

The background features a series of overlapping, semi-transparent blue geometric shapes, primarily triangles and polygons, in various shades of blue, creating a modern, abstract design on the right side of the slide.

# **Lecture 4:**

# **Assessment development and administration process**

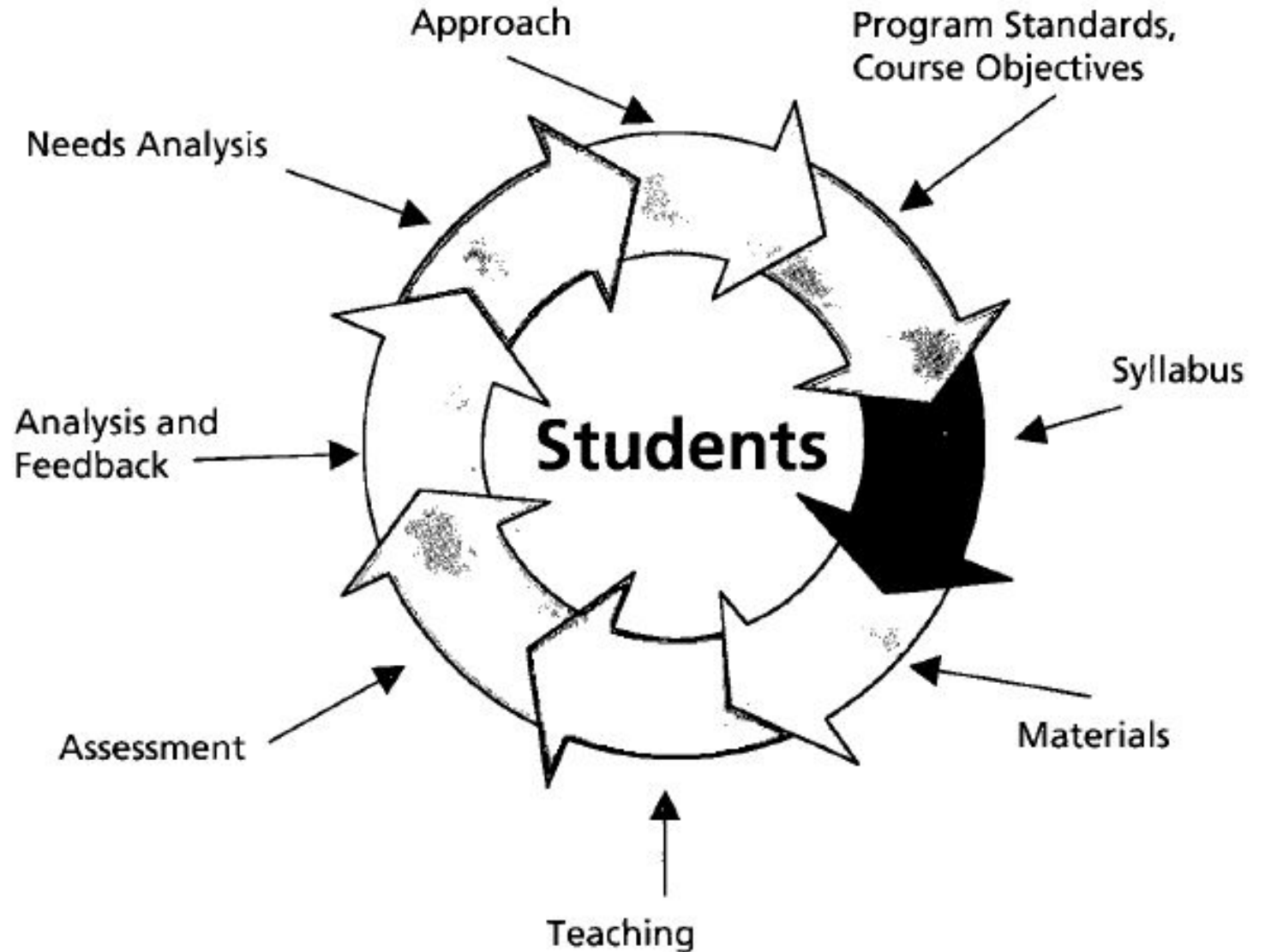
# Objectives of the lecture:

By the end of this lecture you will:

- ▶ understand the process of developing assessment

# Phases of assessment development

- ▶ Planning
- ▶ Development
- ▶ Administration
- ▶ Analysis
- ▶ Feedback
- ▶ Reflection



# Planning Stage 1

## Four Major Purposes for Assessing Learning in Course Design

### Assessing proficiency

*pre course:*  
to place students  
appropriately

*post course:*  
may be done  
to assess  
achievement

### Diagnosing ability/needs

*pre and  
during course:*  
in order to  
identify and  
meet needs

### Assessing Progress

*during course:*  
to assess  
progress

### Assessing Achievement

*at end of  
course or unit:*  
in order to assess  
what has been  
learned and/or  
assign a grade

