Unit 1. Testing and Assessment: The Definitions

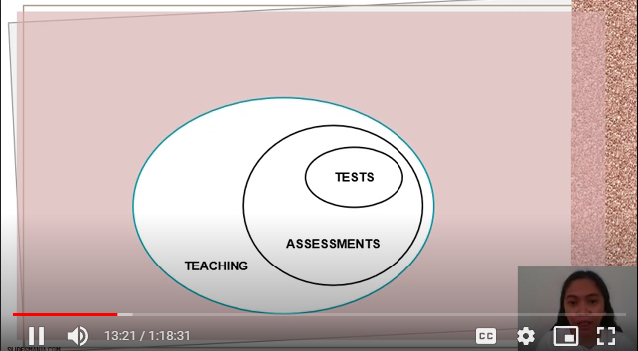
What is a Test?

It is a method of measuring a person’s ability, knowledge, or performance in a given domain.

* The method must be explicit and structured.

Method: is an instrument, it is a set of techniques, procedures, or items that requires performance on the part of the test-taker.

* Measurement on test varies
* Know your test-takers
* Construct adequate tests that would measure the desired criterion



Tests are prepared administrative procedures that occur at identifiable times in a curriculum. When learners must master all their faculties to offer a peak performance knowing that their responses are being measured and evaluated.

Assessment is an ongoing process that encompasses a much wider domain. Whenever a student responds to a question, offers comments on your discussion, or tries out a new word, the teacher is subconsciously making an assessment of the students’ performance. (Observation/judgment). A good teacher should never stop assessing their students. Tests then are actually a subset of assessment.

Does all teaching involve assessment? Are teachers constantly assessing students even without interaction? Depends on your perspective as a teacher.

For optimal learning to take place, students in a classroom must have the freedom to experiment, to try their own hypotheses, without feeling that their overall competence is being judged.

Informal and Formal Assessment

Informal

* It can take a number of forms (accidental, unplanned comments, impromptu feedback)
* It is embedded in classroom tasks. Designed to elicit performance without recording the results

Formal

* These are exercises or procedures. Specifically designed to tap into a storehouse of skills and knowledge.
* They are systematic, planned sampling techniques

Is formal assessment the same as a test? All tests are formal assessments, but not all formal assessment is testing. Formal assessment could be in a form of a tournament, journal, portfolio, etc.

Formative and Summative Assessment

Formative (gina form/ gina hulma)

* Evaluating students in the process of ‘forming’ their competencies and skills with the goal to continue that growth process.
* Informal assessments are formative

Summative

* It aims to measure, or summarize what a student has grasped and typically occurs at the end of a course or unit of instruction.

Norm-reference and Criterion-Referenced

Norm-reference

* Each test-taker’s score is interpreted in relation to a mean, median, standard deviation, percentile rank. Place test-takers along the mathematical continuum.
* The purpose is to place test-takers in a rank.
* Standardized tests. Achievement tests (TOEFEL, CAT, Board Exams)

Criterion-referenced

* It is designed to give test-takers feedback, usually in a form of grades, on a specific course or lesson objectives.
* The distribution of students’ scores may be of little concern as long as the instruments assess appropriate objectives.

Approaches to Language Testing: A Brief History

Discrete-Point

* Are constructed on the assumption that language can be broken down into its component parts and that those parts can be tested successfully.
* Demanded a contextualization that often confuses a test-taker.

Integrative Testing

* Cloze tests and dictations

Cloze Test

* It is a reading passage in which roughly every sixth or seventh word has been deleted.

Dictation

* It is a familiar language-teaching technique that evolved into a testing technique.
* Learners listen to a passage of 100-150 words read aloud by the teacher and write what they hear using correct spelling.
* It has 3 stages:

Oral reading without pauses

Oral reading with long pauses between every phrase

Normal reading speed to give students a chance to see what they wrote.

Communicative Language Testing

* Fundamental principles: correspondence between language test performance and language use.
* In order for a particular language test to be useful for its intended purposes, test performance must correspond in demonstrable ways to language use in non-test situations.

Performance-based Assessment

* Time-consuming but you can actually see the actual efforts.
* Presence of interactive tasks
* Student-centered.
* Oral interview

Current Issues in Classroom Testing

New views on Intelligence:

What is intelligence?

