



Analyzing Content Validity in Designing Assessment for Learning

Abduraxmonova Sevara Ne'matulla qizi

MA 1st-year student of Uzbekistan state world Languages University

Abstract: *This article analyzes content validity and its role in designing assessment for learning which is becoming crucial aspect in FLT. The author gives some definitions and evidences by some scholars who investigated the content validity of the different testing systems. Furthermore, connection between content validity and other characteristics of assessment is analyzed. After reviewing some literature, the reason why the testing system in FLT is still lacking reliability and validity is revealed. The findings of this article can be applied in other researchers which are intended to measure the importance of content validity in assessment design.*

Keywords: *assessment for learning, content validity, validity, content domain.*

Validity is one of the main characteristics of assessment and it refers to the extent that a test measures what it is supposed to measure. The Standards for Educational and Psychological Testing (2014) defines validity as the “degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (p. 11) In English language teaching process validity of the assessment is of paramount importance. To be more precise, assessment validity is more than just attributes of the tests, it is actually what the teachers are going to do with the results. In assessment for learning the outcomes of the assessment really matter. Here are some more definitions by well-known scholars: “A test is valid when it measures what it claims to measure.” (Garrett, 1964 p. 30). “A test is valid when it “measures what it ought to measure.” (Ebel, 1972, p. 436.) “Does the test measure what it is intended to measure? If it does, it is a valid test.” (Lado, 1975 p. 30) “A test cannot be a good test unless it is valid. The essence of validity means the accuracy with which a set of test scores measures what it claims to measure.” (Abbott, 1992 p. 178.) “A test is said to be valid if it measures accurately what it is intended to measure” (Hughes, 1995 p. 22). For that reason, validity is becoming crucial sphere in FLT and a number of scholars have done some researches on it. There are three more subdivisions in validity:

1. Content validity
2. Criterion validity
3. Construct validity

In this article, content validity will be to some extent investigated and clarified its role in providing valid assessment process. Content validity refers to the extent to which the items on a test are fairly representative of the entire domain the test seeks to measure (Neil J. Salkind, 2010). Content validity (Rossiter, 2008) is defined as “the degree to which elements of an assessment instrument are relevant to a representative of the targeted construct for a particular assessment purpose” (Haynes et al., 1995, p. 238). Content validity includes several aspects, e.g., the validity and representativeness of the definition of the construct, the clarity of the instructions, linguistic aspects of the items (e.g., content,

grammar), representativeness of the item pool, and the adequacy of the response format. The higher the content validity of a test, the more accurate is the measurement of the target construct (e.g., de Von et al., 2007). Although it is clear how important content validity is, it has been received little attention in assessment practice and research (e.g., Rossiter, 2008; Johnston et al., 2014). In many cases, test developers assume that content validity is represented by the theoretical definition of the construct, or they do not discuss content validity at all. At the same time, content validity is a necessary condition for other aspects of construct validity. A test or scale that does not really cover the content of the construct it intends to measure will not be related to other constructs or criteria in the way that would be expected for the respective construct. A test is said to have content validity if its contents constitute and represent sample of the language skills, structures, etc. with which it is meant to be concerned. A test is said to have face validity, if looks as if it, measures what it is supposed to measure. A valid test is the test which affords satisfactory evidence of the degree to which the students are actually reaching the desired objectives of teaching, these objectives being specifically stated in terms of tangible behavior. Content validity is an important notion for measurement of skills and knowledge that are frequently employed in evaluation studies. In this context, content validity usually refers to the degree to which a measure captures the program objective (or objectives) (Siddiek, 2018, p. 4) In foreign language teaching process, content validity mostly deals with the content domain as well as grammar, sentence structure and vocabulary. As it is stated, most test-makers do not pay attention to the content validity or just do not use proper methods to make the assessment more valid as well as reliable. Especially, in assessment for learning preparing tests and tasks that match to the requirements of the content validity makes the assessment more productive. First of all, content validity is not a statistical measurement, it actually deals with the quality of the assessment. A test or assessment that is valid should measure the aspects included in the objectives. In formative assessment, teachers examine the quality of the teaching process, pupils' or learners' weaknesses as well as achievements. Besides, mentors decide on which topic they should work further and put goals and objectives. This is another reason why it is important to provide content validity in designing assessment for learning. Every task, every test or exam should answer to the following question:

Does the exam cover all the content domains that need to be assessed?

Content validity is not only about the content or the covering the precise materials that should be assessed, but also it refers to the text or context, level of the learners, grammar and vocabulary, simplicity or complexity of the assessment. These very things later help to provide reliability and practicality, as well as positive washback of the assessment. For example, a classroom assessment should not have items or criteria that measure topics unrelated to the objectives of the course. Instructors can design a table of specifications for tests to ensure and communicate how the content of a course or unit is being measured (<https://ctl.yale.edu/AssessmentDesignValidity>).

So far the validity and content validity seemed to be neglected in the process of foreign language teaching and assessment in Uzbekistan. While analyzing English book 9 for 9th grade pupils in local schools, it can be seen that exercises that are given in the Project section do not coincide with the topics in the previous ones.

Grammar Exercise 3 Write these sentences in full, using the words given. Put the verbs into the correct tense.

e.g. If /I / not / see / the advert / I / not buy / this soap powder. If I hadn't seen the advert, I wouldn't have bought this soap powder.

1 If /I / not / see / the advert / I / not buy / this soap powder.

The exercise is too complicated for the pupils of 9th grade and the topic is not included in the syllabus itself. Why to assess if conditionals, if it is not intended to be taught?

From this example, it can be concluded that in designing any kind of assessment, whether formative or summative, content validity should be provided. Actually, it is the aspect which help to check the achievement of the objectives and the goals that have been put in the beginning of any course. In

order to illustrate content validity, testers investigate the degree to which a test is a representative sample of the content of whatever objectives or specifications the test is originally designed to measure. In that way testers or teachers can check if their assessment is valid or not.

It is not difficult to teach a pupil English language or something else, the matter is how successful the teaching process will be. The results of the teaching and its success can be only seen through proper assessment process. That is achieved by providing content validity in designing a test or a task. Consequently, it brings positive washback which motivates both teachers and learners towards the teaching process.

Reference list:

1. Ebel, Robert. (1972). *Essential Educational Measurements*. Prentice Hall, Inc. Englewood, Cliff. New York.
2. Garrett, Henry. (1959). *Testing for Teachers*. American Book Company. New York. *Glossary of Measurement Terms*. (2003).
3. Hughes, Arthur. (1995). *Testing for Language Teachers*. Seventh Printing. University of Cambridge. Bell & Bain, Ltd. Glasgow.
4. Lado, Robert. (1975). *Language Testing*. Wing Tasi Cheung Printing Company.
5. Siddiek, Ahmed Gumaa. (2009). *Assessment of Sudan School Certificate English Examination*. Published PHD thesis. [Online] Available: www.Lulu.org 2009. UK.
6. Rossiter, J. R. (2008). Content validity of measures of abstract constructs in management and organizational research. *Br. J. Manage.* 19, 380–388
7. Haynes, S. N., Richard, D. C. S., and Kubany, E. S. (1995). Content validity in psychological assessment: a functional approach to concepts and methods. *Psychol. Assess.* 7, 238–247.