



PRINCIPLES AND PRACTICES OF LANGUAGE TESTING

Abdulakhat A. Isakov

ESL/EFL Teacher,

Oriental University in Tashkent, Uzbekistan

Abstract: The evaluation and assessment of language proficiency represent a complex yet essential dimension of language education. While the process demands meticulous planning and implementation, it plays a crucial role in shaping learners' academic trajectories. Traditionally regarded as fundamental elements of language acquisition, testing and evaluation serve not only as tools for measuring linguistic progress but also as mechanisms for enhancing student performance. Through systematic feedback and well-structured assessment practices, educators can foster meaningful learning outcomes and support the overall development of language competencies.

Key words: Assessment, language proficiency, evaluation, feedback, effective testing, practicality, reliability, validity, authenticity, washback.

Introduction

The evaluation and testing of a student's language skills is a tough aspect of the teaching profession since it requires rigorous examination before any test or evaluation can be carried out. On the other hand, this method contributes positively to the academic development of learners. In addition, testing and assessment of language proficiency have traditionally been held in high esteem as being among the most vital and alluring components of language education and acquisition. This is due to the fact that these strategies can not only be used to quantify the progress of student learning through the application of appropriate testing, evaluation, and feedback but they may also be utilized to boost student achievement. It should suffice to mention that the process of learning and teaching relies heavily on various forms of evaluation and testing. On the other hand, the test has the potential to either motivate students to keep studying or cause learners to get disheartened or question their ability to learn a second language. As a result, the planning process for every test needs to be done meticulously while taking into mind the potential negative impacts on the students. According to Brown (2004), in order for a test to be considered "effective, appropriate, valuable, or to put it another way, a good test," it must fulfill a number of requirements (p. 25). The current article provides both a critique and a modification of a test that is currently used as a final assessment for university students in their third year of study.

Learner Assessment Profile

He is a student at the Tashkent State University of Oriental Studies. He is 25 years old and decided to participate in the implementation of the project. He is currently in the second year of his studies at this university's Faculty of English Language Philology. Because of his extensive knowledge of human psychology and the cultures of the world, he stands out among his fellow students. At this point in time, he is proficient in the languages of Uzbek, Russian, Arabic, and English. He has just completed all of the requirements to receive a national certificate and earned a score that is equivalent to B2 on the CEFR level. He is preparing to take the IELTS examination right now. His educational background is really impressive. He was born in an

Uzbek family and raised there. The fact that his grandfather had also been an educator served as a source of inspiration and encouragement for him to continue his education. He attended both private language study centers and a school that specialized in teaching foreign languages because he had a keen interest in picking up other languages from an early age. Later on, he developed an interest in studying a foreign language, specifically English, because being able to communicate in another language paves the way for new prospects. On the other hand, due to the fact that he was born and educated in a backward region of the country, the majority of the education that he received focused on conventional approaches to grammar. In order to get into university, all he did during his school years was memorize numerous test texts and analyze them grammatically. This allowed him to enter university. In spite of the fact that he was unable to develop any of his English skills other than the grammar that was discussed earlier, the program was adequate for the immediate results he needed, which was prepared for entry into a public university.

Educational Setting

The student was able to achieve a passing score on the DTM-administered national standardization examination. The exam included a total of thirty questions, all of which were multiple-choice and focused on grammar, reading, and vocabulary. Simultaneously, thirty questions pertaining to history as well as the Uzbek language and literature were incorporated into the examinations. At the moment, he is enrolled at the Tashkent State University of Oriental Studies in the Faculty of English Language Philology, where he is studying for his third year. Therefore, the institution was a venue for him to engage in the in-depth study of English using cutting-edge instructional strategies. They are exposed to a variety of strategies and procedures for the instruction of the English language as part of the course material. Reading, writing, integrated language skills, grammar, and vocabulary are some of the topics covered in Language Enhancement classes, which are offered to freshmen and sophomores. The institution also provides instruction in a variety of other fields, including Uzbek history, information technology, natural sciences, economics, and others. Beginning in the third year of study, students begin to take an increased number of classes that are concerned with linguistics and the teaching of foreign languages. Methodology, phonetics, English literature, and linguistics are only a few examples of these fields. The student, along with the other students, is required to pass a series of examinations throughout the course of his studies at the university. These include a diagnostic examination at the start of the academic year, as well as midterm and final examinations. Exams given at the midterm and final levels are structured differently. Students are required to take part in a variety of activities, including discussions, debates, role-playing, taking video advertisements, creating pamphlets, and, of course, completing paper-based assessments. I have had talks with his educators, and they have informed me that the student participates quite actively in these assignments. In the academic circles that he has created for students who share the same interests, he demonstrates both his initiative and his ability to organize things effectively.

