2	3	4	$2 \times 2 \times 1$ = 4	$4 \times 2 \times 1$ = 8
	1.4 Washing of hands	NA	NA	NA	4	$3 \times 3 \times 1$ = 9	$4 \times 3 \times 1$ = 12
Score (Percentage)						43 (83%)	52 (100%)

Actual score is derived from raw score (assigned by assessor) multiplied by the Criticality weighting factor multiply by the Difficulty weighting factor. For example, the actual score for Item 1.4 is 9 derived from $3 \times 3 \times 1$. In other words, the candidate has scored 9 marks out of 12 marks.

- 8 As shown in the marking scheme (see above), the performance criteria and items to be measured are obtainable from the skills standards of the module or course. There are 4 items to be measured under Performance Criterion 1 (Safety precautions are observed when handling chemicals). Note that statements under 'Performance Criteria' are written in the passive voice.

Raw Score

- 9 The **Raw Score** is the range of marks assigned to each item to be measured by the assessor. Marks can be assigned based on 0 or 4 marks (system A) or a range of marks from 0, 1, 2, 3 or 4 marks (system B). Under system A, the

item when measured can only be given a mark of either 0 or 4. In other words, the measurement is either 4 (excellent performance or within specification) or 0 (unable to perform or measurement is outside specification) and nothing in between.

- 10 Under system B, the item to be measured can be given marks ranging from 0 (non-performance), 1 (unacceptable performance), 2 (satisfactory or need improvement), 3 (good or acceptable) or 4 (excellent) depending on the performance level of the candidate. An assessor normally circles the number which represents the raw score given.

Weighting Factor

- 11 A task normally comprises a number of skill elements which can vary in terms of importance (criticality) and difficulty levels of performance. Weighting factors for criticality and difficulty are features used to distinguish the more capable and less capable performers. For example, the item **“the emergency door of the bus is opened using the correct approach”** is a very critical task but the difficulty level is not high. Hence, a higher weighting factor (say, 3) is assigned for criticality and a lower factor for difficulty (say, 1).
- 12 The factors generally used for the assignment of weighting are as follows:

Weighting Type	Factor		
	1	2	3
Criticality	Least Critical	Critical	Most Critical
Difficulty	Least Difficult	Difficult	Most Difficult

- 13 An example of a practical test plan incorporating the table of specifications and marking scheme is given at Annex A.

Annex A

TEST PLAN (PRACTICAL)

(An example for illustration purpose)

Certification Title: Certificate in Health Care

Module: Basic Health Care Practices

Date of Approval:

No of Pages: 2

THIS TEST PLAN MUST BE USED IN CONJUNCTION WITH THE SKILLS STANDARD

Scope of Test

- 1 The practical test consists of One assignment
- 2 The specification of the test assignment is set out in the table below:

Task Assignment	Scope of Test	Design Features
Record the height and weight of a patient	▪ Prepare patient for a service task ▪ Prepare equipment ▪ Take height measurement ▪ Take weight measurement ▪ Recording measurements ▪ Perform customer service	Candidate to perform a service task in accordance with scenario provided. Candidate is to: ▪ Prepare patient who is an elderly person, who comes in a wheelchair but is able to stand when assisted. ▪ Take height/weight measurements of elderly person on a machine with digital readout. ▪ Perform the service task with appropriate service techniques ▪ Record the measurements taken in the appropriate record sheet provided.
Duration of Test: 15 minutes		

Marking Scheme (incorporating the rubric)

- 3 Items to be assessed are based on the performance criteria prescribed in the skills standard.
- 4 Use raw scores of 0, 4 and 0, 1, 2, 3, 4, where appropriate, when a rubric is incorporated into the marking scheme.

0, 4	0, 1, 2, 3, 4
0 Outside specification or unable to perform	0 Non-performance or unable to perform
	1 Unacceptable performance
	2 Satisfactory performance
4 Within specification or excellent performance	3 Good or Acceptable performance
	4 Excellent performance

- 5 Weighting factor based on Criticality and Difficulty of task are to be used, as follows:

5.1 Criticality: 1 = least critical; 2 = critical; and 3 = very critical

5.2 Difficulty: 1= least difficult; 2 = difficult; and 3 = very difficult

6 Part of a marking scheme (incorporating the rubric) is shown below as an example:

Performance Criteria	Items to be Measured	Raw Score (R)				Score	
		Award '0' score for non-performance of item or if item performed does not meet specification; otherwise, award score accordingly.					
		1 Unacceptable	2 Satisfactory	3 Good	4 Excellent	Actual R x C x D	Max Marks Possible
2 Height is measured with patient standing up straight without footwear	2.1 Providing assistance to patent to stand up to take measurement	1	2	3	4	__x3x1 = __	4x3x1 = 12
	2.2 Getting patient to take weight measurement bare-footed	NA	NA	NA	4	__x3x3 = __	4x3x3 = 24
	2.3 Getting patient to stand in an up-right position during weight measurement	1	2	3	4	__x3x3 = __	4x3x3 = 36

Pass Mark and Grading System

7 The candidate must score at least 60% marks to pass the test. The grading system is as follows :

Score	Grade
80% and above	Distinction
60% to 79%	Pass
59% and below	Fail

CHAPTER 14: DAVE'S PSYCHOMOTOR CLASSIFICATION

Introduction

- 1 Besides Benjamin Bloom, there are other education psychologists who have done research in the field of psychomotor learning as well. They include R H Dave (1967), EJ Simpson (1972), AJ Harrow (1972), and A Romiszowski. Of all the domains expounded, R H Dave's psychomotor domain is the simplest to understand and one that is easy to apply.
- 2 Dave classifies psychomotor learning into 5 categories. He regards each category as a stage and in each stage a particular behaviour is displayed by the learner. In acquiring practical skills, the learner starts from a stage at the low level and moves up to the next higher level and so on. In the process, he acquires skills in a progressive way until he reaches the highest level attainable.

Five Categories of Dave's Psychomotor Domain

- 3 The 5 categories of Dave's psychomotor domain are:

Stage No	Category	Description of Behaviour or Ability	Action of Learner	Appropriate Active Verbs to describe learning
1	Imitation	Able to copy the action of someone (eg trainer, instructor)	Learn from observation; imitate action or activity	Adhere, copy, follow, identify, imitate, mimic, re-enact, observe, repeat.
2	Manipulation	Able to perform skill from instructions given	Perform task from verbal or written instructions	Build, execute, implement, perform, re-create.
3	Develop Precision	Able to execute skill independently	Perform a task or skill activity without any assistance.	Complete, control, calibrate, show, demonstrate,
4	Articulation	Able to perform with adaption and integration of skills to satisfy stated objectives	Perform activities with use of new methods to meet requirements. Show ability to adapt and innovate.	Adapt, construct, co-ordinate, solve, combine, integrate, develop, modify, improve, teach.
5	Naturalisation	Able to perform at a high level with unconscious mastery of skills.	Perform to meet strategic needs. Show ability to define aim, approach and strategy of activities	Design, invent, manage, project-manage

- 4 What can be learned from Dave's psychomotor domain is that the learning of practical skills essentially covers 3 phases of learning – the first phase being learning by imitation; the second performing after much practice; and the third developing a habit for proficient and consistent performance. At the third phase, the person should display unconscious competence and be able to create new ways to improve on his performance and to teach others.
- 5 The psychomotor domain is originally meant to apply to the learning of skills related to physical movements. However, in the present-day context, it can also apply to skills such as telephone and public-speaking skills as well as to computer operations, (eg data entry and keyboard skills).
- 6 In teaching and learning of practical skills, the ability to perform is proof of the acquisition of skills. In other words, a learner can only be said to have acquired the skills if he is able to display competence of performance without assistance – reach the 'develop precision' stage of Dave's psychomotor domain. Therefore, in the assessment practical skills achievement, it should be aimed at assessing the learner's ability to attain at least stage 3 of Dave's psychomotor domain.
- 7 Some examples of learning objectives under the psychomotor domain are given at Annex A.