* Was once viewed strictly as an ability to perform linguistic and logical-mathematical problem-solving skills.
* Now, there are different intelligences component proposed by researchers (IQ vs. EQ, Howard Gardner)

Traditional and Alternative Assessment

* One-shot standardized items
* Timed, multiple-choice format
* Decontextualized test items
* Scores suffice for feedback
* Norm-referenced scores
* Focus on the ‘right’ answer
* Summative
* Oriented to product
* Non-interactive performance
* Fosters extrinsic motivation

Alternative

* Continuous long-term assessment
* Untimed, free-response format
* Contextualized communicative tasks
* Individualized feedback and washback
* Vriterion0referenced scores
* Open-ended, creative answers
* Formative
* Oriented to process
* Interactive performance
* Fosters intrinsic motivation

Computer-Based Testing

* Computer-assisted or web-based tests
* It could either be small-scale “home-grown” tests available on websites or standardized. Large-scale tests in which many test-takers are involved.
* Computer-adaptive test in which the test-takers receive a set of questions that meet the test specifications and that are generally appropriate for their performance level. (u cant go back to the previous questions, your first click is your final click)

Advantages:

* Classroom-based testing
* Self-directed on various aspects of a language
* Practice for upcoming high-stakes standardized tests
* Some individualization
* Large-scale standardized tests

Disadvantages

* Lack of security
* Occasional “home-grown” quizzes
* The multiple-choice format
* Open-ended responses are less likely to appear
* Human interaction is absent

Unit II: Principles of Language Assessment

How do you know if a test is effective?

* Can it be given within appropriate administrative constraints?
* Is it dependable?
* Does it accurately measure what it wants to measure?

The purpose is to measure the student’s learning, these are all principles

Practicality

An effective test is practical

* Is not excessively expensive
* Stays within appropriate time constraints (consider the difficulty of the test)
* Relatively easy to administer
* Has a scoring/evaluation procedure that is specific and time-efficient

The value and the quality of the test are sometimes hinged on such practical consideration.

Reliability

A reliable test is consistent and dependable (the test should give similar results)

Factors that may contribute to the unreliability of the test:

* Student-related
  + Physical and psychological factors
  + “test-wiseness” (strategy for effective testing)
* Rater
  + Interrater- this occurs when two or more scores yield inconsistent scores of the same test
  + Intra-rater – is a common occurrence for classroom teachers because of unclear scoring criteria, fatigue, and bias. (one solution: skim and scan)
* Test Administration
* Noise (environment)
* Cleanness of the test paper
* Room lighting
* Temperature
* Conditions of desks and chairs
* Test
* Nature of the test
* Timed tests (may discriminate to the students who do not do well on a timed test)
* Poorly written test items

Validity

The most complex and most important principle. It is the extent to which inferences made from assessment results are appropriate, meaningful, and useful in terms of the purpose of assessment (Gronlund, 1998)

Example: A valid test of a reading ability should actually measure the reading ability.

How is the validity of a test established? There is no final, absolute measure of validity but there are several different kinds of evidence that may involve support. No criteria but there are factors.

* Content-related evidence/ content validity
* This can be done through observation if you can clearly define the goal you want to measure.
* Some tests lack validity but they are practical and reliable.
* Consider the difference between direct and indirect testing.

Direct: involves test-takers to perform the target task. (e.g. say the word and perform where the stress is)

Indirect: learners are not performing the task itself but rather in a task-related test in some way. (e.g. paper and pen to tell where the stress of the word is)

* To achieve content validity in classroom assessment is to test performance directly
* Criterion-related evidence
* The extent to which the ‘criterion’ of the test has actually been reached.
* It is best demonstrated through a comparison of results of an assessment with results of some other measure of the same criterion. Can also use this validity by observing the students’ other test results.
* Two categories:

Concurrent: occurring at the same time. Results are supported by other concurrent performances beyond the assessment itself.

Predictive: to assess a test-taker’s likelihood of future success.

* Construct-related evidence
* Does not play a large role for classroom teachers.
* A construct is any theory or hypothesis or model that attempts to explain observed phenomena in our universe of perceptions.
* ‘Proficiency’ and ‘communicative competence’ are linguistic constructs.
* Construct validity is more of asking a question. Is the test based on a theoretical construct?
* It is a major issue in validating large-scale standardized tests of proficiency/
* Can be done informally.
* Consequential validity
* It encompasses all the consequences of a test:

Accuracy in measuring intended criteria.

Its impact on the preparation of test-takers

Its effect on the learners.

Social consequences of a test’s interpretation and use.

* As teachers, we should consider the effects of assessment on our students’ motivation.
* Face validity
* The extent to which students view the assessment as fair, relevant, and useful for improving learning.
* Refers to the degree to which a test looks right and appears to measure the knowledge or abilities it claims to measure. The subjective judgment of the assessment.
* Students perceive the test to be valid when it's:

A well-constructed, expected format

A test doable within the allotted time

Items are clear and simple

Directions that are crystal clear

Tasks related to their other tasks

A difficulty level with a reasonable challenge.

* Is not something that can be empirically tested by a teacher or even testing experts
* An issue on the psychological state of the learner.
* Experts view this as a superficial factor
* Content validity is a very important ingredient in achieving face validity.

Validity in general is a complex concept yet teachers should use their common sense to understand what makes a good test.

Authenticity

The degree of correspondence of the characteristics of a given language test task to the features of a target language task.

