

On Reading Test and Its Validity

Chao Chen

Foreign Languages School, Qingdao University of Science and Technology

Qingdao 266061, China

E-mail: chenchao1977@163.com

Abstract

Reading ability is undoubtedly a useful one when we use a foreign language. Therefore, reading comprehension has long been hold as an important way of testing testees' language ability. Not matter what test it should have validity. Otherwise it cannot test what it intends to test. The present problem is that some reading tests lack validity and are not efficient tests. Test constructors should be well aware of this problem and pay attention to that and try to avoid composing such tests again.

Keywords: Reading comprehension, Validity, Subjectivity

1. Introduction

Reading ability has long been regarded as one of the most important abilities when learning a language. Therefore it is the ultimate goal in college English teaching. Various teaching syllabus set "good reading comprehension" as a goal of training students' language ability. This is a practical goal, because, on one hand, in our country students do not have an authentic English language environment, and it is hard for them to learn English well through listening and speaking. Thus reading remains the most important way to know more about the language itself. On the other hand, for most people, what they need most in their future job is to get as much as information through reading. Suppose Chinese students can speak English and can understand others but cannot read well, they would be like those who can speak and understand Chinese but cannot read and write Chinese, namely, the illiterate. Therefore, it is easy to understand why in a large variety of examinations, reading comprehension holds a large proportion in the whole paper. As we know, these examinations include: non-English major College English Test-4 (CET-4), CET-6, non-English major post graduate entrance examination, Test for English Major-4 (TEM-4) and TEM-8, even in such examinations as TOEFL, GRE and IELTS. Test of reading can help us evaluate the quality of teaching and to tell whether students have mastered certain language ability. Therefore, the validity of a reading test influences directly our evaluation of teaching and it has a guiding significance to our English teaching.

2. Theories of Test Validity

Testing is a matter of using data to establish evidence of learning. But evidence does not occur concretely on the natural state, but is an abstract inference. It is a matter of judgment. The question arises as to who makes this judgment, and how we can decide how valid the evidence is.

Briefly, the validity of a test is the extent to which it measures what it is supposed to measure and nothing else. Every test, whether it is a short, informal classroom test or a public examination, should be as valid as the constructor can make it. The test must aim to provide a true measure of the particular skill, which it is intended to measure: To the extent that it measures external knowledge and other skills at the same time, it will not be a valid test. For example, the following test item is invalid if we wish solely to measure writing ability: 'is photography an art or a science? Discuss.' It is likely to be invalid simply because it demands some knowledge of photography and will consequently be easy for certain students.

There are many different kinds of validity, but only five are vital for the teacher or tester setting his own tests: content validity, face validity, constructs validity, concurrent validity and predictive validity. This paper is going to focus on content validity.

Content validity is a form of validity which is based on the degree to which a test adequately and sufficiently measures a particular skills or behavior. For example, a test of assessing pronunciation skills in a language could not be said to have content validity if it tests only some of the skills which are required for accurate pronunciation, such as isolated sounds, but not stress and intonation. The test would have content validity only if it included a proper sample of the relevant structures.

3. Skills That a Reading Test Intends to Test

There may be different skills at different levels. The following may be thought of as macro-skills, directly related

either to needs or to course objectives:

- Scanning text to locate specific information
- Skimming text to obtain the gist
- Identifying stages of an argument
- Identifying examples presented in support of an argument

And the following are micro-skills:

- Identifying referents of pronouns, etc.
- Using context to guess meaning of unfamiliar words
- Understanding relations between parts of text by recognizing indicators in discourse, especially for the introduction, development, transition, and conclusion of ideas

And the following are recognized as the exercise of straightforward grammatical and lexical abilities:

- Recognizing the significance of the use of the present continuous with future time adverbials
- Knowing that the word 'brother' refers to a male sibling

As mentioned above, the validity of a test is the extent to which it measures what it is supposed to measure. If a reading test cannot measure what it is intended to test or if students can get the answers to the test questions by other means than through reading, then, the judgment that we make about the teaching effect and the students' reading ability would not be accurate. Therefore it is important for testers to compose find appropriate reading materials to compose the reading test.

Successful choice of reading materials depends ultimately on experience, judgment, and a certain amount of common sense. The following are some useful advice:

1. Keep specifications constantly in mind and try to select as representative a sample as possible. Do not repeatedly select texts of a particular kind simply because they are readily available.
2. In order to get acceptable reliability, include as many passages as possible in a test, thereby giving candidates a good number of fresh starts. Considerations of practicality will inevitably impose constraints on this, especially where scanning or skimming is to be tested.
3. In order to test scanning, look for passages, which contain plenty of discrete pieces of information.
4. Choose tests which will interest candidates but which will not overexcite or disturb them.
5. Avoid texts made up of information that may be part of candidates' general knowledge. It may be difficult not to write items to which correct responses are available to some candidates without reading the passage.
6. Assuming that it is only reading ability that is being tested, do not choose texts that are too culturally laden.
7. Do not use texts that students have already read (or even close approximations to them).

