Assessment and International Exams

A.N. KONDAKOVA
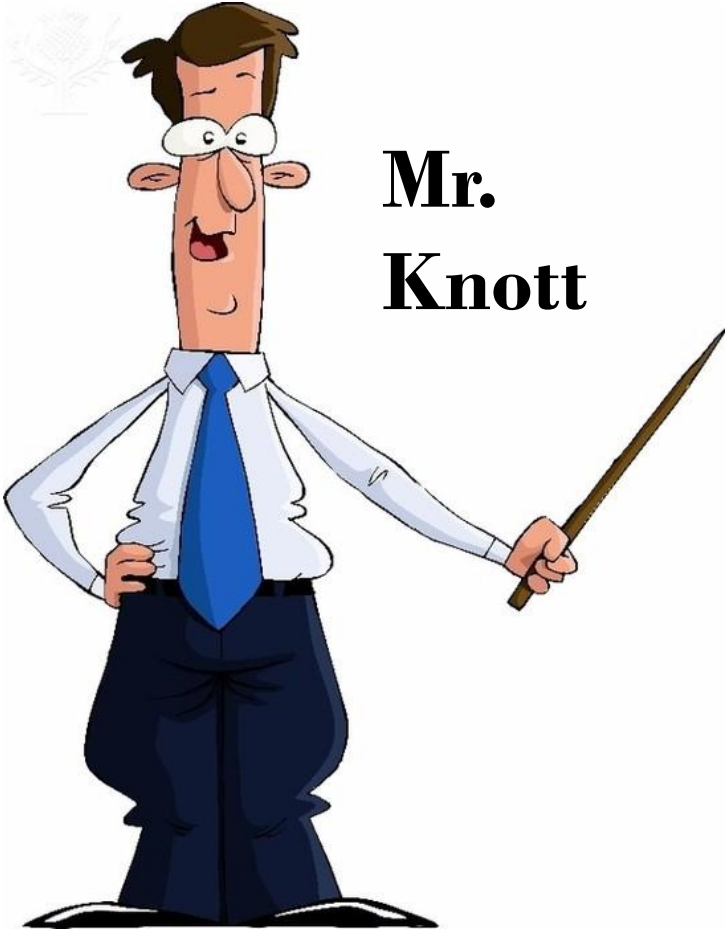
EXPERT, HIGHER SCHOOL OF SOCIAL SCIENCES, HUMANITIES AND INTERNATIONAL COMMUNICATION, PHD STUDENT

Course Outline

Assessment used in this course

Mr. Knott

Class quizzes

Extension activities: for group discussions or self-study

Class presentation on international exams

Student-prepared tasks and tests

Final test

Mrs. Wright

Course literature:

Main course book:

Christine Coombe et al. A practical guide to assessing English Language Learners.

M. Pulverness. A TKT Course. Modules 1, 2 and 3.

Course group:

https://vk.com/club141078612

Websites:

1. http://www.cambridgeenglish.org/exams/cefr/

2. http://www.cal.org/flad/tutorial/index.html

3. http://www.finchpark.com/courses/tkt/unit18.html#

Outline of this lecture

Definition of assessment

Purposes of assessment

What is being assessed

Types of assessment by purpose

Other ways of labelling assessment

Timing of assessment

Practice: observing different types of tests

Generally,

We ASSESS students,

and EVALUATE instruction

Evaluation

 Concerned with the overall program performance (curriculum and syllabuses):
- Are goals and objectives of syllabuses coherent with those of curriculum
- Is the course design effective?
- Do the materials help develop competencies?
- Is there a need to redesign the teaching program?
- How are the SS learning?
- Do the SS develop metadisciplinary competencies?

Assessment

 An ongoing process of gathering, recording, analyzing and reflecting on evidence about pupils' responses to an educational task to make informed and consistent judgements to improve future student learning

(Harlen, Gipps, Broadfoot, Nuttal,1992)

Test

 A test is a formal systematic measuring procedure used to gather information about the student's performance at identifiable times in the curriculum.