Placement test

He took the most recent placement test available to him when he began a five-month IELTS preparation course at a private tutoring facility. The examination included reading, writing, grammatical multiple-choice questions, and speaking, and was aimed at determining whether or not he was qualified to participate in the program that he had chosen. The test's validity was reduced since it lacked listening comprehension questions. When it comes to standardized exam courses with a predetermined curriculum, the test should cover all aspects of the language in order to accurately assess the student's genuine level of language ability. This is very crucial to remember when designing the exam. This segment employed a shortened version of the IELTS

reading and writing test, which was rather easy. The necessary piece of writing was an essay with a word limit of 150. Reading 1 and Task 2 are examples of the easier and more difficult IELTS reading component exams, respectively. This is how the reading component is divided. Because he had studied on his own prior to enrolling in the course, he had a thorough understanding of grammar, making it easy for him to finish the work. He was questioned about personal matters during the speaking section of the role, and his responses were marked by reluctance and halting. Students who do not demonstrate sufficient understanding to pass the course are placed in "pre-IELTS" sessions. Students in these courses study a wide range of literary and intellectual works and have the opportunity to improve their speaking and writing skills. Because this is a private educational institution, candidates can sit for an unlimited number of examinations at any given time. If there is only one goal, the test does not need to be taken more than once. Throughout the course, the center's administration will hold monthly tests. These assessments will be quite similar to the degree exam, but they will be more demanding and will require listening and speaking ability. They were used as accomplishment assessments by students and teachers to determine how much their English skills had progressed over the course of the year. These exams, unlike the ones given at the start of the course, have been designed with reliability and practicality in mind.

Critique of existing assessment or test

They have enrolled in an IELTS preparation course and are studying there at the Navigate training center. Because of this, I decided to go with the word formation challenge that B2 Destination (Unit-25-28), (Original test- Appendix 1) had to offer. Test specifications you can see in Appendix 2. The assignment relates to a position that is currently open. The purpose of the test is to select the appropriate word construction for the provided vocabulary. The student's vocabulary will be put to the test. This activity is assigned ten minutes of my time. There are ten items total on this test. My decision to take on this challenge was based on the fact that the exam in item is the first authentic, reliable, positively washed, practical, and valid. After making sure everything is correct and finding out how they scored in the end, they evaluate the answers to determine where they went wrong.

Practicality. Brown (2004) suggests that in order for a test to be useful, it must satisfy a number of requirements. It is easy to check because of the available keys (Appendix 3). These criteria include cost, test design, administration, and testing time, as well as prior preparation for test administration and publication of results (40-page). It is safe to say that this test satisfies all of the prerequisites for its practicality because it has been thoroughly examined. To begin with, there is no significant financial barrier to getting tested. Second, given that there were word-finding exercises, the compilation process shouldn't have required too much time. Also, certain questions are easily found. Third, because the exam is taken paper-based, it is not overly complicated or time-consuming to administer the tests and publish the results using the classroom. This is the platform that the majority of public institutions utilize to teach and evaluate students in classrooms. As soon as the exam was over, the results were announced in a short time, where they were also presented with the correct answers.

Reliability. There was no possibility of a breakdown in rater reliability because the test was administered and moderated by the author of the chosen book itself rather than by the instructors. The only thing that could go wrong is if someone makes an error while trying to find the correct answers to the question, in which case they should circle the correct option down. The student who took the test stated that he did not notice any inaccuracies in the test's reliability when he was taking it. As a consequence, the evaluation of the test did not contain any subjective variables. The reliability of the student is another factor that is considered.

