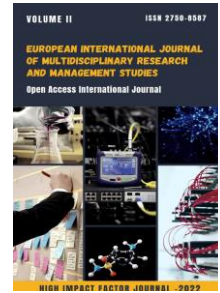


## EUROPEAN INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH AND MANAGEMENT STUDIES

VOLUME03 ISSUE02

DOI: <https://doi.org/10.55640/eijmrms-03-02-14>

Pages: 74-78



### THE VALIDITY OF THE TEST. DEVELOPING STRATEGIES TO IMPROVE VALIDITY

**Salixova Muhayyo Shokirovna**

*Senior Lecturer, Department of Information and Educational Technologies, TUIT, Muhammad al-Khwarizmi, Uzbekistan*

**Raxmonberdieva Gulzira Tolipjinovna**

*Assistant, Department of Information and Educational Technologies, TUIT, Muhammad al-Khwarizmi, Uzbekistan*

#### ABOUT ARTICLE

**Key words:** Validity, evidence, conclusions, domain, constructions, scientific evidence, types of validity.

**Received:** 03.02.2023

**Accepted:** 08.02.2023

**Published:** 13.02.2023

**Abstract:** In the educational process, the test is used to measure the subject area and the constructions contained in this topic. It is impossible to see the level of knowledge of the student, the level of formation of his abilities, abilities, skills: the teacher can see and measure them only indirectly (through questions and answers, observations, assignments).

Based on the conclusions made as a result of such a measurement, certain decisions can be made (for example, on awarding or not awarding a diploma, admission to study or employment) by any person or organization (for example, a teacher, an educational institution, an employer) to do or not to do, etc.).

#### INTRODUCTION

**Validity of the test.** In order for the decisions made to be reasonable, the conclusions drawn on the basis of the test must also be valid (justified).

Validity is the basic concept that determines the quality of a test. Validity based on the purpose of the test refers to the scientific and empirical validity of the conclusions drawn from the test results.

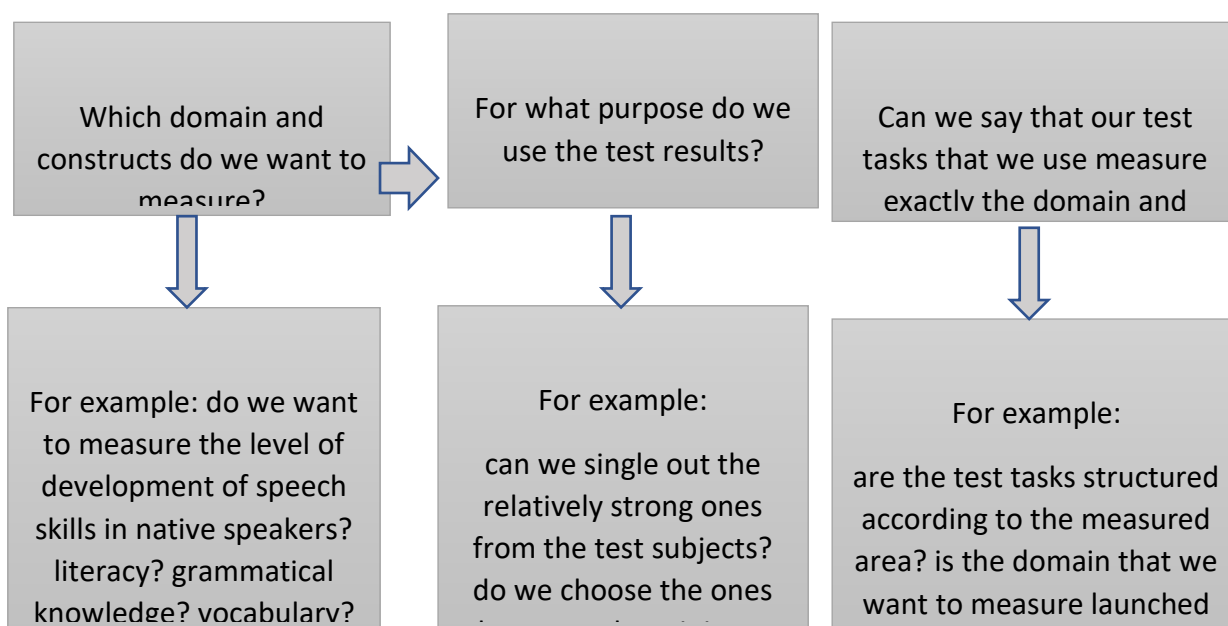
For example, how can we conclude that the student coped well with closed test tasks on the morphology of the native language? Can we say that he has mastered the rules of the morphological department of

linguistics well? Can we say that he can use the learned rules for the correct oral or verbal expression of his thoughts?

This is a controversial issue. Here is another example. Is it possible to conclude that a student who has coped well with test tasks related to reading and understanding texts of level B1 (lower average) in German also understands texts of level B2 (above average)?

Before any test, the persons or organizations responsible for conducting the test should identify several aspects related to the test:

- What area (domain) and constructions do we want to measure? What do these domains and constructs include?
- How are the test results used?
- What is the scientific evidence that the test tasks that we want to use measure exactly the area and constructs that we want to measure?



### Aspects that need to be determined before testing

For example, applicants who are not native English speakers have applied to an English-speaking university. Before admitting applicants, the university admissions committee should assess to what extent they are ready to listen to lectures in English, work with scientific literature, write term papers, communicate with professors and teachers.