Psychomotor Domain – Examples of Learning Objectives

A Imitation (able to copy an action)

- 1 Observe and perform the machining task as demonstrated
- 2 Imitate the hand and body movements of the clown.

B Manipulation (able to perform with given instructions)

- 3 Prepare the dole for making cake according to the procedure given.
- 4 Carry out the activity as instructed in the action plan.

C Develop Precision (able to perform independently)

- 5 Machine the part to tolerance of $\pm 0.005\text{mm}$ accuracy for all dimensions as given in the blue-print. (No procedure and other information will be given.)
- 6 Develop and demonstrate the procedure to perform good customer service.

D Articulation (able to perform with adaption and integration)

- 7 Develop and implement a new manufacturing procedure to increase production rate by 10%.

E Naturalisation (able to show mastery of skills)

- 8 Invent a prototype equipment to reduce heat output to the generator.

CHAPTER 15: TYPES OF ASSESSMENT CRITERIA AND TOOLS

Introduction

1 This chapter covers topics on the following aspects:

- Types of assessment criteria
- Assessment tools for assessing learners
- Assessment of candidates with special needs

Types of Assessments Criteria

2 In the acquisition of practical skills, a learner will invariably be taught the procedures to perform various tasks or activities. The performance of each task will have an outcome, but it may or may not culminate into a tangible product.

3 Thus, in assessing practical skills, the procedure used by the candidate to perform the task will form a part of the assessment. This part is called process assessment. Where the outcome of the task performance by the candidate leads to the completion of a product (ie something made or produced as a result of the process that can be graded or measured), assessment will include grading the product and this part of the assessment is known as product assessment.

4 Assessment of practical skills may include both process and product assessment. Other aspects to be considered as part of assessment are:

- Need to assess candidate's observance and compliance to safety requirements and/or hygiene aspects.
- Time allocation for assessment

Process Assessment

5 In process assessment, an assessor will need to present to observe the way the candidate carries out his task assignment. He will normally have with him a marking scheme or checklist (normally prepared by the setter of the practical test) which will direct him to observe specific aspects of the work procedure used by the candidate. Examples of specific aspects are:

- The candidate has greeted the customer with a smile.
- The candidate has used an appropriate tool to extract the bearing.
- The candidate has applied grease to the moving parts before replacing them.

These specific aspects should conform to the performance criteria in the skills standard of the module or course.

Product Assessment

- 6 In **product assessment**, the performance is scored on the product or project completed by the candidate. Similarly, the assessor would grade the product based on the assessment criteria specified in the marking scheme or checklist. Where necessary, the assessor may use appropriate precision measuring tools and instruments in grading the product. Some examples of assessment criteria are:
- The dimensions of the work piece are machined to within tolerance specified in the blueprint.
 - The electrical unit is assembled and wired up to meet functional requirements.
- 7 As in process assessment, product assessment criteria are also obtainable from the skills standard of the module or course.

Safety and Hygiene Factors

- 8 As work safety is undoubtedly a primary concern to employers and workers alike in any work environment, the inculcation of safe work habits is expected to be featured in any practical skills training programme. Hence, the observance of work safety by a candidate is increasingly becoming a key requirement in skills assessment. On the same basis, it applies to work hygiene as well especially for those who learn and work in a health care environment.

Time allocation for Assessment

- 9 Assessment of practical skills will need time not only to organise but also to allow the candidate time to prepare for the assessment. There are other aspects such as availability of facilities and equipment, acquisition of materials, appointment of assessors and so on. Whether it is, the guidance principle is to allow adequate time for both candidate and assessor to prepare for the assessment. On top of that, it is critical for the test plan developer to allocate adequate time duration for the assessment. For example, if it is stated that the time duration allocated for the assessment is 3 hours, it is essential for the **setter** to ensure that at least 90% of the candidates is able to complete the task within that time.

Assessment Tools to Assess Learners

- 10 Assessors have a variety of tools (or methods) that they can use to assess learners. They are:
- Practical Test papers – commonly used for formal assessment of learner
 - Worksheets – these are work assignments given to the learner in the course of his learning. The worksheets specifies the work or assignments to carried out and the assessment is based the outcomes of the completed assignments.

- Observations guides – these are guides used by an assessor to observe learner's practical work progress over a period of time. The guides contain specific criteria or pointers for the assessor to look out for. The assessor records his inputs his observation findings and remarks into the guides following each observation made.
- Third-party report – this is a report (on a standard format) from a person who is not the assessor. He could be the learner's supervisor or foreman who is assigned by the organisation to oversee the work performance and progress of the learner. This report is normally reviewed by the assessor before he finalises his assessment of the learner.

Assessment of Candidates with Special Needs

11 For learners with special needs, it may necessary to customise the assessments to enable them to participate in the assessments without the need to sacrifice standards. Some of aspects that a test plan developer, setter and assessor should take into account in assessing candidates with special needs could be:

- The type of disability or impairment – The way to test learners with hearing impairment would be different from that for those with sight impairment. Similarly, physical impairment may demand an entirely different approach in assessing skills competence. Considerations could be given to re-design training and assessments to suit such learners. In any case, the setting of practical tests for learners with disabilities will be best done by persons who are specially trained to provided training and conduct assessment to such learners.
- The Age and Gender of Candidates – Because of the need for life-long learning, learners could belong to different age groups and gender. In other words, learners can be young and old, and made up of males and females. Training programmes and duration that are suitable for the young may not suit matured learners. Similarly, programmes that male participants find suitable may not meet with the same agreement from female participants. Hence, in assessing mature learners, a different approach may be necessary eg having many bite-sized tests rather than a one-off test or assessing only the key or major skill elements rather than all the elements. Other considerations could include test breaks and provision of refreshment.
- Religious & Cultural Background of Candidates – As learners may come from many races and nationalities, they may also embrace different religions and cultures. The conduct of practical skills assessment may need to factor in the religious and cultural needs of the candidates. For example, the conduct of assessment should be avoided during noon on Fridays so as not to coincide with the Friday prayers of males Muslims. It is generally acknowledged that

communication by assessors to candidates should avoid all forms of religious connotations.

- Medical Condition of Candidates – Candidates may or may not appear for practical skills assessments in the best of health. And candidates may show signs of fatigue and nerves during a test. When a candidate is observed to be unwell and unlikely to go on with the assessment, it is expected of the assessor to allow him to discontinue with the test and assist him to seek medical attention. If necessary, arrangement can be made to re-schedule his assessment in view of his medical condition.

CHAPTER 16: CONSTRUCT A PRACTICAL TEST PAPER

Introduction

- 1 Based on the practical test plan provided, an appointed setter should be able to proceed to construct the practical test paper. Much of the information needed for setting the test should be available in the test plan. What is needed for the setter to do is set the practical test according to the table of specifications given in the test plan and put them in the form of a test paper. In addition, the setter will need to draw up an appropriate marking scheme (or checklist) to enable the assessor to mark the test.
- 2 The test plan (prepared by the test plan developer) should contain the following information:
 - Course title of test.
 - Design specifications for the test which should include:
 - Description of test task (eg machine a metal component)
 - Scope of test (specifies expected competencies required to be performed)
 - Design features (provides information and parameter to setter to design the test)
 - Test duration (states the time allocated for test)
 - Tools, equipment and materials needed
 - Venue for the test
 - Marking scheme (specifies the type to be used by marker for assessment)
 - Pass mark for test and grading system
- 3 With the information provided in the test plan, a setter should be able to construct the practical test paper. Two (2) examples ie one for a **manufacturing trade** and the other for a **service trade**, given at **Annex A and B**, provide information on the expected requirements of a well-constructed practical test paper.