 Features of test:
- selected representative samples of language
- has explicit structure
- piloted and pre-tested with a group of students
- measuring competence or performance via individual language items
- provide a result (a grade, a numerical score, a rank etc.)
- used for analysis and reflection
- used to re-teach and observe performance

Newer forms of assessment

- Portfolios
- Classroom observations
- Project-based assessment
- Authentic assessment
- Computer-assisted testing
- Peer- or self-assessment

What do we test?

- Language components vs language use (Skills vs subskills)

- Other skills of using language (pragmatic, discourse and strategic skills)

- Language learning skills

- General learning skills

- Other behavioral or social skills

Message and Medium

Teacher: Miguel, where does the President of the United States live?

Miguel (1): He lives in London.

Miguel (2): He live in the White House.

What do we test?

1. He goes to the cinema every day. They?

2. Find a word in the text that means "angry".

3. On the tape, what does John tell Susan what he wants to visit in London?

4. What is the main idea of the paragraph?

5. Dictation: write down the following…

6. That part of the lesson is finished. What do you feel we need to do next?

Why do we assess students' learning?

Assessment is a systematic way of gathering information for the purposes of making decisions.

The act of giving a test always has a purpose.

'The purpose of language testing is always to render information to aid in making intelligent decisions about possible courses of action. But these decisions are diverse, and need to be made very specific for each intended use of a test'.

(Carroll, 1961)

Why do we assess students' learning?

WHO BENEFITS?

- Teachers
- Students
- Parents
- Heads of departments
- School administrations
- Governments

WHY CONDUCT ASSESSMENT?

- Screening and placement
- Progress monitoring
- Assessment informs instruction
- Motivation and learning
- Practice for later assessments
- Certification
- School accountability

Categorization of tests by purpose:

❖ Admission/Placement tests

❖ Diagnostic tests

❖ Progress tests

❖ Achievement tests

❖ Standardized tests

Admission / Placement tests

- Should a student be admitted to the program at all?
- A single test might be used for both purposes: admission and placement
- Commercially available, but will not readily suit any educational institution
- Should be constructed for particular situation
- Try this one: http://www.cambridgeenglish.org/test-your-english/

Diagnostic tests

- Identify learners' areas of strength and weakness

- "Other types of tests are based on success, while diagnostic tests are based on failure" (Harris and McCann, 1994)

- Straightforward, but at the level of subskills – less informative

Progress tests

- Are Ss mastering course content and meeting course objectives?

- Many progress decisions are made informally

- Formal vs informal assessment

Achievement tests

- How well have Ss met <u>course objectives</u> or mastered <u>course content</u>?

- Accumulate the material from an entire course

- Administered by ministries of education, official examining board or members of other teaching institutions

Proficiency testing

- Do Sts have sufficient command of the language for a particular purpose (studying or working abroad)?

- Not based on a particular curriculum or a language program

- Measure Tts' ability in a language regardless of any language training program they may have received

- Developed by external bodies

Types of assessment

- Formal vs Informal
- Formative vs Summative
- Subjective vs Objective
- Direct vs Indirect
- Classroom/"low-stakes" vs Standardized, "high-stakes"
- Criterion-referenced vs Norm-referenced
- Alternative, authentic vs Traditional tests

Normative vs Criterion-referenced testing

- **NRT**
  Norm referenced tests
- Standardized tests in which the students' proficiency levels are compared to other students in the normative group
- Proficiency tests
  - TOEFL, Cambridge exams, IELTS
  - Broad spread of scores with normal distribution (bell curve)
- Goal: determine S's level
- Expressed as percentiles

- **CRT**
  Criterion referenced tests
- Compares students' performances to stated criteria or outcomes
- Focus on the individual and his/her attainment, competency
- Achievement or progress tests
  - in-course and final assessments
  - qualifying examinations
- Narrower spread of scores
- Goal: determine if S has achieved competencies at particular level
- Expressed in percentages

Reading test score

Student A obtained a score, that placed her on the 25th position among the candidates who have take the test (i.e. she did better than 75% of those who took it).