Validity. At this point, I am beginning to have serious reservations about the value of the student's final exam. According to Brown (2004), validity explores what the test is supposed to measure and how it should be measured in order to determine whether or not it is effective. Brown calls validity "the most involved criterion and possibly the most fundamental idea of effective testing." To begin with, although the test is highly realistic and reliable, it is unable to evaluate the goals that were supposed to be achieved by taking the course. According to the syllabus, by the end of the class, students should be able to read with comprehension, listen with comprehension, speak on target issues, and write about target topics. As a consequence of this, the examination should include not only questions focused on vocabulary but also on listening, speaking, writing, and reading, just as the students exercise these skills in the classroom. Because we need to acquire language without separating it from its context, which is a necessary component of reality, it would be great if vocabulary assessments were included as part of listening or reading assignments. Another factor that contributes to the test's lack of validity is the fact that it is only capable of gauging the students' actual levels of knowledge to a certain extent. Word recognition for fill-in-the-blank questions is based solely on matching existing terms. Only the students' capability to develop synonyms, meanings, and phrases that are pertinent to the context is tested; there are no questions that test the students' ability to recognize these things. The exam is not prepared in such a way that "the greatest pupils mildly confront rather than overpower the poor students." This is important for the test's consequential validity. If you do not know all of the words, you will not pass the test; in that case, you will fail.

Authenticity. According to Bachman and Palmer (1996), authenticity may be defined as "the degree to which the features of a specific language test task match the qualities of a language learning task." Therefore, the constructed exam needs to reflect on activities that occur in real life. Natural language is going to be a part of the jobs that you will do in real life situations, and the exam that is being discussed definitely has that component. However, each component is considered as separate rather than in the context of the whole, which makes it seem less real. In addition, despite the widespread use of multiple-choice questions, we sometimes come across them while we are attempting to use language for the purpose of having meaningful conversations. Instead, we might find ourselves in a situation where we have to explain something to someone else, give an alternative word when we forget the one we wanted to use, or know when and how to use a certain word or phrase.

Washback. The validity of the information being tested has a direct bearing on whether or not the washback that is being provided by the test is of any use. In this particular scenario, the use of washback will be of no benefit to either the students or the teachers because it is unable to evaluate the students' true language proficiency. With these statistics, teachers are able to forecast their students' success in other areas where it is not assured that they will be successful. As a result, it is not possible for it to provide pupils with comprehensive instructions on how teachers should teach and how students should learn. The test provides feedback in the form of the number of correct answers; if the result is high, the feedback inspires a sense of self-assurance; if the result is low, the feedback motivates the student to study. The exam also has a means of telling students about particular themes from the topics presented during the course that they should revisit later on.

Three strengths

Appropriate to the intellectual capacity of the test takers as well as their age category;

The items on the test are arranged in a manner that is both logically obvious and appropriate, which contributes to the clarity and relevance of the test's organization

The test is practical since it is simple to give and score, and it does not require a lot of time from either the students or the instructors to solve the tasks.

Three weaknesses

There are no detailed guidelines or illustrative examples provided for any of the tasks;

Because there is only one word offered and no other possibilities, the test is sufficiently challenging to evaluate the test-taker's genuine understanding and ability to recognize word formation.

Provided with time limitation: 8 minutes.

It is essential to provide examinees with a time limit within which they must do each task that is part of a test; this is especially true for tests that measure continual progress. When there is a time constraint, it is typically more difficult for students to finish the assignment. In addition, it is helpful for the people who are in charge of organizing the test since it gives them a clear window of time after finishing the assignment to grade the test. In the event that this is not the case, the test-taker being tested as well as the examiner will be required to finish the assignment for an unspecified amount of time. This is supported by the assertions made by Mousavi (2009), who writes that "... the amount of time required for construction and management to facilitate cost, performance evaluation, and reporting."

Sample answer is provided in order to be easier for the test-taker.

As a result of this improvement, the test-taker will have an easier time getting the exam started right away. I made things simpler for the test-taker who was going to take the test by including a sample test that was a prepared version of one of the tests. If a sample is provided with the test, the individual who is taking the test will have a better understanding of how to perform on the test. In this case, the test-taker has a good probability of achieving a satisfactory level of performance.

Changing structures of the test. Questions have been changed from filling the gaps with word formations to multiple-choice vocabulary.