Construction of Practical Test Paper

- 4 The test paper should comprise the cover page and the test question paper itself. The cover page provides the general information and instructions to the trainee taking the test such as the certification title, the module, duration of test and general instructions to candidate taking the test.
- 5 The contents of the question paper should comprise the following specifications:
 - **Test objective.** State the objective using an active verb eg 'Measure and record the height and weight of a patient'. Make sure it matches with the specification in the test plan.

- **Duration of the test.** Adhere to the test duration stated in the test plan. Design the test such that the candidate has a good chance to complete it within the time allowed. Where possible, carry out a test or pilot run to verify it.
- **Tools. Equipment and Materials.** List the major items that will be provided to the candidate taking the test. Need to provide accurate information so that the right items can be procured for the test. For example, a weighing machine can be a simple mechanical type or one that is with digital readouts.
- **Instructions.** Any instructions given to the candidate should be spelt out clearly and in a concise manner. Use words like 'follow the instructions given in the test in carrying out your task'; 'observe safety precautions'; and 'stamp your identification number on your work pieces'. Do advise the candidate to read the instructions and study the drawings before commencing work on the task. Any special precautions needed, especially those relating to safety, must be clearly stated.
- **Description of Assignment.** Develop the assignment based within the scope of test and the design features given in the test plan. Make sure it can be carried out with available resources within the stipulated test duration. For the performance of a service task, it may be necessary for the setter to describe in the test paper a scenario under which the test task is to be performed. In the case of constructing or making an object, it is best to provide a technical drawing or blueprint. Examples below show 2 assignment formats.

Example of Assignment with Scenario

Assignment

An elderly patient is wheeled to the clinic. She is able to stand when assisted. You are required to take his height and weight measurements. You are to perform the following task elements:

1. Identify the patient against record
2. Explain purpose and procedure to patient
3. Prepare the measuring equipment
4. Assist patient to stand on machine
- 5.
- 6.
- 7.

Example of Assignment without Scenario

Assignment

1. With the power supply off, construct the circuit shown in Figure 1.
2. Turn on the power supply to +15VCD.
3. Test and troubleshoot the circuit.
- 4.
- 5.
- 6.

- **Assessment Criteria.** The candidate will need to know how he is to be assessed for the test. Hence, the assessment criteria will need to be clearly spelt out so that he is aware of the expected performance requirements (see example below for some of the criteria specified).

Assessment Criteria

You will be assessed on:

- Correct work procedure
- Correct use of equipment and tools
- Observance of safety precautions
-
-
-

Marking Scheme

- 6 A marking scheme will need to be drawn up following the setting of the test paper. The marking scheme format is specified in the test plan. Follow it. The purpose of the raw scores of (0, 4) and (0, 1, 2, 3, 4), and criticality and difficulty weightings have already been discussed in Chapter 12. Devise the marking scheme format and input the performance criteria and items to be measured into the respective columns of the scheme. Thereafter, assign the appropriate raw scores, criticality weighting and difficulty weighting to each item to be measured.
- 7 Take note of the following points when developing an objective marking scheme:
 - Make reference to the skills standard to identify the performance criteria and items to be measured.
 - Measure only the critical items and avoid the trivial elements. (However, too few items may affect the reliability of the eventual test scores.)
 - Specify the items clearly; avoid ambiguous statements;
 - Write the items in the passive voice eg 'All components are connected correctly'.
 - Arrange the items in a logical manner to facilitate ease of scoring by the marker.
 - The items should include both process and product items, where appropriate.
 - Assign appropriate raw scores, criticality weighting and difficulty weighting for each item.
 - Sum the possible scores of all the items. The final possible score will equal to 100% of marks.

- 8 A marking scheme that is well-structured and objectively prepared will allow the assessor (marker) to understand clearly the scoring requirements and award scores to the candidate strictly based on:

- The items to be measured as specified in the marking scheme (and nothing else).
- The raw scores as indicated ie select 0 or 4 for raw score of [0, 4] or select 0, 1, 2, 3, or 4 for raw score of [0, 1, 2, 3, 4]. The interpretation of the raw scores is as follows:

Raw Score Interpretation

Raw Score of 0, 4		Raw Score of 0, 1, 2, 3, 4	
0	Unable to perform (outside specification)	0	Unable to perform or non-performance
		1	Unsatisfactory performance
		2	Satisfactory performance
		3	Good performance
4	Excellent performance (within specification)	4	Excellent performance

- 9 With an objective marking scheme, scoring consistency can be maintained even if different markers are engaged to carry out the scoring for different candidates and at different occasions. In this way, the scores would be more reliable because marking is objectively done.

Pilot Testing of Practical Test Paper

- 10 Where feasible, it is always preferred that a practical test paper is pilot tested after it has been set. Suitable learners could be identified and given the test in as close an environment as possible to that of an actual test. If properly conducted, monitored and evaluated, the pilot or test run could yield the following benefits:

- To confirm that the test can be completed within the test duration.
- To ascertain the clarity of the task assignment.
- To establish the adequacy of the tools, equipment and materials.
- To test out the use of the marking scheme.
- To obtain information on the difficulty level of the test.

- 11 The feedback from the pilot run could be used to fine-tune the test paper before finalising it for use. Areas of weakness could be addressed to increase the reliability of the test. Also, other test organizational arrangements could be reviewed and improved upon, if necessary.

Confidentiality of Test Paper

- 12 The developed practical test paper, inclusive of the marking scheme, is to be submitted to the proper authority in the organization for moderation (if necessary) and safe keeping. Maintaining confidentiality of test paper information is a responsibility of any setter. In other words, information of the test items and marking scheme must not leak out to anyone and anytime.

An example for illustration purpose

NRIC No./FIN

--	--	--	--	--	--	--	--	--	--

CERTIFICATE in HEALTH CARE

MODULE: BASIC HEALTH CARE PRACTICES

PRACTICAL TEST

20 MINUTES

JANUARY 2010

INSTRUCTIONS TO CANDIDATES

- 1 Do not open this question paper until you are told to do so by the examiner
- 2 Read the question paper and seek clarifications, if any, with the examiner.
- 3 The Practical Test consists of ONE assignment only
- 4 Check the equipment, tools and materials provided before you start work
- 5 Observe safety and health practices
- 6 Write your NRIC No on work submitted.
- 7 Return this question paper at the end of the test.

This question paper consists of 2 printed pages (including this page)

Practical Test

Test Description: Measure and record height and weight of a patient

Test Duration: 20 minutes

Tools, Equipment & Materials:

- 1 Digital height and weighing machine
- 2 Ruler
- 3 Disposable paper

Instructions:

Read this question paper. Seek clarifications from the Examiner if you have any questions. Do not begin until he tells you to do so. Listen to his instructions. Observe workplace safety and health practices.

