KINDS OF TESTS AND TESTING

Василевская В.Л.

Белорусский государственный университет

Липовка В.Г.

Военная академия Республики Беларусь

Test construction is a matter of problem solving, every teaching situation sets a different testing problem. In order to arrive at the best solution for any particular problem it is important to choose the most appropriate test or testing system.

In our work we use four types of tests: proficiency tests, achievement tests, diagnostic tests, and placement tests.

Proficiency tests

Proficiency tests are designed to measure people's ability in a language, regardless of any training they may have had in that language. The content of a proficiency test, therefore, is not based on the content or objectives of language courses that people taking the test may have followed. It is based on a specification of what candidates have to be able to do in the language in order to be considered proficient, it means having sufficient command of the language for a particular purpose. The function of such tests is to show whether candidates have reached a certain standard with respect to a set of specified abilities.

Though there is no particular purpose in mind for the language, these general proficiency tests should have detailed specifications saying just what it is that successful candidates have demonstrated that they can do. Each test should be seen to be based directly on these

specifications. All users of a test (teachers, students, employers, etc.) can then judge whether the test is suitable for them, and can interpret test results.

Despite differences between them of content and level of difficulty, all proficiency tests have in common the fact that they are not based on courses that candidates may have previously taken.

Achievement tests

In contrast to proficiency tests, achievement tests are directly related to language courses, their purpose being to establish how successful individual students, groups of students, or the courses themselves have been in achieving objectives. They are of two kinds: final achievement tests and progress achievement tests.

Final achievement tests are those administered at the end of a course of study. They may be written and administered by ministries of education, official examining boards, or by members of teaching institutions. Clearly the content of these tests must be related to the courses with which they are concerned.

In the view of some testers, the content of a final achievement test should be based directly on a detailed course syllabus or on the books and other materials used. This has been referred to as the <u>syllabus-content approach</u>. It has an obvious appeal, since the test only contains what it is thought that the students have actually encountered, and thus can be considered, in this respect at least, a fair test. The disadvantage is that if the syllabus is badly designed, or the books and other materials are badly chosen, the results of a test can be very misleading. Successful performance on the test may not truly indicate successful achievement of course objectives. For example, a course may have as an objective the development of conversational ability, but the course

itself and the test may require students only to utter carefully prepared statements about their home town, the weather, or whatever. Another course may aim to develop a reading ability in English, but the test may limit itself to the vocabulary the students are known to have met. In each of these examples test results will fail to show what students have achieved in terms of course objectives.

The alternative approach is to base the test content directly on the objectives of the course. This has a number of advantages. First, it compels course designers to be explicit about objectives. Secondly, it makes it possible for performance on the test to show just how far students have achieved those objectives. This in turn puts pressure on those responsible for the syllabus and for the selection of books and materials to ensure that these are consistent with the course objectives. Tests based on objectives work against the perpetuation of poor teaching practice, something which course-content-based tests, almost as if part of a conspiracy, fail to do. They will provide more accurate information about individual and group achievement, and it is likely to promote a more beneficial backwash effect on teaching.

It might be argued that to base test content on objectives rather than on course content is unfair to students. If the course content does not fit well with objectives, they will be expected to do things for which they have not been prepared. In a sense this is true. But in another sense it is not. If a test is based on the content of a poor or inappropriate course, the students taking it will be misled as to the extent of their achievement and the quality of the course. Whereas if the test is based on objectives, not only will the information it gives be more useful, but there is less chance of the course surviving in its present unsatisfactory form. Initially some students may suffer, but

future students will benefit from the pressure for change. The long-term interests of students are best served by final achievement tests whose content is based on course objectives.

Is there any real difference between final achievement tests and proficiency tests? If a test is based on the objectives of a course, and these are equivalent to language needs on which a proficiency test is based, there is no reason to expect a difference between the form and content of the two tests. Two things have to be remembered, however. First, objectives and needs will not typically coincide in this way. Secondly, many achievement tests are not in fact based on course objectives.

Progress achievement tests, as their name suggests, are intended to measure the progress that students are making. They contribute to formative assessment. There should be established a series of well-defined short-term, objectives, progress tests based on short-term objectives will fit well with what has been taught.

In addition to more formal achievement tests that require careful preparation, teachers should feel free to set their own 'pop quizzes'. These serve both to make a rough check on students' progress and to keep students on their toes.

Diagnostic tests

Diagnostic tests are used to identify learners' strengths and weaknesses. They are intended primarily to ascertain what learning still needs to take place. These tests will help to see who is particularly weak in, say, speaking as opposed to reading in a language.