* It can be enacted in the real world
* Authenticity means:

The language in the test is natural

Items are contextualized

Topics are meaningful

Thematic organization of items

Tasks represent real-world tasks

Washback

The effect of testing on teaching and learning

* “cram” courses and “teaching to the test” are examples
* Teachers should create classroom tests in which washback can be achieved
* It enhances intrinsic motivation, autonomy, self-confidence, language ego, interlanguage, among others.
* Comment generously and specifically. Give strategic hints on how to improve their performance
* Positive washback= feedback
* Consider the difference between formative and summative tests
* It implies that students have ready access to you to discuss the feedback and evaluation given. Asking your teachers why your grades are like that is a form of washback. Give students the chance to give feedback on your feedback.

Unit III: Constructs and Models of Language Testing

Construct: a theory. An abstract thing. Describing the components of a model

Model: relatively abstract. Theoretical descriptions of what it means to be able to communicate in a second language. Both are different from the framework. A framework is the selection of skill and ability from a model that is relevant to a specific context.

I.The Nature of Models

All models of language ability have 3 dimensions (McNamara. 1996):

* Model of knowledge: what it means to know a language.
* Model of performance
* Actual language use

These help articulate the ”theoretical rationale”

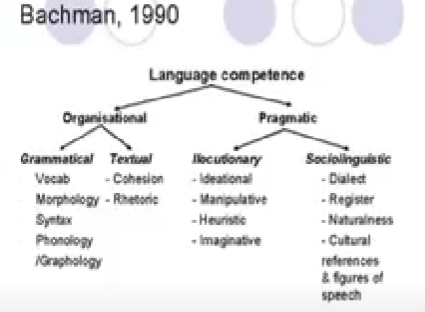
Canale and Swain’s Model (1980)

* Communicative competence leads to more useful and effective second language teaching and allows more valid and reliable measurement of the second language communications skills.
* It is the interaction between grammatical and sociolinguistic competence.
* For assessment, tests need to tap both aspects of communicative competence.
* Its relevance to language testing:
* The distinction between communicative competence and actual performance
* Is viewed as knowledge, discrete-points tests
* Is used to develop criteria for the evaluation of language performance. At different levels of proficiency.
* Communicative competence: grammatical and sociolinguistic (sociocultural rules), strategic (set of compensatory strategies)

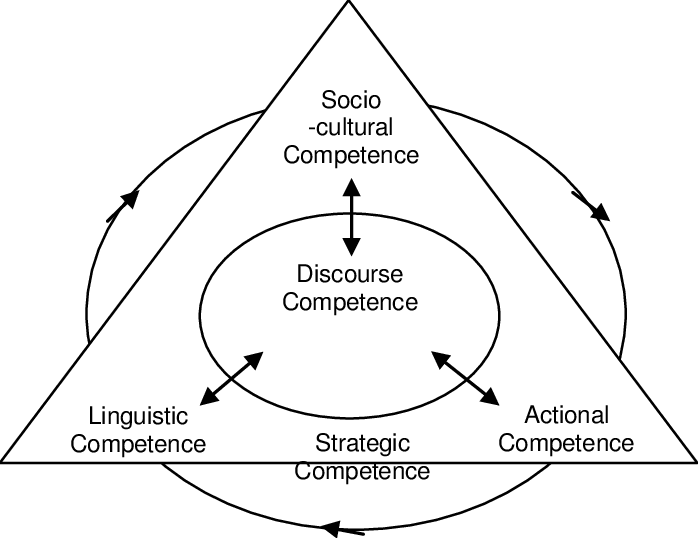
Canale’s Adaptations (1983)

* Communicative competence refers to both knowledge and skill in using this knowledge when interacting in actual communication.
* Actual communication means the realization of such knowledge and skill under psychological and contextual conditions.
* Communicative Competence (Knowledge and skills)
* Grammatical
* Sociolinguistic (sociocultural rules and the rules of discourse)
* Strategic (strategies that enhance the effectiveness of communication)
* Discourse (a unified spoken or written texts in different genres using cohesion in form and coherence in meaning)
* Actual Communication (instances of language use)

Bachman’s Model of Communicative Language Ability (CLA)

* It does two things:
  + Distinguishes knowledge and skill
  + Attempt to characterize the processes and the interaction of various components and the context in which language use occurs.
* Strategic competence: the capacity for implementing the components of language competence in contextualized communicative language use. Psycho-physiological mechanisms which enable the actual execution of language as a physical phenomenon.
* 
* Pragmatic competence: the acceptability of utterances within specific contexts of language use, and the rules determining the successful use of language within specified contexts.
* Illocutionary competence: concern with the performance of the language functions
* Sociolinguistic competence: the sensitivity to, or control of the language conventions of language use that are determined by the features of used contexts.