# Planning stage 2

- ▶ Needs assessment → establish the TLU domain of your students → write goals and objectives / learning outcomes for your course
- ▶ Bachman and Palmer (1996):  
Target language use (TLU) domain is "tasks that the test taker is likely to encounter outside of the test itself, and to which we want our inferences about language ability to generalize"
  - ▶ Real-life domains
  - ▶ Language instruction domains

# Different courses to meet different learner needs

- ▶ A 3-day intensive course on exam strategies for students, taking a university entrance exam
- ▶ A series of one-to-one lessons over 8 weeks on business presentations
- ▶ A six-month course for future tourists focusing on social and daily survival English
- ▶ A short summer course in an English-speaking country for teenagers, involving lots of sports, trips to tourist sites and chatting with English teenagers
- ▶ A once-a-week course for a small group of accountants focusing on English for accountancy, and held in the learners' company
- ▶ A four-week online course on writing business letters in English

# Examples of learning goals / learning outcomes

## General English:

- ▶ communicate with friends in informal situations;
- ▶ read adapted texts and summarize the content;
- ▶ write letters in formal and informal style.

## English for Academic purposes:

- ▶ Improve students' writing of academic papers of different genre (chapters of dissertations, abstracts, conference papers, journal articles)

# Hotel English

After successful completion of this course, students will be able to:

- ▶ Improve their ability to communicate with hotel guests and other staff.
- ▶ Learn the vocabulary of all the main areas in the hotel
- ▶ Gain confidence in using English in the hotel environment.

# English for tourism

- ▶ To improve learners' English so that they can use English effectively and confidently in their future work, especially in the field of tourism
- ▶ To provide students with key vocabulary in tourism in order to help students develop the skills of reading and writing.
- ▶ To develop the communication skills needed to answer the telephone call in a hotel, to deal with guest inquiries and complaints.



# What is the construct?

<b>Construct</b>	<b>Elements often included in construct definitions</b>
<b>Listening</b>	Listening for the main idea, listening for major points, listening for specific details, listening for the gist, inferencing, predicting, determining the meaning of unfamiliar vocabulary from context, distinguishing fact from opinion, determining speaker's intent, note-taking
<b>Reading</b>	Reading for the main idea, reading for major points, reading for specific details, reading for the gist, inferencing, predicting, skimming, scanning, determining the meaning of unfamiliar vocabulary from context, distinguishing fact from opinion, sensitivity to rhetorical organization of the text, sensitivity to cohesion of the text, identifying author purpose or tone, paraphrasing texts
<b>Speaking</b>	Grammar, vocabulary, segmental pronunciation (vowels and consonants), suprasegmental pronunciation (for example, stress, rhythm, intonation, prominence, connected speech phenomena), content, organization, cohesion, task performance, appropriate use or performance of language functions, sociolinguistic appropriacy
<b>Writing</b>	Grammar, vocabulary, content, rhetorical organization, cohesion, task performance, use of appropriate rhetorical mode, register
<b>Grammar</b>	Ability to use structures accurately, ability to comprehend structures, control at the sentence level, control at the discourse or suprasentential level, accuracy of forms, accuracy of meaning (Purpura 2004); may relate to grammar in general, or to specific forms covered in the course or program
<b>Vocabulary</b>	Ability to recognize and understand words and phrases, define or explain them, use them appropriately in context (at the sentential and/or discourse/suprasentential levels); ability to recognize and understand collocations, define or explain them, and use them appropriately in context; may relate to vocabulary in general, or to specific words and phrases taught in the course or program

# Planning stage 3

## Content validity and practicality issues

- ▶ Create an inventory of course and program content, materials used
- ▶ Decide on weighting and scoring/grading (analytical or holistic)
- ▶ Consider validity and practicality issues
- ▶ Autonomy or collaborative effort?

# Stage II - Test Development

- ▶ Map out the assessment instrument
  - ▶ Create test specifications

Reading Specification    Time: 40 mins.