We may be able to analyze samples of a person's in writing or speaking in order to create profiles of the student's ability with respect

to such categories as 'grammatical accuracy' or 'linguistic appropriacy'.

Diagnostic tests are extremely useful for individualized instruction or self-instruction. Learners can be shown where gaps exist in their command of the language and can be directed to sources of information, exemplification practice.

Placement tests

Placement tests, as their name suggests, are intended to provide information that will help to place students at the stage (or in the part) or the teaching program most appropriate to their abilities. Typically they are used to assign students to classes at different levels.

Placement tests that are most successful are those constructed for particular situations.

Direct versus indirect testing

Testing is said to be *direct* when it requires the candidate to perform precisely the skills that we wish to measure. If we want to know how well the candidates can write compositions, we get them to write compositions. If we want to know how well the candidates pronounce a language, we get them to speak. The tasks and the texts that are used, should be as authentic as possible.

Direct testing is easier to carry out when it is intended to measure the productive skills of speaking and writing.

Direct testing has a number of attractions. First, provided that we are clear about what abilities we want to assess, it is relatively straightforward to create the conditions which will elicit the behaviour on which to base our judgments. Secondly, at least in the case of the productive skills, the assessment and interpretation of students' performance is also quite straightforward. Thirdly, since practice for

the test involves practice of the skills that we wish to foster, there is likely a helpful backwash effect.

Indirect testing attempts to measure the abilities that underlie the skills in which we are interested.

Perhaps the main appeal of indirect testing is that it seems to offer the possibility of testing a representative sample of a finite number of abilities which underlie a potentially indefinite large number of manifestations of them.

The main problem with indirect tests is that the relationship between performance on them and performance of the skills in which we are usually more interested tends to be rather weak in strength and uncertain in nature. We do not yet know enough about the component of, say, composition writing to predict accurately composition writing ability from scores on tests that measure the abilities that we believe underlie it. We may construct tests of grammar, vocabulary, discourse markers, handwriting, punctuation, and what we will. But we will not be able to predict accurately scores on compositions (even if we make sure of the validity of the composition scores by having people write many compositions and by scoring these in a valid and reliable way).

As far as proficiency and final achievement tests are concerned, it is preferable to rely on direct testing. Of course, to obtain diagnostic information on abilities, such as control of particular grammatical structures, indirect testing may be perfectly appropriate.

Some tests are referred as *semi-direct.* The most obvious examples of these are speaking tests where candidates respond to tape-recorded stimuli, with their own responses being recorded and later scored. These tests are semi-direct in the sense that, although not direct, they simulate direct testing.

Discrete point versus integrative testing

Discrete point testing refers to the testing of one element at a time, item by item. *Integrative testing,* by contrast, requires candidate to combine many language elements in the completion of the task. This might involve writing a composition, making notes while listening to a lecture, taking a dictation, or completing a cloze passage. Discrete point tests will almost always be indirect, while integrative tests will tend to be direct.

Objective testing versus subjective testing

The distinction here is between methods of scoring. If no judgment is required on the part of the scorer, then the scoring is *objective*. A multiple choice test, with the correct responses unambiguously identified, would be a case in point. If judgment is called <u>for, the </u>scoring is said to be *subjective.* There are different degrees of subjectivity in testing. The impressionistic scoring of a composition may be considered more subjective than the scoring of short answers in response to questions on a reading passage.

Computer adaptive testing

In most paper and pencil tests, the candidate is presented with all the items, usually in ascending order of difficulty, and is required to respond to as many of them as possible. This is not the most economical collecting information on someone's ability. People of high ability (in relation to the test as a whole) will spend time responding to items that are very easy for them - all, or nearly all, of which they will get correct. We would have been able to predict their performance on this items from their correct response to more difficult items. Similarly, we could predict the performance of people of low

ability on difficult items, simply by seeing their consistently incorrect response to easy items.

Computer adaptive testing offers a potentially more efficient way of collecting information on people's ability. All candidates are presented with an item of average difficulty. Those who respond correctly are presented with a more difficult item; those who respond incorrectly are presented with an easier item. The computer goes on in this way to present individual candidates with items that are appropriate for their apparent level of ability raising or lowering the level of difficulty until a dependable estimate of their ability is achieved. Oral interviews are typically a form of adaptive testing, with the interviewer's prompts and language being adapted to the apparent level of the candidate.

BIBLIOGRAPHY

Fulcher, G. An English language placement test: issues in reliability and validity. / G.Fulcher. - Cambridge University Press, 2007.- 114 c.

Hughes, A. Testing for Language Teachers / A.Hughes. - Cambridge University Press, 2003.- 58 c.