

Testing

For some of these questions you will have your own ideas and suggestions. If you would like to discuss any of these, you can leave a message in the Forum, join in the scheduled live chats or look at the transcripts from previous sessions.

Part 1: Kinds of Tests by **H.Douglas Brown**
from **Teaching by Principles: 2nd edition, Longman, 2001**

In note form complete the following table

Type of tests	Description
<i>Proficiency tests</i>	<i>Test global competence not isolated skills. Examples include m/c testing of grammar/ vocab/ reading comp/ aural comp. plus a little writing. TOEFL is an example.</i>
<i>Diagnostic tests</i>	<i>Diagnose a particular aspect of a language e.g. pronunciation. Usually, such tests offer checklist of features for administrator to use in pinpointing difficulties. Give teacher information on areas that need to be worked on.</i>
<i>Placement tests</i>	<i>Purpose is to place student into appropriate level or section of language curriculum or school. Typically includes a sampling of material to be covered in the curriculum. Aims to identify point at which student is at in the curriculum.</i>
<i>Achievement tests</i>	<i>Limited to particular material covered in a curriculum within a particular time frame, and are offered after a course has covered the objectives in question. Primary role is to determine acquisition of course objectives at the end of a period of instruction.</i>
<i>Aptitude tests</i>	<i>Designed to measure person's capacity or general ability to learn a foreign language and to be successful in that. Considered to be independent of a particular language. Examples require students to perform such tasks as memorizing numbers and vocabulary, listening to foreign words, and detecting spelling clues and grammatical patterns.</i>

Part 2: **Types of Test Item** by **Jeremy Harmer**
from **The Practice of English Language teaching: 3rd edition, Pearson Ed. Ltd. 2001**

- In your own words, say what is meant by the following:
 - a. Indirect test items: *Indirect test items try to measure a student's knowledge and ability by getting at what lies beneath their receptive and productive skills. They try to find out about a student's language knowledge through more controlled items, such as multiple choice questions or grammar transformation items.*
 - b. Direct test items: *A test item is direct if it asks candidates to perform the communicative skill which is being tested. They try to be as much like real-life language use as possible.*

- Write your own examples of:

[personal answers]

- a. two multiple choice items
 - b. a short cloze test
 - c. a transformation exercise
 - d. two sentence or question re-ordering items.
- Why is it important to create a level playing field?
To ensure that all candidates have the same chance of success. So it is important that the test items do not call on a student's specialist knowledge.
 - Why is it important to replicate real-life interaction?
In real life when people speak or write they generally do so with some real purpose. Language does not occur hypothetically or in a vacuum.

- Complete the following table with your most and least favourite test types from the ones given in the extract:

[personal answers]

	Favourite t-type	Least favourite t-type
Listening		
speaking		
Reading		
writing		

Part 3: **Testing Vocabulary** by **Scott Thornbury**
from **How to Teach Vocabulary**, Pearson Ed. Ltd. 2002

- Complete the table with information about different test items

Type of test	Describe the item type	Give an example
Multiple choice	<i>Can be isolated words, in sentence context or whole text. Easy to score. Seem to be easy to design. But can be guessed or got right by chance. Only test recognition and may be more difficult to design than they seem.</i>	<i>tangle means</i> <i>a. a type of dance</i> <i>b. a tropical forest</i> <i>c. a confused mass</i> <i>d. a kind of fruit</i>
Gap-fill	<i>Test both production and recognition.</i> <i>1. Cloze: write in every nth word</i> <i>2. First letter Cloze: Complete each word after the first letter</i> <i>3. C-test: Complete the second half of every other word.</i> <i>4. Word conversion: convert words from one form to another to fit context.</i>	<i>...another (1) _____ the traveller...</i> <i>...the traveller may e_____ is...</i> <i>...the trav_____ may encou_____ is...</i> <i>1. compose On one occasion the opera was conducted by the _____.</i>
Sentence-writing	<i>Learners write sentences of their own that show the meaning of targeted words. But difficult to assess extent of S's understanding of the word.</i>	<i>Tony was so tired he slept for ten hours.</i>
Paragraph writing	<i>Learners set the task of writing a whole text that includes the selected vocabulary items. This is feasible only if the words themselves are likely to co-occur.</i>	<i>Write a paragraph of about 100 words to include at least six of the following ten words.</i>

- Compare your table with the information in the extract '**Types of Test Item**'. Are the same items described or is there a difference between the two extracts?

[personal answer]

- What is the difference between qualitative and quantitative assessment?

Qualitative evaluation is carried out according to the impression made on the reader set against broad bands or scales reflecting range and accuracy. Quantitative assessment involves actual counting or measuring of, e.g. lexical density, lexical variety, lexical sophistication.

- Explain the relevance of the following to quantitative assessment:
 - a. lexical density: *a measure of the proportion of content words in a text. Content words – as opposed to function words – are words that carry a high information load, such as nouns, adjectives and verbs. Written text that contains a high proportion of such words is characteristic of proficient writers.*
 - b. lexical variety: *a measure of the different words in the text. A high proportion of different words is an indicator of an extensive vocabulary knowledge – what is often called range.*
 - c. lexical sophistication: *assessed by counting the number of relatively infrequent words in a text – such as the number of words that fall outside of a list of the top 2,000 most frequent words.*

Part 4: **Some Practical steps to test construction** by H. Douglas Brown
from **Teaching by Principles: 2nd edition, Longman, 2001**

- Put the following in order [by writing the number in the brackets] and briefly describe what the author means in each case.