Assignment:

- 1 An elderly patient is wheeled to the clinic. He is able to stand when assisted. You are required to measure his height and weight. You should perform the following task elements.
 - 1.1 Identify patient against record.
 - 1.2 Explain purpose and procedure to patient
 - 1.3 Prepare the equipment.
 - 1.4 Assist patient to stand on machine
 - 1.5 Take height and weight measurements
 - 1.6 Record data
 - 1.7 Help patient to wheelchair.
- 2 Submit The height and weight records to the examiner

Assessment Criteria

- 3 You will be assessed on:
 - 3.1 Correct work procedure
 - 3.2 Handling of patient with care
 - 3.3 Correct use of equipment and tools
 - 3.4 Observance of safety
 - 3.4 Accuracy in measurements

*** End ***

MARKING SCHEME
(For use by Marker/Assessor/Examiner only)

Test Assignment: Measure and record height and weight of patient

Test Duration: 20 minutes

Candidate's Index Number _____

Performance Criteria	Items to be Measured	Raw Score (R)				Score	
		Award '0' score for non-performance of item or if item performed does not meet specification; otherwise, award score accordingly.					
		1 Unacceptable	2 Satisfactory	3 Good	4 Excellent	Actual R x C x D	Max Marks Possible
1 The weighing machine is set to zero before patient steps on it to measure	1.1 Setting of Weighing m/c	NA	NA	NA	4	__ x 3 x 2 = ____	4 x 3 x 2 = 24
	1.2 Cleaning of weighing m/c	1	2	3	4	__ x 2 x 1 = ____	4 x 2 x 1 = 8
	1.3 Identifying correct patient for weighing procedure	NA	NA	NA	4	__ x 3 x 1 = ____	4 x 3 x 1 = 8
	1.4 Explaining purpose of weight-taking to patient	1	2	3	4	__ x 3 x 1 = ____	4 x 3 x 1 = 12
2 Height is measured with patient standing up straight without footwear	2.1 Providing assistance to patient to stand up to take measurement	1	2	3	4	__ x 3 x 2 = ____	4 x 3 x 2 = 24
	2.2 Getting patient to take weight measurement bare-footed	NA	NA	NA	4	__ x 3 x 3 = ____	4 x 3 x 3 = 24
	2.3 Getting to stand in up-right position during weight measurement	1	2	3	4	__ x 3 x 3 = ____	4 x 3 x 3 = 36
3 The level on the patient's head is flat before recording the height	3.1 Checking flat level of patient's head	NA	NA	NA	4	__ x 3 x 2 = ____	4 x 3 x 2 = 24
	3.2 Checking patient's comfort level during height measuring procedure	1	2	3	4	__ x 2 x 1 = ____	4 x 2 x 1 = 8
4 The height and weight of patient are recorded accurately	4.1 Recording the height of patient	NA	NA	NA	4	__ x 3 x 1 = ____	4 x 3 x 1 = 12

	4.2 Recording the weight of patient	NA	NA	NA	4	___ x 3 x 1 = ___	4 x 3 x 1 = 12
	4.3 Providing assistance to patient to return to wheelchair	1	2	3	4	___ x 2 x 1 = ___	4 x 2 x 1 = 8
Score (Percentage)						___ ()	200 (100%)

Name and Signature of Marker/Assessor/Examiner _____ Date _____

NRIC No./FIN

--	--	--	--	--	--	--	--	--	--

CERTIFICATE in ELECTRONICS

MODULE: INDUSTRIAL ELECTRONICS

PRACTICAL TEST

ONE HOUR

JUNE 2010

INSTRUCTIONS TO CANDIDATES

- 1 Do not open this question paper until you are told to do so by the examiner
- 2 Read the question paper and seek clarifications, if any, with the examiner.
- 3 Check the tools and materials provided before you start work
- 4 Observe safety and health practices
- 5 Stamp an identification on your workpiece. The examiner will advise you what to do.
- 6 Return this question paper at the end of the test.

This question paper consists of 4 printed pages (including this page)

PRACTICAL TEST

Test Description: Construct a simple 555 timer circuit.

Test Duration: 1 hour

Tools: Equipment and Materials

- 1) Multimeter
- 2) Oscilloscope
- 3) DC Power Supply
- 4) LED X 1
- 5) Resistor (680Ω X 1, $10K\Omega$ X 1, $150K\Omega$ X 1)
- 6) Capacitor ($10nF$ X 1, $1\mu F$ X 1, $47\mu F$ X 1)
- 7) IC555 X 1

Instructions: Read the question paper. The examiner will give a briefing to test candidates on the assignment and assessment criteria. Seek clarifications from him if you have questions on this test. Do not begin until he tells you to do so. Observe safety and health practices.

Assignment

1. With the power supply off, construct the circuit shown in FIG. 1. ($R_1=150K\Omega$, $R_2=10K\Omega$, $C=1\mu F$)
2. Turn on the power supply +15VDC.
3. Test and troubleshoot the circuit.
4. Measure the Vp-p of Vout, pin 2 of IC 555 and record its value in **TABLE 1**.
5. Show your circuit and signal waveforms to the examiner.
6. Turn off the power supply.
7. Submit the answer sheet in page 2 to the examiner.

Assessment Criteria

- 1 You will be assessed on:
 - 1.1 Accuracy in connection of the circuit;
 - 1.2 Component layout and wiring;
 - 1.3 Waveform measurement; and
 - 1.4 Answering of two short questions.

Your Answer Sheet (Please submit this answer sheet to the examiner)

Your NRIC No:

Date:

Waveform Measurement.

Test point	Waveform	Vp-p
Vout		
Pin 2		

TABLE 1

Answer the following 2 questions by filling in the blanks with a word or phrase.

1) The 555 timer circuit is connected up as a/an _____.

2) The frequency may be adjusted by _____.

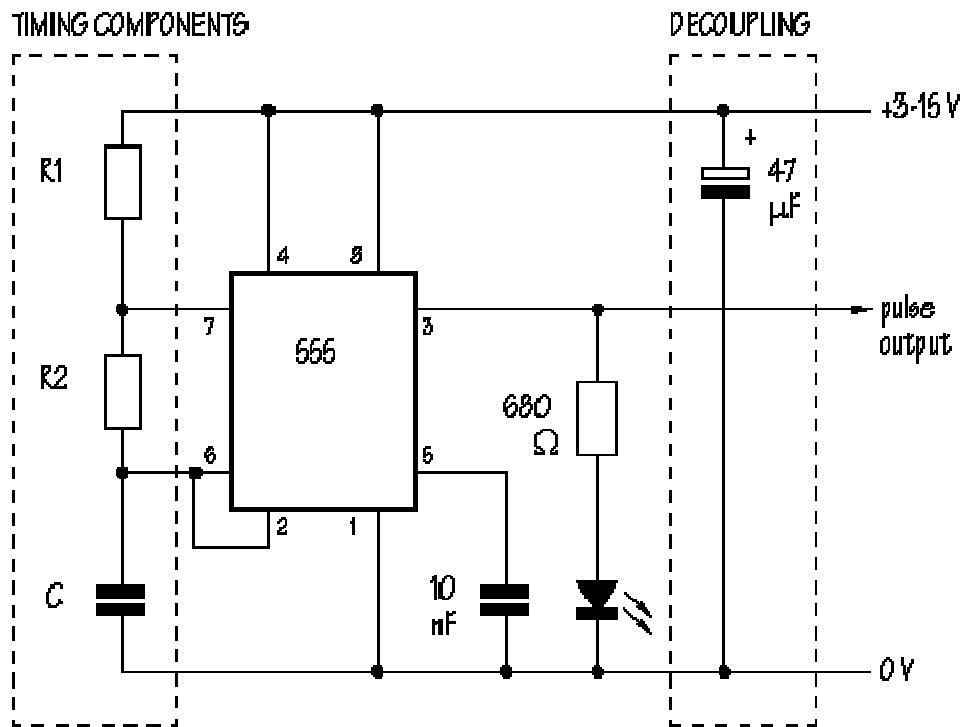


FIG 1 TIMER CIRCUIT

*** End ***

MARKING SCHEME

(For use by Marker/Assessor/Examiner only)