Celce-Murcia, et. al

* + Attempt to characterize the processes and the interaction of various components and the context in which language use occurs.
* 

Interactional Competence

* Interest in the social context of speech and how communication is understood and constructed in the specific context
* Kramsch (1986): Talk is co-constructed by the participants in communication
* The use of pair or group modes in the testing of speaking.
* Challenge: the reliability of scoring (scoring should be done as a group)

From Models to Framework: Validity Models and Performance Conditions

* Models of communicative competence and performance form an important basis for what test scores may mean, and they help us decide the extent to which the score can be generalized to other performance.
* “Effective-driven” testing: a creation in which the ultimate test design decisions are driven by the impacts that the test will have on

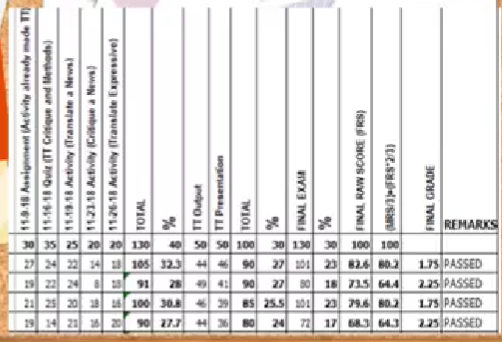
Unit IV: Scoring, Grading, and Giving Feedback

Scoring

When you design a classroom test you must consider how to score it and how to grade it. Your scoring plan reflects the relative way that you place in each section. It will depend on what is the driving force in your overall objectives. Reflects the weight of your test.

Grading

Has something to do with computing all the scores acquired by your students for a certain period. This would also depend on the percentage distribution of the subject.



105 divided by 130 times x 100 x .40 = 32.3%

* Total scores of the student ÷ actual total score of the tests x percent distribution = total percent of the student

32.3 + 27 + 23 = 82.6

* Add all total percent of the student = raw score

MRS x .33

FRS x .67

Add the product = final term score

Giving Feedback

A section of scoring and grading would not be complete without some consideration of the forms in which you will over feedback to your students. The feedback you want to become beneficial feedback.

There are many ways to provide feedback. Washback is achieved when students can through testing experience identify their areas of success and challenge. When a test becomes a learning experience, it achieves washback. Do not delay feedback.

Unit V: Grading and Student Evaluation

Computing grades is not easy as there are some issues or considerations that you have to follow before you give them grades.

1. Philosophy of Grading: what should grades reflect?

* Make a quick, intuitive allocation of the factors to be included in deciding the final grade of a course.
* Factors:
  + Language performance of the students as formally demonstrated on tests, quizzes, and other explicitly scored procedures
  + Your intuitive, informal observation of the student’s language performance (important to inform your students in advance how those observations will be recorded throughout the semester)
  + Oral participation in class
  + Improvement
  + Behavior in class
  + Effort
  + Motivation
  + Punctuality and attendance
  + Gift giving
* Some educational assessment experts state definitively that none of these items should be a factor in grading.
* If you want to incorporate these, you should quantify it properly.
* Gronlund (1998):

“Base grades on students achievement, and achievement only. Grades should represent the extent to which the intended learning outcomes were achieved by students. They should NOT be contaminated by student effort, tardiness, misbehavior, and other extraneous factors. If they are permitted to become part of the grade, as an indicator of the achievement is lost.”

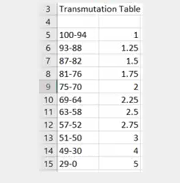
* There are also students whose abilities may not be apparent on achievement tests and measured performance tests. That’s why it’s important to have varied assessment forms.

1. Guideline for selecting grades criteria

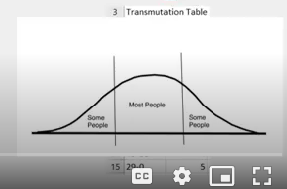
* It is essential for all components of grading to be consistent with an institutional philosophy and/or regulations
* All components of a final grade need to be explicitly stated in writing.
* Challenge yourself to create checklists, charts, and note-taking systems that allow you to convey to the student the basis for your conclusions.
* Consider allocating relatively small weights on other factors so that a grade reflects achievement.

1. Calculating Grades: Absolute and Relative Grading

* It is important to specify your approach to grading
* Absolute grading pre-specify standards of performance on a numerical point system.



* Relative grading is more commonly used than the other. It is accomplished by ranking students in order of performance and assigning cut-off points for grades.



* Issues:
* Is there room in a grading system for a teacher’s intuition, for you ‘hunch’ that the student should get a higher or lower grade than is indicated by performance?
* Should teachers ‘massage’ grades to conform to their appraisal of students beyond the measured performance assessment that has been stipulated as grading criteria?

We can’t. With so many alternative assessments now available to us, we are capable of designing numerous observed performances as criteria for grades. In doing so we can ensure that a final grade fully captures a summative evaluation of a student.

1. Teacher’s Perceptions of Appropriate Grade Distributions

* Teachers have preconceived notions of their own standards for grading.
* Teachers have pre-selected expectations, but in practice may not conform to those expectations.