[3] Draft your test

Will give you a good idea of what the test will look like, how students will perceive it (face validity), the extent to which authentic language and contexts are present, the length of the listening stimuli, how well a storyline comes across, how things like the cloze testing format will work, and other practicalities.

[5] Final edit and type your test

Imagine that you are one of your students. Go through each set of directions and all items slowly and deliberately, timing yourself as you do so. Check components for timing and difficulty.

[2] From your objectives draw up test specifications

Your specifications will indicate how you will divide up the test period, what skills you will test, and what the items will look like.

[4] Revise your test

Work through all the items you have devised and ask a number of important questions.

[1] Test towards clear unambiguous objectives

Don't just test because, for example, 'It's Friday.' Carefully list everything that you think your students should "know" or be able to "do," based on the material the students are responsible for.

[6] Utilize your feedback after administering the test

Doing the test will give you some information about how easy or difficult it was, about the time limits, and about your students' affective reaction to it and their general performance. Take note of these forms of feedback and use them for making your next test.

- In what way are the following examples of 'alternative assessment options'?

They all offer opportunities for assessment that are not traditional, formal tests.

- a. Self and peer-assessment
- b. journals
- c. Conferences
- d. Portfolios
- e. Co-operative test construction

- In your own words describe the cooperative pair work and self-evaluation oral test designed by Tim Murphey, and which the author cites.

[personal answer]

Part 5: **Writing and Marking Tests** by **Jeremy Harmer**
from **The Practice of English Language teaching: 3rd edition, Pearson Ed. Ltd. 2001**

- What do you understand by the following:

a. assess the test situation

Remind yourself of the context in which the test takes place. Decide how much time should be given to the test-taking, when and where it will take place, and how much time there is for marking.

b. decide what to test

List what you want to include in the test. Take a conscious decision to include or exclude skills such as reading comprehension or speaking (if speaking tests are impractical). Be clear what syllabus items can be legitimately included (in an achievement test), and what kinds of topics and situations are appropriate for your students.

c. balance the elements

Decide balance of direct to indirect test items. Estimate how long each section of the test should take and then write test items within those time constraints. The amount of space and time we give to the various elements should also reflect their importance in our teaching.

d. weight the scores

Make sure marks reflect the relative importance of various test items.

e. make the test work

Try out individual items and/or whole tests on colleagues and students alike before administering them to real candidates.

- How will the following help to get round the problem of test-marking unreliability?

a. training

If scorers have seen examples of scripts at various different levels and discussed what marks they should be given, then their marking is likely to be less erratic than if they come to the task fresh. If scorers are allowed to watch and discuss videoed oral tests, they can be trained to 'rate the samples of spoken English accurately and consistently in terms of the pre-defined descriptions of performance'

b. more than one scorer

The more people who look at a script, the greater the chance that its true worth will be located somewhere between the various scores it is given.

c. global assessment scales

'Pre-defined descriptions of performance' which say what students need to be capable of in order to gain the required marks.

d. analytic profiles

Marking gets more reliable when a student's performance is analysed in much greater detail. Instead of just a general assessment, marks are awarded for different elements.

e. scoring and interacting during tests

Separating the role of scorer (or examiner) from the role of interlocutor (the examiner who guides and provokes conversation). This may cause practical problems, but it will allow the scorer to observe and assess, free from the responsibility of keeping the interaction with the candidate or candidates going. In many tests of speaking, students are now put in pairs or groups for certain tasks since it is felt that this will ensure genuine interaction, and will help to relax students in a way that interlocutor-candidate interaction might fail to do on its own.

Part 6: **Speaking Tests** by Lynda Taylor
from **Modern English Teacher**, 12/3, 2003

- Jargon buster! Read these descriptions and then find the words or phrases in the article which apply to them. The first letter of each word is given.
 - a. A close study of a written record of what someone said: *detailed transcription analysis*.
 - b. A list of qualities for an assessor to use when examining a candidate: *observational checklist*.
 - c. Written statement saying what abilities to expect at different levels: *performance descriptor*.
 - d. The fact that the tester has more influence than the candidate: *asymmetrical power relationship*.
 - e. A group of people who are very like the people who will actually take the test: *representative sample of the test candidature*.
 - f. The tester who speaks to the candidates: *interlocutor*.
 - g. The script that the tester who speaks should follow: *interlocutor frame*.
 - h. A way of measuring the performance of oral examiners: *oral examiner monitoring checklist*.
 - i. A person who listens to candidates and decides what scores to give them: *assessor*.

- Complete the table with information from the article, and your own opinions.

	Advantages	Disadvantages
One-to-one	<i>[personal answers]</i>	<i>[personal answers]</i>
Paired candidate format	<i>Reflects communicative approaches.</i> <i>Gives a more varied sample of spoken interaction.</i> <i>Reduces potentially negative effect of asymmetrical power relationship.</i> <i>Generates richer and more varied sample from each candidate.</i> <i>Encourages greater interactivensness.</i>	<i>[personal answers]</i>



Teacher Development Pack Suggested Answers

Part 7: **Teaching the Test** by **Jeremy Harmer**
from **The Practice of English Language teaching: 3rd edition, Pearson Ed. Ltd. 2001**

- Of the six ways suggested for teaching exam students which is the one that:

[personal answers]

- a. you think most important?
- b. is the most difficult
- c. you don't agree with
- d. students will appreciate most