Test Assignment: **Construct a simple 555 timer circuit.**

Test Duration: 1 hour

Performance Criteria	Items to be Measured	Raw Score (R)				Score	
		Award '0' score for non-performance of item or if item performed does not meet specification; otherwise, award score accordingly.					
		1 Unacceptable	2 Satisfactory	3 Good	4 Excellent	Actual R x C x D	Max Marks Possible
1 Connection is done accurately	1.1 Connecting components	1	2	3	4	__ x 2 x 2 = __	4 x 2 x 2 = 16
	1.2 Performing wiring work	1	2	3	4	__ x 2 x 3 = __	4 x 2 x 3 = 24
2 Components are laid out in an orderly manner	2.1 Laying out of components	1	2	3	4	__ x 2 x 2 = __	4 x 2 x 2 = 16
	2.2 Utilising of board space	1	2	3	4	__ x 2 x 2 = __	4 x 2 x 2 = 16
	2.3 Neatness of wiring	1	2	3	4	__ x 2 x 2 = __	4 x 2 x 2 = 16
3 Waveform measurement is accurate	3.1 Measuring work (Vout)	NA	NA	NA	4	__ x 3 x 2 = __	4 x 3 x 2 = 24
	3.2 Measuring work (Pin 2)	NA	NA	NA	4	__ x 3 x 2 = __	4 x 3 x 2 = 24
4 Answers to questions are correct	4.1 Providing answer 1	NA	NA	NA	4	__ x 3 x 2 = __	4 x 3 x 2 = 24
	4.2 Providing answer 2	NA	NA	NA	4	__ x 3 x 2 = __	4 x 3 x 2 = 24
Score (Percentage)						__ ()	184 (100%)

Name and Signature of Marker/Assessor/Examiner _____ Date _____

CHAPTER 17: CONDUCT OF PRACTICAL ASSESSMENT

Introduction

- 1 This chapter cover areas related to the conduct of practical assessment of learners as well as evaluation of the assessment process.

Conduct of Practical Assessment

- 2 In order to conduct a practical assessment well, the assessment will need to be well-organised and implemented. The key aspects that will contribute to a well-organised practical assessment to meet assessment objectives are:
 - Adherence to the assessment plan;
 - Adherence to the practical assessment process;
 - Meeting the needs of candidates with special needs;
 - Adherence to the skills standards for the practical assessment;
 - Observing the assessor's code of practice
 - Collecting evidence of candidates' competence; and
 - Providing feedback and remedial actions

Adherence to Assessment Plan

- 3 The scope and design of a test plan for the conduct of a practical test have been discussed in Chapter 4. The development of the Table of Specifications (TOS), which is a component part of the scope of the test, has been detailed in Chapter 6. An example of a typical practical test or assessment plan is given in Annex A of Chapter 13.
- 4 A test plan is expected to be drawn up for each practical assessment. The number of practical assessments required for a course will determine the number of test plans needed to be prepared. .
- 5 A formal practical assessment is usually scheduled following a stipulated training period, which often has provision for learners to have adequate time to learn and practise their skills. It is expected that the training period would have prepared the learners adequately for the practical assessment.
- 6 Before the assessment can be conducted, a setter will need to be appointed to prepare the following items:
 - Practical test paper based on the test plan (refer to Chapter 16);
 - Tool, Equipment & Materials list setting out the tools, materials and equipment/machines needed (including quantities); and

- Marking scheme.

- 7 The preparation of the test paper is to ensure that the candidates are assessed only on the competencies (based on the skills standards) as prescribed in the test plan while the marking scheme guides the assessor (examiner) to score the test during and/or after the assessment. The preparation of the tools, equipment and materials list is necessary so that the administrator or assessor can prepare the necessary items for the candidates to use during the assessment.

Adherence to Practical Assessment Process

- 8 For a practical assessment to be well-organised and conducted, it should follow a process. The practical assessment process could incorporate the following 3 phases:

- Pre-test Preparation;
- Test Implementation; and
- Post-test Activities.

- 9 Under the Pre-test Preparation stage, the activities to be carried out include:

- Formulation of test rules and regulations setting out the conditions under which the test is to be conducted and that which candidates have to adhere to. These should include provision for candidates to make an appeal for their grades to be reviewed if they are not satisfied with the grades obtained;
- Preparation of an assessment schedule followed by informing the candidates of the schedule in advance;
- Briefing candidates on relevant test information eg, test rules and regulations, assessment schedule and appeal provision, etc. The briefing helps clarify issues and concerns that candidates may have on the practical assessment and procedures. At the same time, the process of appeal can also be explained to ensure that candidates understand the steps to take in making an appeal;
- Appointment of test paper setter. In addition to setting the test paper, the setter also needs to prepare the tools, equipment and materials list and the marking scheme;
- Appointment of assessor (examiner or marker) and moderator, if necessary;

- Appointment of invigilators, if necessary. The invigilators are personnel who will oversee the conduct of the practical assessment. They do not have a role in marking the assessment;
- Preparation of tools, equipment and materials based on the list provide by the setter;
- Preparation of facilities for the test eg, workshops, laboratories and facilities needed for the assessment; and
- Preparation of necessary forms such as candidates' attendance sheet, invigilator's report and examiner's report.

10 Next, is the Test Implementation phase. In this phase, the test is conducted. The activities include:

- Verifying the identities of the candidates to ensure that the right candidates are admitted to take the practical test;
- Assigning candidates to their work stations and clarifying issues that may arise;
- Conduct of test with the issue of test papers to candidates;
- Invigilation of test with the deployment of invigilators, who will attend to the candidates should issues arise;
- Collection of evidence by assessors who have been assigned to be present during the test to carry out process marking guided by a marking scheme or assessment checklist;
- Conclusion of test at the end of the allocated test duration or when all candidates have concluded their practical performance. Candidates will be required to submit their test products (if any) in the manner as instructed by the invigilators.

11 Post-test Activities under the third phase include:

- Completion of invigilator's report, where comments on the test proceedings are provided and any anomalies observed are indicated;
- Submission of test projects submitted by candidates to assessors or markers for product marking;

- Finalisation of test scores and grades by assessors or markers. If necessary, moderation could be carried before the scores and grades are finalised;
- Submission of scores and grades for approval by the proper authority;
- Review of results for appeal cases, if any; and
- Provision of feedback to candidates.

Meeting Needs of Candidates with Special Needs

12 For candidates with special needs, it may be necessary to make alternative arrangements in order for them to take the practical test. These candidates could include those with hearing difficulty, physical limitation and visual impairment. Where possible, special arrangements should be made, within reason and scope of the skills standards allowable, to enable the special needs' candidates to take the test. Some of the special arrangements could include:

- Adjustment made to the assessment process eg, the test duration could be lengthened slightly. Another adjustment could be to arrange for the test to be taken at the training or work area where the special needs' candidates receive their training since they will be expected to be familiar with the tools, equipment and environment at their own training location;
- Provision of support services eg, allocation of additional personnel to be on hand to provide assistance to the special needs' candidates. Such personnel could include trainers and counsellors who could provide assistance such as assisting the candidates to visit the toilet, take certain medication or provide support of a personal nature, etc; and
- Use of special equipment (adaptive technology) eg, use of equipment, machines or tools that have been specially designed and engineered for use by special needs' candidates. Where such special items have been used in training by special needs' candidates, it is within reason for such items to be used for the practical assessment.

Adherence to Skills Standards for Practical Assessment

- 13 If a practical test is set according to requirements as specified in the test plan, the practical test paper set by the setter should contain:
- Information on the specific task assignment or activities to be performed by the candidates;
 - The time allocated to complete the task assignment;
 - The resources to be used (such as tools, materials and equipment) to perform the task assignment; and
 - The assessment criteria (ie how the task assignment or activities will be assessed and scored).