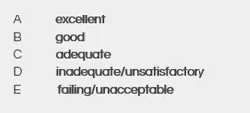
1. Institutional Expectations and Constraints

* Point system is more common in most institutions than letter grades.
* Some institutions offer narrative evaluations

1. Cross-cultural Factors and the Question of Difficulty

* It is important for the teachers to understand the context in which they are teaching.
* Factors:
* It is unheard of to ask a student to self-assess performance
* The teacher assigns a grade, and nobody questions the teacher’s criteria
* A good teacher is one who can design a very difficult test
* Grades of 100, A, 1.0 are reserved to only a highly select few
* One single final exam is a determinant of a student’s entire course grade
* The notion of a teacher’s preparing students to do their best on a test is an educational contradiction.
* How do you design a difficult test?
* Teacher experience
* Adeptness
* Special care in framing items that are clear and relevant
* Mirroring in-class tasks
* Variation of tasks
* Reference to prior tests
* Thorough review and preparation
* Knowledge of your student’s collective abilities
* A little bit of luck

1. What do letter grades ‘mean’?



* Letter grades connote a holistic score that sums up a multitude of performances throughout the course.
* It underscores the meaninglessness of the adjectives. Those letters have come to mean almost everything in their gatekeeping role in admission decision and employment acceptance. Thus as educators, we must explain properly to other people that letter or numerical grades are simply one side of a complex representation of a student’s ability. Even though these are holistic scores, these are their ways to be employed.

1. Alternatives to Letter Grading

* Formative
* Teacher’s comments
* Written reaction to student’s self-assessment
* Review
* Peer-assessment
* Self-assessment
* Teacher-student conference
* Summative
* Teacher comments
* Summative written evaluation remarks
* Written reaction
* Summative checklist
* Narrative evaluations
* Teacher-student conference

1. Summative Alternatives

* Self-assessments
* Checklists, guided journal entry, essay, and teacher-student conference
* Narrative evaluations
* could be used as a reference to future applications
* Could not be easily quantified to grades
* Checklist evaluations (practicality, reliability while maintaining washback)
* Conferences

1. Principles and guidelines for Grading and Evaluation

Characteristics

* Grades are not necessarily based on a universally accepted scale
* It is sometimes subjective and context-dependent.
* Grading tests are often done on the ‘curve’
* It reflects a teacher’s philosophy of grading.
* It reflects the institutional philosophy of grading.
* Cross-cultural variations need to be understood
* Tests do not always yield an expected level of difficulty
* Letter grades may not mean the same thing to all people
* Alternatives to grades are highly desirable additions

Principles

* Develop a personal philosophy of grading that is consistent with your philosophy of teaching and evaluation
* Conform to the institution’s philosophy
* Design tests that conforms to the institutional and cross-cultural expectations of the difficulty that students should experience
* Select appropriate criteria for grading
* Communicate criteria for grading to students at the beginning of the course
* Triangulate grade evaluations with alternative that are more formative and gives more washback.

So these models seek to transform your tests from anguish ordeals into challenging and intrinsically motivating learning experiences

Unit VI: Designing Classroom Language Tests

1. What is the purpose of the test?
2. What are the objectives of the test?
3. How will the test specifications reflect both the purpose and the objectives?
4. How will the test tasks be selected and the separate items be arranged?
5. What kind of scoring, grading, or feedback is expected?

Test Types

* Language aptitude
* Language proficiency
* Diagnostic
* Achievement

Language Aptitude Tests (not common)

* Predicts a person’s success prior to exposure to the second language.
* Designed to measure capacity or general ability to learn a foreign language and ultimate success in that undertaking
* Are ostensively designed to apply to the classroom learning any language. As long as any language you can use this. The student’s scores will determine their success in learning the language.
* Modern Language Aptitude Test: consists of number learning, phonetic script, spelling clues, words in sentences, and paired associates.
* Pinsleur Language Aptitude Battery
* Any tests that claim to predict success in learning a language are undoubtedly flawed.

Proficiency Tests

* Tests global competence in a language
* It tests the overall ability
* Standardized multiple-choice items on grammar, vocabulary, reading comprehension, and aural comprehension
* Almost always summative and norm-referenced
* Since they provide a single score or one for each section of the test, this is a sufficient result for the gatekeeping role they play of accepting or denying someone on accepting or someone’s passage into the next stage of a journey. Usually not equipped feedback.
* Test of English as a Foreign Language
* How the constructs of language ability are specified

Placement Tests

* To place a student into a particular level or section of a language curriculum
* English as a second language placement test: the students read a short article and then write a summary essay; test 2 a composition in response to that article, test 3 read an essay and identify any grammatical errors.
* May come in many varieties: assessing comprehension and production, responding through written and oral performance, open-ended and limited responses, selection, and gap-filling formats.