Part	Input text	Question type	Marking	Lexis - all questions
1	<p>A Famous Person or Place</p> <p><u>Multi-paragraph length text.</u> Length: 275-350 words.</p> <p>FK readability grade level 9-10.</p> <p><u>From 5-15 words in each question stem.</u> No more than 10 words in response options</p>	<p><b>Multiple choice</b> with 4 response options</p> <p>Question Types:</p> <ul style="list-style-type: none"> <li>• main idea (whole text) (1)</li> <li>• main idea (paragraph) (1)</li> <li>• local information (1)</li> <li>• supporting detail Qs (3)</li> <li>• vocabulary in context (1)</li> <li>• reference word (1)</li> <li>• <u>inference</u> (1) "Where would this text most likely be found?"</li> </ul> <p>Difficulty index: 3 difficult, 3 easy, 3 <u>mid range</u></p>	<p>9 items 1 mark each</p>	<p>General lexis</p>
2	<p>A Global Problem or Human Interest Story</p> <p><u>Multi-paragraph length text.</u> Length: 425-550 words.</p> <p>FK readability grade level 11-12.</p> <p><u>From 5-15 words in each question stem.</u> No more than 10 words in response options</p>	<p><b>Multiple choice</b> with 4 response options</p> <p>Question Types:</p> <ul style="list-style-type: none"> <li>• main idea (whole text) (1)</li> <li>• main idea (paragraph) (1)</li> <li>• local information (1)</li> <li>• supporting detail Qs (3)</li> <li>• vocabulary in context (1)</li> <li>• reference word—more than one word ok (1)</li> <li>• <u>inference</u> (1)</li> </ul> <p>Difficulty index: This text is the most difficult text on the exam. 3 difficult, 3 easy, 3 <u>mid range</u> Linear text in article format</p>	<p>9 items 1 mark each</p>	<p>General lexis</p>

# Test Specification

- ▶ A detailed description of exactly *what* is being assessed and *how* it is being done
- ▶ Purposes for using specifications:
  - ▶ if multiple versions are used, specifications provide QC to create comparable versions
  - ▶ if used by different teacher, can be a roadmap for test development and grading
- ▶ Usually, specifications include:
  - ▶ a general description of the assessment
  - ▶ a list of tested skills and operations students should be able to do
  - ▶ the techniques for assessing those skills (formats and tasks to be used, types of prompts, expected type of response, timing for the task)
  - ▶ expected level of performance and grading criteria

# Davidson & Lynch Model (1995)

- ▶ No single best format or magic formula for a spec
- ▶ Innumerable ways to design one
- ▶ Davidson & Lynch model is based on the Popham Model (1978) and has five components
  - ▶ General Description
  - ▶ Prompt Attributes
  - ▶ Response Attributes
  - ▶ Sample Item
  - ▶ Specification Supplement

<b>Item No.</b>	<b>General description</b>	<b>Prompt attribute</b>	<b>Response attribute</b>	<b>Sample item</b>	<b>Specification suppl.</b>

# Section 1: General Description

- ▶ The GD section is the object or focus of assessment
  - ▶ Indicates behavior or skill to be tested
  - ▶ Statement of purpose or reason or motivation for testing
- ▶ Normally a capsule summary that can be read quickly is best
- ▶ Example:

The Ss will be able to guess the meaning of certain vocabulary words from context. The texts and words will be of either a scientific, academic or general nature.

## Section 2: Prompt Attributes

- ▶ Called the ‘stimulus’ attributes in Popham model
- ▶ Component of test that details what will be given to test taker
  - ▶ Selection of an item or task format
  - ▶ Detailed description of what test takers will be asked to do
  - ▶ Directions or instructions
  - ▶ Form of actual item or task
- ▶ Isn't usually long or complicated



# Example of Prompt Attribute

- ▶ The student will be asked to write a letter of complaint about a common situation. Each student will be given a written prompt which includes his role, the role of the addressee, and a minimum of three pieces of information to include in the complaint letter.

# Section 3: Response Attributes

- ▶ Part of the specs that details how the test taker will respond to the item or task
- ▶ Often difficult to distinguish from the PA
- ▶ Example:
  - ▶ The test taker will write at least a three paragraph business letter, max. 250 words.
  - ▶ The test taker will select the one best answer from the four alternatives presented in the test item.
  - ▶ The test taker will mark their answers on the answer sheet, filling in the blank or circling the letter of the best alternative.

# Section 4: Sample Item

- ▶ Purpose is to ‘bring to life’ the GD, PA and RA
- ▶ Establishes explicit format & content patterns for the items or tasks that will be written from specs

# Example of Sample Item

- ▶ On a recent flight back home to the UAE, Emirates Airlines lost your baggage.
- ▶ Write a complaint letter to Mr. Al-Ahli, the General Manager, telling him about your problem. Be sure to include the following:
  - ▶ Your flight details
  - ▶ A description of the baggage lost and its contents
  - ▶ What you would like Mr. Al-Ahli to do for you

# Section 5: Specifications Supplement

- ▶ Optional component
- ▶ Designed to allow the spec to include as much detail & info as possible
  - ▶ References or lists of something
- ▶ Anything else that would make the spec appear unwieldy