If the task assignment contained in the practical test paper is set according to the specifications in the test plan, it will ensure that the candidates are assessed on their ability to meet the required skills competency as stipulated in the skills standard.

- 14 In other words, if workers of a particular occupation are required to take performance tests and if the tests are set according to the test plan requirements (as specified in a structured skills standard for the occupation) then, those who pass would have acquired the performance competency to meet the performance criteria as stated in the skills standard. In short, they have the ability to do the job.
- 15 Hence, it is important that the skills standard of an occupation prescribe the competencies required for the job and the performance criteria to be achieved for each competency. In this way, the standard and level of performance of workers in that occupation can be assessed or measured objectively.

Observing Assessor's Code of Practice

- 16 As mentioned in Chapter 11, the formulation of a code of practice for assessors is to guide the assessors to conduct themselves in an appropriate manner that is consistent with the performance expectations, values and culture of the an organisation. For practical assessments, the code of practice could include:
- An assessor shall have the skills and knowledge to carry out each practical assessment assignment;
 - An assessor shall be impartial and objective in discharging his practical assessment duties.

- An assessor shall display due diligence in the evidence collection process in a practical assessment; and
- An assessor shall award scores based on evidence collected or sighted.

17 A code of practice varies from one organisation to another. However, in whatever form it takes, the code provides a set of “rules” or “do’s and don’ts” for every assessor in the organisation to comply with. If assessors adhere fully to the code, this will result in uniformity of practice in the conduct of practical assessment. For example, if assessors score the candidates in an objective manner, based on evidence collected and without bias, it will increase the reliability of the test scores. Moreover, the conduct of practical assessments in a systematic, well-organised and fair manner will greatly lessen the need for candidates to be unhappy over the assessments or consider raising an appeal for a review of the test results.

Collecting Evidence of Candidates’ Competence

- 18 As assessment results are based on scores obtained by candidates in a practical assessment and as scores are entered by an assessor based on evidence collected, it is important, therefore, that evidence collection be carried out carefully, accurately and as objectively as possible. In process marking, for example, the scoring by an assessor on the methods used by a candidate will depend very much on the assessor sighting the evidence in the course of the candidate’s practical test performance. The evidence may or may not be totally present. On the other hand, the assessor may have missed sighting the evidence or certain part of the evidence. When that happens, the scoring can be put into question.
- 19 In certain cases, it may be necessary for other forms of evidence to be used to support assessment. Videography could be used, or there could be more than one assessor present during the test.
- 20 In assessing candidates in a practical test, it is common for assessors to be assigned to carry out both process and product assessment. For process marking, in particular, it is required of an assessor to observe each candidate’s performance in relation to the performance criteria and items to be measured while the assessment is in progress. For each measurable item, he is expected to look for evidence as displayed by the candidate before giving the item the appropriate score. If there is a good ‘match’ (ie, the evidence as observed tallies with the item’s requirement), the assessor will score the item with an appropriate high score. On the other hand, if the evidence shows that there is no match, the assessor will award an appropriate low score to the item.

- 21 The process of seeking evidence and scoring is not an easy task in process marking. This is especially so if the assessment assignment is of a short duration and there many items to assess. It is also possible for an assessor to miss out observing evidence on an item inadvertently and, subsequently, assign an inappropriate score to the item. When this happens, it could have an adverse effect on the outcome of the overall score. For this reason, it is, therefore, important for an assessor to focus his entire attention in observing a candidate during his test performance, for example, to display due diligence in the discharge of his duties.
- 22 Product marking is less of a hustle as the marking is done after the practical test is completed. The submitted end products (ie, physical test pieces submitted by the candidates) can be checked, measured and marked at a suitable or appointed following the completion of the test. Hence, the pressure imposed on product marking is relatively less compared to process marking.
- 23 It is recognised that there are limitations in the collection of evidence, particularly, for process marking. Some of the limitations and precautions that could be taken to overcome them are discussed as follows:
- Unclear instructions in assessment papers – For a practical test, instructions refer to the information printed on the test paper eg, information on the tools, equipment and materials to be used, the task assignment to be carried out for the test and the criteria used for the assessment. If instructions are not clear, candidates could misinterpret the instructions and this could affect the way they approach or attempt the test. Hence, a setter of the practical test paper needs to ensure that all instructions on the test paper are specific and easy to understand. Short sentences are preferred over long sentences;
 - Assessment criteria not explicit enough – The candidates are informed through the criteria on how their task assignment will be assessed or marked. If the criteria are not specific enough, vague or misleading, the candidates could misread or misinterpret the criteria. This could cause them to under-perform or perform differently in their task assignment. Hence, it is important that the assessment criteria are explicit. The other approach is the use of a pilot run to test out the papers and gather inputs which could then be used to improve the test paper including the assessment criteria;
 - Insufficient time for assessment – this is a common grouse of candidates sitting for a test. Insufficient time could lead to candidates rushing to complete their task assignment and, therefore, take short-cuts in performing their task elements. This could then lead to the assessor not being able to collect adequate evidence on those task

elements which have not been performed. Grouse aside, it is expected of a setter to ensure that the practical test be set such that the candidates will have enough time to complete. The conduct of a pilot run could provide useful feedback to the setter to make necessary changes to the task assignment or to the duration of the test. If 90% of the pilot run candidates are able to complete the test within the time set, there may not be a need to make changes to the practical test paper;

- Assessment activities not integrated – If assessment activities such as those requiring candidates to move from station to station or perform many sub-tasks (ie one after another), these activities will need to be well integrated and co-ordinated to allow candidates to perform their test activities in a progressive manner. If the flow of the activities or sub-tasks is disjointed, candidates may encounter difficulties in their test performance. This could also pose problems to the assessor in collecting evidence for scoring purposes. On the other hand, if the activities are well integrated and co-ordinated (by those responsible for making arrangements for practical tests) such that the assessment activities are arranged with a logical flow, it will enable continuity in test performance and allow the assessor to collect evidence in a smooth and logical manner; and
- Marking scheme is too complicated – In a practical test, the marking scheme forms the basis for marks to be awarded. A complicated marking scheme will invariably make it more demanding for an assessor to score the task, particularly for process marking which requires collection of evidence while the test is in progress. One way to reduce this difficulty is for the setter to draw up a marking scheme that is easy for the assessor to use (eg, use of specific rubrics in the marking scheme). Consequently, this could enable the assessor to match the evidence collected to the required performance criteria with greater ease, resulting in more accurate scoring and at a faster pace.

- 24 The appointment of more assessors to collect evidence during a practical test will certainly help to reduce or even remove any limitation in evidence collection. In reality, however, it may be easy said than done as there is a cost factor and the issue of manpower availability. To deal with the issue of limitations in evidence collected for a practical test, one logical approach to adopt is to conduct a moderation of the test scores. The moderation process would involve the appointment of another group of assessors to review the marking scheme scores for all the candidates. In the process, the moderators could also re-examine and re-mark the test pieces submitted by the candidates.

Providing Feedback and Taking Remedial Actions by Candidates

- 25 Feedback is a positive way to provide inputs to learners on their practical performance. It can be provided to learners who have performed well as well as those who need to make improvements. Positive feedback given to the better performers can help to motivate them to seek ways to improve on their strengths. Likewise, feedback should also be given to those who have achieved a low level of performance. The feedback should be given constructively to help them to agree to take remedial actions to address their performance gaps.
- 26 Feedback should be a two-way process. In a feedback session with a learner, there is a need to ask the learner if there are areas that could be improved. If the learner asks questions, appropriate responses should be given.
- 27 Constructive feedback calls for dialogue to be directed at the problem and not the person ie, the learner. In this way, feedback can be provided in a way to seek agreement from the learner on remedial actions to be taken to address the performance gap. The suggested approach on 'Giving Constructive Feedback', as detailed in Chapter 11, is a useful way to provide constructive feedback to learners who need to level up their competency level.

