INNOVATIONS IN TECHNOLOGY AND SCIENCE EDUCATION
ISSN 2181-371X
https://sr-journals.org/index.php/ITSE

IMPROVING VALIDITY OF TESTING LANGUAGE SKILLS

Scientific adviser: ***Saparniyazov Rustem Shaniyazovich*** *PhD,*
Associate Professor, English Linguistics Department
Master student: ***Xushvaqtova Dilnora Boypo'lat qizi****,*
Second-year Master`s student in English Linguistics
Karakalpak State University named after Berdakh

Annotatsiya: Ushbu maqolada tilni baholashdagi validlik (asoslash) tushunchasi, uning turli xillari hamda test natijalarining aniqligini oshirish metodologiyalari tadqiq etiladi. Maqolada test natijalarining ma'noli talqin qilinishini ta'minlash uchun test konstruktlarini (tuzilmalarini) real hayotdagi kommunikativ kompetensiya (muloqot qobiliyati) bilan muvofiqlashtirish muhimligi alohida ta'kidlanadi.

Kalit so'zlar: Tilni baholash,Kommunikativ kompetensiya,test konstrukti,baholash metodologiyasi,test natijalari tahlili

Abstract This article explores the concept of validity in language assessment, its various types, and methodologies for enhancing the accuracy of test results. It emphasizes the importance of aligning test constructs with real-world communicative competence to ensure meaningful interpretations of test scores.

Keywords: Language assessment, communicative competence, test construct, assessment methodology, test results analysis.

Аннотация: В данной статье исследуется концепция валидности в языковом тестировании, её различные виды и методологии повышения точности результатов тестов. Особое внимание уделяется важности согласования тестовых конструктов с реальной коммуникативной компетенцией для обеспечения содержательной интерпретации полученных баллов.

Ключевые слова: Оценка языковых навыков, коммуникативная компетенция, тестовый конструкт, методология оценки, анализ результатов тестирования.

Introduction

In today's era of globalization, language learning and its assessment systems have become an integral part of the educational landscape. However, measuring language proficiency requires more than just administering a test; the focus must be on the extent to which the results are accurate and justifiable—namely, the *validity* of the assessment. Validity serves as the fundamental pillar of language testing, determining how precisely test tasks reflect a candidate's actual language skills (listening, reading, writing, and speaking). In many instances, traditional assessment methods are limited to measuring theoretical grammatical knowledge rather than a learner's practical communicative ability. This article aims to analyze the various types of validity in language testing and identify ways to enhance assessment quality by aligning test constructs with real-world communicative needs.

In the field of applied linguistics and educational measurement, the assessment of language proficiency has evolved from simple vocabulary recall to the evaluation of complex communicative competence. As global mobility and international communication intensify, the scores derived from language tests increasingly carry high-stakes consequences for academic admissions, professional licensing, and immigration. Consequently, the most critical question in modern assessment is not how much a student knows, but how accurately a test score reflects their real-world ability—a concept fundamentally known as validity.

Validity is no longer viewed as a static property of a test, but as a "unified framework" of evidence that supports the appropriateness and usefulness of test-based inferences [1;55]. Despite its importance, a significant gap persists between theoretical constructs and practical test design. Traditional assessments often rely on "discrete-point" testing, focusing on isolated grammatical rules or lexical items. Such approaches frequently result in *construct under-representation* (failing to test the full range of a skill) or *construct-irrelevant variance* (test scores being affected by factors other than language ability, such as test-taking anxiety or background knowledge). The shift toward the Common European Framework of Reference for Languages (CEFR) and task-based pedagogy demands a re-evaluation of how validity is established. To ensure that test results allow for meaningful interpretations, the test constructs must be rigorously aligned with authentic, real-world communicative tasks. This article provides an in-depth exploration of validity's multifaceted nature, examines the methodological challenges in contemporary language testing, and proposes a strategic framework for improving the accuracy and fairness of language assessments in diverse educational contexts.