Diagnostic Tests

* Designed to diagnose specified aspects of a language
* Diagnostic and placement tests may sometimes be indistinguishable from each other. Similarly, there’s a fine line between this and a general achievement test
* It will typically offer more detailed subcategorized information
* Any placements test offers information beyond simply designating a course level may also serve diagnostic purposes.
* Achievement test – gitudlo na, diagnostic= before pa sa pagtudlo
* Have students become aware of their errors

Achievement Tests

* It is related directly to classroom lessons, units, or even a total curriculum.
* It is limited to particular material addressed in a curriculum within a particular time frame and is offered after a course has focused on the objective in question.
* Can serve the diagnostic role
* Its primary role is to determine whether course objectives have been met and appropriate knowledge and skills acquired.
* It is often summative and plays a formative role
* The specifications include objectives, the relative importance of the objectives, the tasks, practicality issues, and the extent of the test structure.

Practical Steps to Test Construction

Equip you with the tools that you need to create classroom-oriented tests.

Assessing Clear, Unambiguous Objectives

What is you want to test?

* Examine the objective for the unit you are testing. Example, students will learn tag questions.
* Determine appropriate testable objectives

Sample Performance Objectives:

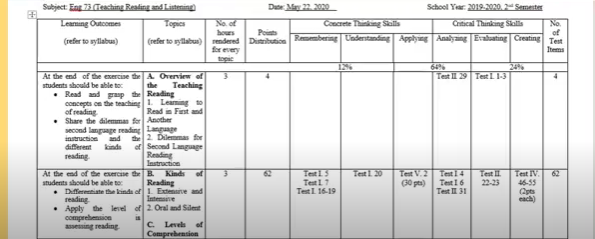
The student will:

* Recognize and produce tag questions, with correct grammatical form and final intonation pattern, in simple conversations.
* Recognize and produce information questions with correct final intonation pattern.

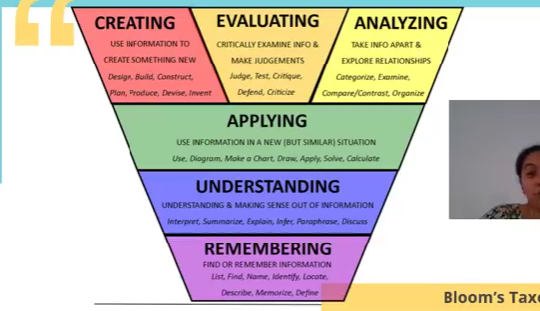
Kung unsay naa sa imong lesson plan or objectives ka, diha ka magkuha/magbase og quiz

Drawing Up Test Specifications

* It could be a simple and practical outline of your test.
* Things to consider: outline of the test, skills included, item types, and tasks.
* Consider the number of test within your time frame



Lesson nga daghay time gi allot = didto magkuha og daghan items sa test



it would not depend on the type of test given but it would depend on the type of question given.

Design Multiple-Choice Test Items

Weakness of Multiple-Choice Items:

* The technique test only recognition knowledge
* Guessing may have a considerable effect on test scores.
* The technique severely restricts what can be tested.
* It is very difficult to write successful items.
* Washback is harmful
* cheating may be facilitated
* It is receptive or selective response
* Stem, Key and distractors
* Difficult to design correctly. Easy to check hard to make
* Avoid making patterns
* Practicality and reliability

Guidelines in Designing Multiple-Choice Test Items

* Design each item to measure a specific objective.

Where did George go after the part last night?

a. Yes, he did.

b. Because ehe was tired.

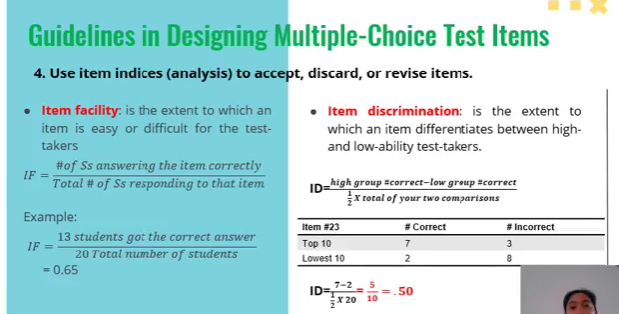
To Elaine;s place for another party.

Around eleven oclock

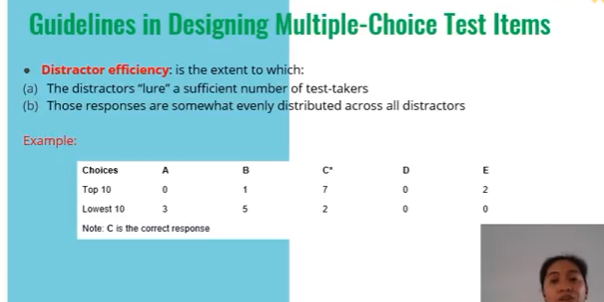
* State both stem and options as simple and direct as possible.
* Make certain that the intended answer is clearly the correct response.
* Distractor efficiency: is the extent to which the distractors ‘lure’ a sufficient number of test-takers; those responses are somewhat evenly distributed across all distractors.
* Use item indices (analysis) to accept, discard, or revise items.