Evaluate the Assessment Process

- 28 Evaluation is important activity in assessment. By way of evaluation, an organisation implementing the assessment system will be able to determine the relevance and effectiveness of its assessment system.
- 29 For a practical assessment system to remain relevant and effective, evaluation of the system will need to be carried out in a timely manner. Timely evaluation or review of the assessment system will ensure that the assessment process remains current and is aligned to the needs of the organisation and its learners. The assessment process can also be streamlined to enhance its effectiveness.
- 30 A general rule is for a review to be conducted periodically eg, from 1 to 3 years. Notwithstanding this, a timely review would include conducting a review when the situation calls for it. For example, a significant number of issues related to the practical assessment system have occurred and actions are needed to address them promptly and appropriately. Procrastination could lead to disastrous consequences.

Evaluation Criteria

31 The following criteria could be considered for use to evaluate the practical assessment process in an evaluation exercise:

- Rules of Evidence - refer to specific criteria to be used in the evaluation process. Specifically, these criteria are as follows:
 - Validity – refers to evidence of acquisition of practical skills ie, whether the test assesses the learners based on the learning objectives or competence topics that have been taught. For example, whether the specifications in a practical test paper (ie, task assignment; tools, equipment & materials listing and assessment criteria) are consistent with that specified in the corresponding test plan. A good match between the test plan and the test paper points to a high level of validity. This means that the practical test has been set to meet learning objectives (as specified in the skills standard) and cover the prescribed competencies as specified in the test plan. For learners who have learned well and possess the skills competence, their scores would be expected on the high side (high validity);
 - Reliability – refers to the interpretation of evidence and assessment outcome. In other words, whether the evidence of the test scores can be relied on to determine the skills competence of the learners. Factors such as objectivity of scoring (whether the scores have been awarded based on evidence collected), conduciveness of test environment and quantity of items for scoring (whether the marking scheme has been drawn up to cover all the items that need evidence to be collected and scored) will have an effect on the reliability of test scores. Reliability of test scores will increase with fewer errors of measurement. Moderation by comparing the scores of 2 or more assessors can help to enhance reliability;
 - Fairness – refers to the extent of fairness being exercised in preparing the test paper and awarding scores to all the items in the marking scheme. For example, the choice of words used (whether simple or bombastic), language difficulty (whether direct and easy-to-understand or complex and long-winded) and test duration (whether the time allocation for the test is too short or too long), relative to the language ability and comprehension level of the learners, will have an effect on the fairness of the test paper. Other aspects that can have an impact on fairness could include: whether adequate and appropriate assistance is provided during the test; test procedure being complied with; and fairness of judgement made by assessor when scoring the evidence collected;

- Flexibility – refers to whether the assessment of learners is carried out in more ways than one. For example, a practical assessment can be termed as flexible if the outcome of the assessment takes into account test results taken from 2 or more types of assessment (eg practical course work performance, periodic phase test results, project work performance and summative practical test scores). On the other hand, a practical assessment that comprises only one final practical test may be regarded as inflexible by the learners;
- Sufficiency of evidence – refers to amount of evidence collected. In other words, whether the assessment has collected evidence to satisfy the competency requirements as taught. For example, if the practical test fully complies with the test plan requirements and yet does not satisfy the sufficiency requirement, then, the evidence collected during the test will not be enough to ascertain that the learners have acquired the expected skills competency. In turn, this will mean that the test plan will need to be reviewed to ensure that, when a test paper is set based on the test plan, competencies taught are adequately covered in the test plan;
- Types of evidence – refers to evidence pertaining to process and product. It is common in a practical test for learners to be required to carry out a process in performing the task assignment which may culminate with an end product. For process assessment, an assessor will look for evidence on whether the learners have carried the process according to the assessment criteria before scoring on the marking scheme. Where there is an end product, the evidence will be collected at the post-test phase ie marked after the test with scores entered into the marking scheme;
- Assessment process – refers to the systemic process used for the practical assessment. For example, the process mentioned in Para 8 of this Chapter outlines 3 phases, namely Pre-test Preparation, Test Implementation and Post-test Activities. These phases and process activities could be examined, reviewed and revised as appropriate for the purpose of improvement;
- Assessment method – refers to the current method used in assessing learners' knowledge and skills. Project work, practical phase tests, field work, on-the-job training assessment and practical tests are some examples of the types of test used in practical assessment. Individual organisations may select one or more types of test for their assessment method. Similarly, in an evaluation exercise, the current assessment method could be reviewed to improve on the methodology; and

- Assessment duration – refers to the duration allocated for the practical test. The duration is usually expressed in terms of minutes or hours or both. Test duration for a particular test could be viewed as adequate or inadequate depending on the various factors such as learners' education background, learning ability, perception and special needs, etc. For this reason, assessment duration could be reviewed periodically to align the duration to the needs of current learners.

Methods to Evaluate Assessment Process

- 32 Methods that could be considered for use in evaluating the assessment process are as follows:

Method	Purpose
Evaluation form	To obtain feedback from current learners and assessors on assessment aspects as well suggestions for improvements.
Questionnaire	To obtain responses from current learners who have been exposed to the assessment system for some time.
Survey form	To obtain feedback from current as well as past learners and assessors on assessment aspects and suggestions for improvements.
Interview	To interview learners and assessor separately to obtain views and comments on the assessment system. The interviews could include past learners and assessors.
Focus group	To hold discussions in small groups to address specific issues related to assessment matters. Attendees invited to the discussions could include both present and past learners and assessors;
Assessment Reports	To obtain information on past assessment results and result analyses. The reports will be useful to examine the strengths and weaknesses of the assessment system.

- 33 The selection of one or more methods for the evaluation exercise will depend on the feasibility of usage of the methods, some of which may be difficult to conduct such as surveys and focus group discussions.

Limitations of Assessment Evidence

- 34 It is not uncommon for assessors to encounter problems when assessing learners in practical assessments. Where problems can be anticipated, these could be looked into with actions taken to eliminate or minimise the effects of the problems. When a problem occurs and is not addressed, the effect will have an impact on the extent of assessment evidence that can be collected.

35 Some of the limitations encountered in the collection of assessment evidence and probable solutions are as follows:

No	Limitation	Example	Probable Solution
1	Failure to collect evidence of competence	Assessor has too many candidates to collect evidence from for process marking (wide coverage of candidates).	• Increase the number of assessors to allow each assessor to have enough time to collect the evidence. • Alternatively, spread the assessment over a number of sessions, thus reducing the number of candidates for each session such that the assessor will be able to effectively collect the evidence of competence from all candidates.
2	Faults in design of assessment tools	Difficult or wrong task assignment set or assignment is set outside the scope and design of the test plan by the setter	• Provide training to setter on designing task assignments for practical assessment or provide setter with specimen copies of good test papers used previously for his reference. • Appoint a moderator to vet the draft practical test paper before finalising it for use. This will ensure that the task assignment is set in accordance with test plan and test the candidates at the appropriate skill competence level.
3	Assessment criteria not sufficiently explicit	Words used on how the task assignment will be assessed are either too brief or vague. Words used could be beyond the level of the candidates to comprehend.	• Appoint a moderator to vet the draft practical test papers to ensure that the assessment criteria are clearly stated ie consistent with the test plan and words used are appropriate, simple and easy-to-understand. The moderator can also ensure that the choice of words is within the comprehension level of the candidates.
4	Assessment instructions not clearly communicated to candidates	Instructions on the practical test papers are not well written and with missing information. This could result candidates mis-reading the instructions.	• Standardise the format to be used for practical test papers. Also, introduce standardised instructions to be used on all test papers. These will ensure that whatever general instructions that need to be communicated to candidates are printed on all practical test papers. • Require the invigilator or assessor at the assessment venue to brief candidates on the