Moreover, in most tests, especially tests of general proficiency, it is useful to include a variety of test types for reading comprehension in addition to the usual, more literary prose extracts: e.g. newspaper, instructions for using appliances and machinery, directory extracts, public notices, timetables and maps, advertisements, etc. The inclusion of such text types will not only provide a more realistic and reliable means of assessment but will also help to motivate students by demonstrating how the target language is used in real-life situations. Consequently, it becomes important that the actual presentation of the material should be as authentic as possible. When composing a test paper, the advice mentioned above should be followed.

4. Problems with Present Reading Tests

However, in present tests, there are various kinds of problems. Sometimes the students who take the test give the correct answers to the questions, but the results cannot show the ability that was originally wanted. In present tests, multiple-choice is one of the favorite forms of testing. It offers a useful way of testing reading comprehension. This is because it is regarded as an objective form. Generally, people think that it is not easy for the multiple-choice to be influenced by the scorers and the answers to multiple-choice questions are comparatively fixed. However, this is not really the case. Not all multiple-choice reading tests are necessarily good tests of reading comprehension. This kind of test is sometimes influenced by the subjectivity of the testers, and therefore, lacks validity, especially content validity.

The subjectivity of this form is because of the subjectivity of the test questions themselves. We usually think that multiple-choice questions have only one answer, but the problem arises from the composing of the questions. Sometimes the questions reflect the subjective idea of test constructors and it is quite not obvious. This subjectivity of the test questions influences the validity of the test paper.

First of all, test constructors take it for granted that the answers to the questions should and can only be found in the reading materials. Therefore, they sometimes include some general knowledge or some background knowledge in the choices but are not aware of it. For example:

- (1) According to the author, if a person is found guilty of a crime, _____.
- Society is to be held responsible
 - Modern civilization is responsible for it.
 - The criminal himself should bear the blame.
 - The standards of living should be improved.

(NON-MAJOR POST GRADUATE EXAM-91)

According to our general knowledge, if a person committed a crime, the criminal himself should be responsible for it. 'A, b and d' are not the correct answer.

Secondly, there would be a time when the so-called standard answer is not the only correct answer. From a certain point, the other answers are also correct. This is because the 'standard answer' is the view of the test constructor and hence influenced by their subjective idea, while the testees may have their own point of view or worldview and then choose another answer. For example:

- (2) Which of the following, according to the author, should play the leading role in the solution of environmental problems?
- Legislation and governmental intervention
 - The industry's understanding and support
 - The efforts of environmental health professionals
 - The cooperation of ecologists, environmental activists and conservationists

(CET6-99)

The standard answer of example (2) is 'c'. Because in the original reading text there is the sentence "However, it is those of us who are paid to make the decisions to develop, improve and enforce environmental standards, I submit, who must lead the charge." From the context we can see that the "us" here refers to the "environmental health professional". But it seems that we have no enough reason to believe that 'c' is the only correct answer, because not all the environmental health professional play a major role. To enforce environmental standards can also be done by government. Therefore, 'a' can also be a possible answer. There is a dilemma to decide which is the answer.

Thirdly, the questions themselves narrow the scope of the choice. This helps the testees to make the right choice without referring to the reading materials. For example:

- (3) Which is not mentioned in the passage as a disadvantage of the use of nuclear energy?
- Accidents which have caused great loss of money
 - Direct threat to the environment
 - Technological advances
 - The possibility of high costs

(CET4-95)

The word 'disadvantage' suggests the negative characteristic of the answer. Obviously 'a, b and d' are all the disadvantages and the only answer is 'c'.

The above examples only show how multiple-choice reading tests are influenced by testers' subjective idea. Of course there are other problems with reading test and they cannot all be included in this paper. It is important to be aware of these problems and measures should be taken to avoid these problems. Only in this way can we strengthen the validity of reading test and hence the validity of the whole test.

References

- Heaton, J.B. (2000). *Writing English Language Tests*. Beijing: Foreign Language Teaching and Research Press.
- Hughes, Arthur. (2002). *Testing For Language Teachers*. Beijing: Foreign Language Teaching and Research Press.
- McNamara, Tim. (2003). *Language Testing*. Shanghai: Shanghai Foreign Language Education Press.