Item facility: is the extent to which an item is easy or difficult for the test-takers

Item discrimination: is the extent to which an item differentiates between high- and low-ability test-takers



We need item facility because sometimes the item is too easy/difficult. There is a reason why one must also put easy items every once in a while to build up the student’s feeling of success (they’re in the very beginning/warm-up).



For CERA: daghan koy narealize sa mga teachers nako before on how they conduct tests. They don’t have the same thorough principles like this sa pagbuhat sa test.

Devising Test Tasks

Do what you can bring to your students an instrument that, to the best of your ability, practical and reliable.

* Imagine you are the student
* Time yourself
* Make sure the test is neat.

Unit VII: Assessing Listening

*These language skills should not be assessed in isolation. The integration of these skills is of paramount importance to language learning. Moreover, assessment is more authentic and provides a more beneficial washback when skills are integrated.*

Observing the Performance of the Four Skills

* You assess the person’s competence, but you observe the person’s performance.
* Important principle: consider the fallibility of the results of a single performance.
* Teachers should triangulate their measurements
* Another principle: rely as much as possible on observable performance

All language users perform the act of listening, speaking, reading, and writing. These skills require competence in order to accomplish these performances. If you want to assess someone’s ability in one or a combination of the four skills, you assess their competence by observing their performance.

Observable Performance of the Four Skills

|  |  |  |
| --- | --- | --- |
|  | Can the teacher directly observe |  |
|  | the process? | the product? |
| Listening | NO | NO |
| Speaking | YES | NO |
| Reading | NO | NO |
| Writing | YES | YES |

^^ can either be receptive or productive.

Listening: the product is not in the form of written or spoken response but within the structures of the brain (hence, the no) which is hard because we can’t observe what’s inside their brain

Speaking: process (bc temporary) and product yes if i-record.

The Importance of Listening

* Occurs when someone is speaking
* Listening is second only to speaking but it is only as good with the latter.
* The input of listening accounts for a large portion of successful language acquisition.

Basic Types of Listening Performance

* Intensive: is listening for a particular purpose (ex. listening based on where the words are stressed, intonation, and pronunciation). Within the language. This is often assessed by measuring an individual’s phonological, morphological, and ability to paraphrase.
* Responsive: responsive listening is listening to a small amount of language such as a command, question, or greeting. After listening, the student is expected to develop an appropriate short response.
* Selective: similar to scanning in reading
* Extensive: listening in the normal or usual thing we do (ex. Listening to a whole passage)

Micro skills of Listening

* Discriminate distinct sounds.
* Retain chunks of language in short-term memory
* Recognize English strestterns
* Recognize reduced forms of words.
* Distinguish word boundaries
* Process speech at different rates of delivery
* Process speech containing pauses, errors, etc.
* Recognize grammatical word clauses
* Detect sentence constituents
* Recognize a particular meaning
* Recognize cohesive devices

Macro skills of Listening

* Recognize the communicative functions of utterances.
* Infer situations, participants, and goals using real-world knowledge
* Form ideas, predict outcomes, causes, and effects, relationship between main and supporting idea, exemplification
* Distinguish literal and implied meanings
* Use non-verbal cues to decipher meanings
* Develop and use a battery of listening strategies

What makes listening difficult?

* Implied in the micro and macro skills of listening is a notion of what makes many aspects of listening difficult.
* Clustering
* Redundancy
* Reduced forms
* Performance variables
* Colloquial language
* Rate of delivery
* Stress, rhythm, and intonation
* interaction

Designing Assessment Tasks: Intensive Listening

* recognizing phonological and morphological elements (minimal pairs)
* paraphrase recognition

Designing Assessment Tasks: Responsive Listening

* question-and-answer format
* appropriate response to a question (wh questions)
* open-ended response to a question

Designing Assessment Tasks: Selective Listening

* listening cloze tasks
* information transfer
* sentence repetition

Designing Assessment Tasks: Extensive Listening

* Dictation
* Communicative stimulus-response tasks
* Authentic listening tasks
* Note-taking
* Editing
* Interpretative tasks
* Retelling

Unit VIII: Assessing Speaking

Challenges in designing speaking tests

* Interaction of speaking and listening
* Design of elicitation tech niques because most speaking is a product of the creative construction of sentences. The speaker has to make some choices as to what words, what meaning, what lexicon, what structure, and what discourse should be used.
* Scoring procedures

Basic Types of Speaking

* Imitative
* Intensive
* Responsive
* Interactive
* Extensive

Micro skills of speaking

* Produce differences among English phonemes and allophonic variants
* Produce chunks of the language of different lengths
* Produce English stress patterns, words in stressed and unstressed positions, rhythmic structure, and intonation contours
* Produced reduced forms of words and phrases
* Use an adequate number of lexical units to accomplish pragmatic purposes
* Produce fluent speech at different rates of delivery
* Monitor one’s own oral production and use various strategic devices: pauses, fillers, self-corrections, and backtracking, to enhance the clarity of the message.
* Use grammatical word classes, systems, patterns, rules, and elliptical forms.
* Produce speech in natural constituents: in appropriate phrases, pause groups, breath groups, and sentence constituents.
* Express a particular meaning in different grammatical forms
* Use cohesive devices (but, first, foremost, however)in spoken discourse.