			general instructions before start of test. This will offer the candidates the opportunity to clarify any doubt if they have.
5	Time for assessment not enough	Task assignment set by the setter requires many activities to be carried or is too difficult to be completed within the test duration specified.	• Setter to conduct a field test or test run to fine-tune the test duration. • Moderation of test duration by a moderator will help to ascertain whether the test duration allocated in the test plan is both reasonable and realistic. • Review the test duration in the test plan in relation to the stipulated scope and design features. Alternatively, review the features.
6	Assessment activities not integrated	Task assignment has many activities built into it but the activities are not linked or connected in a sequence (no integration)	• Review the test plan such that the scope and design features clearly spell out the integration of activities if the task assignment is set with multiple activities. • Moderations of task assignment to check that the task assignment set is consistent with the intention of the scope and design features of the test plan.
7	Model answers too limited	Setter has not provided model answers with adequate guidance given to assessors. Assessors will have to use their initiatives to interpret the model answers in the award of scores.	• Require the setter to provide model answers that will guide the assessor to award scores for different levels of evidence or competence. • Moderation of test paper to check the marking scheme is provided and that the scheme allowed for ease of scoring by the assessor ie with adequate guidance.

(Note – solutions may not be exhaustive)

Personnel for the Evaluation Process

36 The evaluation of the assessment process will need to be undertaken by the appropriate authority. Relevant personnel to be involved in the evaluation process or exercise could include:

- Management Representative (MR) of the organisation or assessment centre – the MR is normally a high-ranking person and with his participation, the significance of the evaluation exercise can be greatly

enhanced. The MR could play a lead role in setting the roles and responsibilities (terms of reference) for the evaluation team;

- Assessment Manager – the person who takes charge of assessment administration. He could either assist the MR or be the lead person in the evaluation process. If he chairs the evaluation team, he will likely be the one who chairs meetings, deliberate on issues with team, provide guidance and direction to team members and makes decisions (possibly in consultation with the MR);
- Assessment Plan Developer – The person who develops the test plans to be used by setters to set papers for assessment of learners. He will invariably be a content expert or subject matter expert (SME). Where necessary, more than one SME could be appointed to the evaluation team; and
- Curriculum Developer – The person who develops the curriculum for a particular occupation (trade area). The curriculum he develops would include development of the skills standard for the occupation, training specifications, learning objectives, learning materials and so on.

- 37 The key personnel in the review team are the Curriculum Developer and Assessment Plan Developer as they are the content experts. If there are actions to be taken to make improvements to the curriculum documents and test plans, they will be the most appropriate persons to be responsible in carrying out the follow-up actions.

Feedback and Follow-up Actions

- 38 For the outcomes or findings of the evaluation of assessment process to take effect, they will need to be translated into actions to be taken to make improvements to the assessment system. Relevant personnel (besides members of the evaluation team), who need to follow up on the actions needed, will have to be informed and directed on the actions required.
- 39 The relevant personnel could include other curriculum developers and assessment plan developers (who are not on the evaluation team), setters, assessors, moderators, invigilators and other personnel who need to be part of the feedback loop. Communication with them by the evaluation team members could take the form of briefings, meetings and/or dialogue session. Where specific documents are required to be changed, updated or revised, these should be done with specific time-frame for submission and compliance.
- 40 Monitoring of completion of follow-up actions should remain the function of the evaluation team. Where improvements are required to be made to the curriculum and test plans, these should be monitored to ensure completion

according to the time-line. It will further ensure that improvements are, indeed, made so that the tests will be able to fulfil the rules of evidence in assessment. Other areas for improvements could involve specific aspects of the assessment process. The actions to be taken should likewise be monitored so that the overall assessment system could be enhance for greater effectiveness in the implementation of the system.

SUGGESTED READING

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: the classification of educational goals; Handbook I: Cognitive Domain* New York, Longmans, Green, 1956.

Anderson, Lorin W. & Lauren A. Sosniak, eds. (1994). *Bloom's Taxonomy: A Forty-Year Retrospective*. Chicago National Society for the Study of Education.

L. W. Anderson, D. R. Krathwohl, Peter W. Airasian, Kathleen A. Cruikshank, Richard E. Mayer, Paul R. Pintrich, James Raths, and Merlin C. Wittrock (eds) (2000). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives* Allyn and Bacon.

Krathwohl, D. R. (2002). *A revision of bloom's taxonomy: An overview – Theory into Practice*.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives; the classification of educational goals*. Handbook II: The affective domain. New York: Longman.

Krathwohl, D. R. *Methods of Educational & Social Science Research: An Integrated Approach*. 1st Ed. 1993, 2nd Ed. 1998, New York: Longman, also Long Grove, IL: Waveland Press; 3rd Ed 2009, Waveland Press.

Martell, K., & Calderon, T. (Eds.). (2005). *Assessment of student learning in business schools: Best practices each step of the way* (Vols. 1-2). Tallahassee, FL: Associate for Institutional Research.

Walvoord, B. E. (2010). *Assessment clear and simple: A practical guide for institutions, departments, and general education* (2nd ed.). San Francisco: Jossey-Bass.

R. J. Armstrong et al. *Developing and Writing Behavioral Objectives* (Tucson, AZ: Educational Innovators Press, 1970).

Dave, R.H. (1970). *Psychomotor levels*. In R.J. Armstrong (Ed.) *Developing and writing educational objectives*. Tucson AZ: Educational Innovators Press.

Simpson, E.J. (1972). *The Classification of Educational Objectives in the Psychomotor Domain* Washington DC: Gryphon House.

Harrow, A.J. (1972). *A Taxonomy of the Psychomotor Domain* New York.

Likert, R. (1932). A Technique for the Measurement of Attitudes, *Archives of Psychology*, No.140..

Armstrong, Robert (1987). "The midpoint on a Five-Point Likert-Type Scale". *Perceptual and Motor Skills*: Vol 64.

Louis Guttman, 1944. *A Basis for Scaling Qualitative Data*, American Sociological Review 9.

Andrich, D. (1978). *A rating formulation for ordered response categories*. *Psychometrika*.

Guttman, L. (1950). The basis for scalogram analysis. In Stouffer *et al.* *Measurement and Prediction*. The American Soldier Vol. IV. New York: Wiley.

Writing Multiple-Choice Test Items. ERIC Digest. by Kehoe, Jerard (www.ericdigests.org/1997-1/test.html).

Authentic Assessment toolbox (<http://jfmuellet.faculty.noctrl.edu/toolbox/>).

<http://rubistar.4teachers.org/>

<http://www.csus.edu/indiv/p/peachj/330ab/samples/rubric/index.html#criteria>

<http://tep.uoregon.edu/resources/assessment/multiplechoicequestions/mc4critthink.html>

<http://www.ncsu.edu/sciencejunction/route/profession/Assessment/assess.html>

<http://www.slideshare.net/suequirante/item-and-distractor-analysis>

Kelley TL. The selection of upper and lower groups for the validation of test items. J Educ Psychol 1939.

Institute of Technical Education, Singapore, Notes on Test Construction and Assessment Module, January 2009 edition,

Institute of Technical Education, Singapore, .Assessment standards of Test and Examination Division.