Or:

Student A: Sufficient comprehension to read simple authentic written material within a familiar context. Can locate and understand the main ideas in materials written for the general reader. Does not have a broad active vocabulary but is able to use contextual clues to understand the text.

What is the major drawback of NRTs?

Summative vs Formative assessment

	Summative	Formative
Time reference	At the end of a learning period	During the process of learning
Purpose of assessment	To measure competency, to determine how well students can do relative to a given concept or skill.	To improve instruction (how to revise or modify instruction, when to move on to new concepts)
Use of results	To give grades and to move levels

Can be used in administrative planning (internal use of results). | Teachers: to plan for and modify instruction, students: to self-monitor and self-assess their understand of new concepts
Results used internally. |
| Example | End-of course test, public exam | Correction (mini-formative assessment) |

Objective vs Subjective testing

- The distinction here lies in the methodology of scoring.

- **An objective test** is one that can be scored objectively and uses selected-response questions (for example, multiple choice or true-false statements);

- A **subjective test** is one that involves human judgment to score, as in most tests of writing or speaking (writing or speaking).

Direct vs Indirect testing

- Direct tests require the test-takers to use the ability (skill) that is being assessed
  - Test skills and subskills

- Indirect tests examine the test takers' knowledge of individual language items
  - Test knowledge of individual language items

Direct test items

Speaking?

Writing?

Reading?

Listening?

Indirect test items

Gap fills: She had a quick shower, but she didn't _____ time to put on her makeup.

Clozes or multiple-choice clozes (every 5ᵗʰ, 6ᵗʰ, 7ᵗʰ, or 8ᵗʰ word is omitted):

The Netherlands

Welcome to the Netherlands, a tiny country that only extends, at its broadest, 312 km north to south, and 264 km east to west - **(1)** … the land area increases slightly each year as a **(2)** … of continuous land reclamation and drainage. With a lot of heart and much to offer, 'Holland,' as it is **(3)** … known to most of us abroad - a name stemming **(4)** … its once most prominent provinces - has more going on per kilometre than most countries, and more English-speaking natives. You'll be impressed by its **(5)** … cities and charmed by its countryside and villages, full of contrasts. From the exciting variety **(6)** … offer, you could choose a romantic canal boat tour in Amsterdam, a Royal Tour by coach in The Hague, or a hydrofoil tour around the biggest harbour in the world - Rotterdam.

Indirect test items

Sentence reordering (or jumbled sentences):

eating (b) cookies (c) his mother's (d) under the tree (e) sat (f) a young fellow (g) fresh-baked

Sentence transformation:

When she got home, Brittany was still tired so she lay down to have a bit of rest (**because**).

If you do not hurry up, you will miss the bus (**unless**).

Indirect test items

Proofreading (underline a mistake in a sentence):

Luckily, she doesn't wearing much makeup.

Matching

A	B
1 How many brothers has Mary got?	a) He's one metre 82.
2 How often does John go to the gym?	b) He's twenty three.
3 How old is John's brother?	c) She's completely fluent.
4 How old is Mary?	d) He's nineteen.
5 How tall is John?	e) Three times a week.
6 How well does Mary speak Chinese?	f) Three.

Dictations?

High-stakes and low-stakes tests

 High-stakes tests are those in which the results are likely to have a major impact on the lives of the Sts

 Low-stakes have a relatively minor on the lives of individuals

Timing of assessment

- Before or outside program?

- At the start of a program?

- During a program?

- End of a program?

Consider a number of tests. For each of them, answer the following questions (if applicable):

o Can you comment on the teaching context and the timing of assessment?

o What is the purpose of the test, and what decisions can be made after the administration of such a test?

o Is it formative or summative?

o Does it contain direct or indirect test items (or a mixture of both)?

o Which test items are objective, and which are subjective?
  * (How can you make subjective test items make less subjective?)

o Is it a high-stakes or a low-stakes test?

o Just looking at the test, can you tell if it is norm-referenced or criterion-referenced?