# Bachman & Palmer Model

- ▶ Bachman & Palmer (1996) spec divided into two parts
  - ▶ Structure of the test
    - ▶ How many parts or subtests; their ordering & relative importance; number of items/tasks per part
  - ▶ Test task specifications
    - ▶ Purpose & Definition of the construct
    - ▶ Time allotment
    - ▶ Instructions
    - ▶ Characteristics of input & expected response
    - ▶ Scoring method

# Alderson, Clapham & Wall Model

- ▶ Alderson, Clapham & Wall (1995) Model
- ▶ Specs should vary in format & content according to audience
- ▶ Different Specs for
  - ▶ Test writer
  - ▶ Test validator
  - ▶ Test user

# Test Writer's Specs

- ▶ General statement of purpose
- ▶ Test battery
- ▶ Test focus
- ▶ Source of texts
- ▶ Test tasks and items
- ▶ Rubrics



## ▶ Test Validator Specs

- ▶ Focus on model of language ability/construct
- ▶ Grading and marking info

## ▶ Test User Specs

- ▶ Statement of purpose
- ▶ Sample items or complete tests
- ▶ Description of expected performance at key levels

# Selecting test items

- ▶ Use tests that come together with the textbook
- ▶ Draw inspiration from professionally designed exams, but do not forget to make necessary modifications
- ▶ Create your own item bank:
  - ▶ a large collection of test items classified according to topics, scale of difficulty, level

# II Stage - Test Development

- ▶ Map out the assessment instrument
  - ▶ Create a specification
    - ▶ decide on form, formats, weighting, components
  - ▶ Construct a draft of the instrument
  - ▶ Establish grading criteria
    - ▶ Prepare an answer key, task descriptors
  - ▶ Pilot the instrument with a representative group of students
  - ▶ Analyze pilot
  - ▶ Do all necessary modifications

### 4 points

- the main idea was clearly stated
- the essay was well organized
- the choice of words was good
- very few minor grammatical errors

### 3 points

- the main idea was fairly clear
- the essay was moderately well organized
- the vocabulary was good
- some minor grammatical errors

### 2 points

- the main idea was indicated, but not clearly
- the essay was not so well organized
- the vocabulary choice was fair
- some major grammatical errors

### 1 point

- the main idea was hard to identify
- the essay was poorly organized
- the vocabulary was weak
- many grammatical errors

## Originality of Content

- **4 points:** interesting ideas were stated clearly
- **3 points:** interesting ideas were stated fairly clearly
- **2 points:** ideas somewhat unclear
- **1 point:** ideas not clear

## Organization

- **4 points:** well organized
- **3 points:** fairly well organized
- **2 points:** loosely organized
- **1 point:** ideas disconnected

## Vocabulary

- **4 points:** very effective choice of words
- **3 points:** effective choice of words
- **2 points:** fairly good vocabulary
- **1 point:** limited vocabulary range of vocabulary

## Grammar

- **4 points:** almost no errors
- **3 points:** few minor errors
- **2 points:** some errors
- **1 point:** many errors

## Cohesion & Logical Consistency

- **4 points:** sentences logically combined
- **3 points:** sentences fairly logically combined
- **2 points:** sentences poorly combined
- **1 point:** many unfinished sentences

# Stage III - Test Administration

- ▶ Before administration
  - ▶ Provide information to students on:
    - ▶ task, coverage, formats, weighting, timing, logistics
    - ▶ assessment schedule throughout the year
  - ▶ For exams, prepare students via
    - ▶ learner training in test-taking strategies
    - ▶ practice test activities

# Stage IV - Administration of assessment

- ▶ Schedule carefully
- ▶ Follow established procedures
  - ▶ use documents from your institution
  - ▶ ensure that all participants are clear on roles, responsibilities, time schedule
- ▶ Organize and check out equipment
- ▶ Have backup supplies and equipment

# After the Administration

- ▶ Grade assessment instrument
  - ▶ Calibrate scorers, if several
  - ▶ Use answer key, criteria for marking
  - ▶ Agree on correction codes and marking
- ▶ Use computer basic statistics or conduct own analysis
  - ▶ overall, by section, item
- ▶ Get results to administration, students, teachers
  - ▶ provide feedback for remediation
  - ▶ channel washback to teachers on curriculum



# Stages 5 and 6 - Analysis and reflection

- ▶ **Reflect on assessment process**
  - ▶ Make time to write impressions while event still fresh
  - ▶ Learn from each assessment
    - ▶ Did it serve its purpose?
    - ▶ What was the “fit” with the curricular outcomes?
    - ▶ Was it valid and reliable?
    - ▶ Was it part of the students’ learning experience?
    - ▶ How could you improve the assessment?
    - ▶ Do not forget to use statistics to analyze your data!