To do this, the university will have to provide applicants with a test evaluating their ability to listen and read scientific texts in English, verbally and in writing to express their thoughts on scientific topics, as well as establish minimum requirements for the language level of applicants, and select those who meet these minimum requirements.

The results of such a test allow the university to make valid and reasonable conclusions about which applicant is ready to teach in English. If a university gives applicants a test of knowledge of English grammar, such a test will allow us to conclude about the level of proficiency of applicants in the rules of English grammar, but not about the ability of applicants to understand scientific texts or express their opinions on scientific topics in English.

Another example: if an educational institution that trains future linguists wants to assess the level of grammatical knowledge that students have acquired after completing an English grammar course, then it would be most appropriate to conduct a test for them aimed at checking the grammar of the language.

So, to what extent the conclusions drawn as a result of the audit are valid (justified) depends on the purpose of the audit, and a check that allows you to draw reliable conclusions in one situation may not give such an opportunity in another situation.

Evidence and conclusions of validity. The purpose of the test plays a key role in determining the validity of the conclusions made as a result of the test. If the same test is used for different purposes, to obtain different conclusions and make different decisions, it is necessary to make sure that the conclusions drawn as a result of each application of the test are reasonable.

As we saw above, the test is usually used to make a decision, that is, according to the test results, students (pupils, students), educators (teachers, educational institution) are accepted and even important decisions for the education system are made. In order for these decisions to be correct, the conclusions drawn as a result of the audit must be reliable.

In order to be able to say that the conclusions are correct, we need evidence. These arguments can be:

- evidence related to the content of the test;
- evidence related to the process of test subjects' performance of test tasks;
- evidence relating to the internal structure of the test;
- evidence based on the relationship of the test with other variables;
- evidence related to the consequences of the test.

This evidence is collected during the validation process.

Validation is the process of providing scientific evidence confirming the correctness of the conclusions made as a result of testing.

Despite the fact that validity is the most basic concept of pedagogical measurements (testology), there is no consensus among specialists on the definition of this concept.

Stages of test validity development. In the course of the development of science, definitions have also changed. The concept of validity can be divided into several stages:

- 1920-1950 – the emergence and development of the concept of "validity";
- 1950-1970 – fragmentation (disintegration) of the concept of "validity";
- 1970-2000 – "integration of the concept of "validity";
- from 2000 to the present day – modern interpretation of the concept of "Validity"

When the concept of "validity" first appeared, its definition was as follows: "validity is the degree to which a test measures the characteristic that it should measure." As can be seen from this definition, validity was initially considered as a property of the test, that is, the test itself could be valid or invalid, and this view persisted until the 1970s and 1980s.

By the 1950s, an idea of three types of validity had been formed:

- - validity criteria is the compliance of a test with a certain criterion, in turn, it is believed that there are two types of criteria validity:

- predictive validity is the ability of a test to predict results according to specified criteria. For example, entrance tests to a higher educational institution should predict to what extent applicants will be able to successfully master their chosen fields. If students accepted to study with high results of entrance tests also achieve high academic results, then the requirements of predictive validity are met.

- - comparative validity - compatibility of the test result with the results of other measurements according to a given criterion. For example, if we create a test that evaluates students' English level, we can determine its comparative validity with other existing tests that measure the same area (for example, IELTS or TOEFL exams). To do this, we test the test we created in a group of IELTS subjects, and if the results of the subjects in our test are similar to the results of the IELTS test, the requirements of comparative validity are met.

- - content validity - compatibility of the test with the content of the measured area.

- constructive validity - the test correctly measures the construct (characteristic) that should be measured.

An understanding has been formed that a certain type of validity condition must be met, depending on the purpose of the test and what type of test is used. For example, predictive validity for predictive tests, substantial validity for achievement tests, and constructive validity for psychological tests were considered sufficient.

By the 1970s and 1980s, Lee Cronbach (Lee Cronbach) and Samuel Messick (Samuel Messick), who are now considered classic representatives of the science of pedagogical measurements, put forward the opinion that this is not a characteristic of the test, but of the conclusions drawn on the basis of the test. In their opinion, the test itself cannot be valid or invalid, because the test is only a tool. This tool can be used in different situations and for different purposes. For example, a scalpel in the hands of a skilled surgeon can save someone's life, and in the hands of an ignorant one it can end someone's life, but the blame lies not on the scalpel, but on those who use it for other purposes.

By the 2000s, the concept of "types of validity" was abandoned, and instead the concept of the five aforementioned types of evidence of validity was officially adopted. This is also reflected in the "Standards of educational and psychological tests", which serve as an official document for organizations engaged in pedagogical measurements. Now, in order to prove the validity (validity) of the conclusions made on the basis of their tests, organizations involved in the creation and conduct of tests are required to provide five types of evidence during the validation process.

## REFERENCES

1. Komil Jalilov. Baholash nazariyasi (tesrologiya, pedagogic o'lchovlar) asoslari/ Toshkent. AKADEMNASHR. 2020
2. Cizek, G. J. and Bunch, M B. (2011) Standard Setting, Thousand Oaks: SAGE.
3. Green, A. (2014) Exploring Language Assessment and Testing: Language in Action, New York, NY: Routledge.
4. Воробьева, С. В. (2019) Современные средства оценивания результатов обучения в общеобразовательной школе. Изд. 2-е, пер. и доп. Москва: Юрайт.
5. Haladyna, T. and Rodriguez, M. (2013) Developing and Validating Test Items. London – New York: Routledge.