Macro skills of Speaking

* Appropriately accomplish communicative functions according to situations, participants, and goals.
* Use appropriate styles, registers, implicature, redundancies, pragmatic conventions, conversation rules, floor-keeping, yielding, interrupting, and other sociolinguistic features in f2f conversations.
* Convey links and connections between events
* Convey facial features, kinesics, body language, and other nonverbal cues along with verbal cues.

Important issues:

* No speaking task is capable of isolating a single skill of oral production
* Eliciting the specific criterion you have designated for a task can be tricky
* Specify scoring procedures.

Designing Assessment Task: Imitative Speaking

* Short words

Designing Assessment Task: Intensive Speaking

* Directed response tasks
* Read-aloud tasks
* Sentence/dialogue completion tasks and oral questionnaires
* Picture-cued tasks
* Translation

Designing Assessment Task: Responsive Speaking (short)

* Question and answer
* Giving instructions and directions
* paraphrasing

Designing Assessment Task: Interactive Speaking (lengthy, multiple parti)

* interviews
* Roleplay
* discussion and conversation
* games

Designing Assessment Task: Extensive Speaking

* oral presentations
* picture-cued story telling
* retelling a story, news event
* translation (extended prose)

Unit IX: Assessing Reading

Hurdles to be cleared in order to become an efficient reader

* master bottom-up (data-driven) and top-down(concept-driven) approaches
* develop content (background knowledge) and formal schemata (cultural experience)

The assessment of reading ability does not end with comprehension. Strategies in reading are also important factors to help students summative tests.

Types (genres) of Reading:

* Academic: magazines, newspaper, manuscript, reference materials, journals, thesis, textbooks, editorials, opinion writing
* Job-related: text messages, emails, letters, memos, reports, schedules, labels, signs, announcements, applications, financial documents, directories, manuals
* Personal: greetings, invitations, notes, recipe, maps, advertisement, fiction

Efficient readers also have to know what their purpose is in reading a text. They also have to know the strategies for accomplishing that purpose and how to retain the information.

Micro skills of Reading:

* Discriminate among the distinctive graphemes and orthographic patterns of English.
* Retain chunks of language of different lengths in short-term memory
* Process writing at an efficient rate of speed to suit the purpose.
* Recognize a core of words, and interpret word order patterns and their significance

Macro skills of Reading

* Recognize rhetorical forms of written discourse and their significance of interpretations
* Infer links and connections between events, deduce cause and effect, etc
* Recognize the communicative functions of written text, according to form and purpose.
* Distinguish between literal and implied meanings
* Detect culturally specific references and interpret them in a context of appropriate cultural schemata
* Infer context that is not explicit.

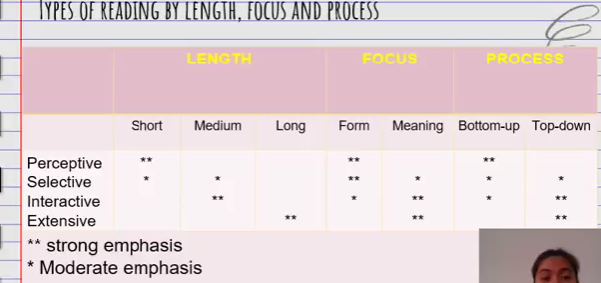
Principal Strategies for Reading Comprehension:

* Identify your purpose in reading a text.
* Apply spelling rules and conventions for bottom-up decoding.
* Use lexical analysis to determine meaning
* Guess at meaning when you aren’t certain
* Skim the text for the gust and for main idea
* Scan the text for specific information
* Use silent reading techniques for rapid processing.
* Use marginal notes, outline, charts, or semantic maps for understanding and retaining information
* Distinguish between literal and implied meanings
* Capitalize on discourse markers to process relationships

Types of Reading:

* Perceptive
* Selective
* Interactive
* Extensive

Types of Reading by length, focus and process



Designing Assessment Task: Perceptive

* Reading aloud
* Written Response
* Multiple-choice
* Picture-Cued Items

Designing Assessment Task: Selective

* Multiple-choice (form-focused criteria)
* Matching tasks
* Editing tasks
* Picture-cued
* Gap-filing tasks

Designing Assessment Task: Interactive

* Cloze tasks
* Impromptu: reading plus comprehension questions
* Short-answer tasks
* Editing (longer texts)
* Scanning
* Ordering tasks
* Information transfer: reading charts, maps, graphs, and diagrams

Designing Assessment Task: Extensive

* Skimming
* Summarizing and responding
* Note-taking and outlining