The evolution of validity theory in language assessment has undergone a significant transformation, moving from a fragmented view to a more holistic, unified framework. Historically, Lado [2;57] laid the groundwork for structuralist testing, focusing on discrete language points. However, this approach was later criticized for its lack of focus on the interactional nature of language. The most influential shift occurred with the work of Bachman [3;69], who introduced the "Communicative Language Ability" (CLA) model. Bachman argued that validity is not just about the content of the test, but how well it reflects the complex interaction between a learner's language competence, strategic competence, and the psycho-physiological mechanisms. Following this, Messick [5;70] revolutionized the field by defining validity as a unified concept, where social consequences and the ethics of test use (consequential validity) became central to the validation process. Furthermore, Weir [7;90] emphasized a "socio-cognitive" approach, suggesting that validation must account for both the internal mental processes of the test-taker and the external social context of the language task. These theoretical developments form the basis for modern attempts to align test constructs with real-world communicative needs.

TYPES OF VALIDITY

Validity is not a monolithic concept; rather, it is established through various types of evidence. In the context of language testing, three primary types are essential:

- Content Validity: This refers to the extent to which a test represents the specific domain of language it is intended to cover. For example, if a test is designed to measure "Academic Writing," its content should include tasks like synthesizing information or summarizing lectures, rather than creative storytelling.
- Construct Validity: This is the most crucial type, as it examines whether the test accurately reflects the underlying theoretical construct (e.g., "communicative competence"). It ensures that the test scores are not influenced by construct-irrelevant factors, such as background knowledge or digital literacy in computer-based tests.
- Criterion-Related Validity: This involves comparing test scores with an external criterion. This can be concurrent (comparing results with an established test like IELTS) or predictive (how well the test scores predict future performance in a real-world setting, such as a university course) [7,91].

Methodologies For Improving Validity Of Testing Language Skills

To bridge the gap between theoretical constructs and practical assessment, researchers and test developers employ several rigorous methodologies. The following strategies are essential for enhancing the validity of language tests:

1. **A Priori Validation (Test Specification):** Before creating test items, developers must draft a "Table of Specifications." This ensures a balance between different language skills and cognitive levels, preventing the over-representation of easy tasks or the omission of critical skills.

2. **Authenticity and Situational Context:** Validity is improved when test tasks mirror real-life language use. According to Bachman and Palmer [2,67], authenticity is the degree to which a test task's characteristics match those of a Target Language Use (TLU) domain. For instance, using real-world audio samples for listening tests instead of scripted, studio-recorded materials increases validity.
3. **The Role of Rubrics and Rater Training:** In subjective skills like Speaking and Writing, validity often depends on the scoring criteria. Developing descriptive, transparent rubrics and conducting "Rater Calibration" sessions ensures that scores reflect the student's actual ability rather than the examiner's subjective bias.
4. **A Posteriori Validation (Statistical Analysis):** After the test is administered, Item Response Theory (IRT) or Classical Test Theory (CTT) analysis is used to identify items that are either too ambiguous or fail to discriminate between high and low-ability learners.

Conclusion

In conclusion, the validity of language skills testing is not a fixed attribute of a test but a continuous process of gathering evidence to support the accuracy and appropriateness of score interpretations. This article has demonstrated that achieving high validity requires a deliberate alignment between theoretical test constructs and real-world communicative demands. As language assessment continues to evolve, it is clear that traditional, discrete-point testing is no longer sufficient to capture the multifaceted and dynamic nature of human communication.

To ensure meaningful, fair, and ethical assessment, test developers must prioritize task authenticity, rigorous rater training, and the integration of diverse language skills. Furthermore, the ongoing evaluation of "consequential validity" is essential to ensure a positive washback effect on classroom instruction and learner motivation. By adopting a socio-cognitive and evidence-based approach to validation, educators can develop assessment instruments that not only measure linguistic knowledge but also accurately reflect a learner's true capacity to function effectively in a globalized society

References:

1. Alderson, J. C. Assessing Reading. Cambridge University Press. 2000.p 55
2. Bachman, L. F., & Palmer, A. S. Language Assessment in Practice: 2010 Designing and Developing Useful Language Tests. Oxford University Press.pp 60-90.
3. Fulcher, G. Practical Language Testing. Hodder Education. 2010.pp 60-77
4. Hughes, A. Testing for Language Teachers. Cambridge University Press. 2003.pp 40-45
5. Messick, S. Validity. In R. L. Linn (Ed.), Educational Measurement. Macmillan. 1989.pp 67-70
6. Shohamy, E. The Power of Tests: A Critical Perspective on the Uses of Language Tests. Longman.2001. pp33-35
7. Weir, C. J. Language Testing and Validation: An Evidence-based Approach. Palgrave Macmillan.2008.pp89-90