This is done because the test-taker taking the test might have forgotten the different types of word formations that appear on this kind of test, in which case making this adjustment is not simple. Because students waste time trying to think of the correct word structure, I wanted to shift the test to a multiple-choice format for the vocabulary portion. In this case, the test-taker taking the test will not only have a greater probability of finding words, but they will also save time. The reading level of the exam takers calls for the use of straightforward sentence structure.

Learner's response

At first, I gave the student an original test that consisted of ten questions on the topics that had been discussed, specifically filling in the gaps with word formations, and I allowed him five minutes to do the test. While taking the exam, the student reported feeling under a lot of pressure and having a lot of trouble. The reasons behind this were that there was a lack of time, he was not provided with an example of the correct response, and it was difficult to convert words into their proper word-formation. Only four out of ten questions were able to be answered correctly by the student who took the test (Appendix 4). Because of this, I modified and the original test in order to simplify it and make it easier for the student to understand.

The taker of the test obviously had a considerably better overall performance as a result of the test being administered in a format that had been adjusted, and as a result, he felt both more confident and more at ease while completing the task. In addition, the display of the example of the correct answer is the first part of the test that has been modified. The student was able to do the task through the test without facing any difficulties as a result of this, but it did require him to give careful thought to the questions and care about how the test was answered. In the second variant, the test was redesigned to have multiple-choice questions rather than the previous fill-in-the-gaps with word-formation vocabulary. This was done to ensure that the individual taking the

exam would have a clearer understanding of the task at hand and be able to complete it without any confusion. The participant in the test regarded all of the answers as being equally meaningful because the scoring system was the same for all of them. And finally, the time restriction was one of the most important factors in determining how quickly the activity could be finished. This was because it had an impact not only on the student's pace but also on his psychology. I added three more minutes to the allotted time, bringing the total to eight minutes, which assisted him in checking and rechecking his answers. He had more focus and determination, and he was less hesitant to pick one of the options. The individual who took the test had a very favorable impression of the revised version of the test and was able to achieve an eight out of a possible nine-point score.

Conclusion: One of the most important parts of education is the assessment system. Students need to have accurate evaluations done on them in order for them to be able to grow and enhance their skills. It also means that it is extremely essential for instructors to have the ability to create and appropriately administer a variety of tests based on the requirements and understanding of the students in their classrooms. Throughout the process of completing this task, I was able to gain knowledge on how to effectively design and administer a test. It dawned on me that exams are not only tasks for checking one's grammar or vocabulary; rather, in order for them to be successful, they need to satisfy essential criteria such as practicality, reliability, and validity. My instructors at Webster University taught me how to adapt various tests to the abilities and requirements of the students I might eventually teach.

References

1. Aripova, G. T. (2021). LINGUISTIC APPROACH IN CREATING BILINGUAL MUSIC DICTIONARY. *Актуальные научные исследования в современном мире*, (1-6), 11-14.
2. M.R.Abdullayeva. FORMATION OF ALTERNATIVE VARIANTS IN THE PROCESS OF ARTISTIC TRANSLATION (ON THE EXAMPLE OF A WORK OF “SUCH IS LIFE”). (2025). *International Journal of Artificial Intelligence*, 5(04), 1287-1291.
3. Aripova, G. T. (2018). ADULTS TEACHING ENGLISH IN THE SPHERE OF ART AND CULTURE. In *Язык и культура* (pp. 72-81).
4. Brown, H. D. (2004). *Language Assessment: Principles and Classroom practices*. NY: Pearson Education.
5. Malcolm, M., & Steve, T.N. (2006). *Destination B2 Grammar & Vocabulary*. Macmillan Education.
6. Abdullayeva, M., & Bekmahammadova, F. (2025). INGLIZ TILINI O'RGANISHDA SUNIY INTELLEKTDAN FOYDALANISH. *PEDAGOGIK ISLOHOTLAR VA ULARNING YECHIMLARI*, 12(01), 136-137.
7. Mousavi, S. A. (2009). *An encyclopedic dictionary of a language testing* (4th Edition). Tehran: Rahmana Publications.
8. Brown, D. H. (2010). *Language assessment: Principle and Classroom Practices* (2nd Edition). White Plains, NY; Longman Pearson 9780138149314
9. Isakov, A. . (2025). O'ZBEK VA INGLIZ TILLARIDAGI UNLILARDA FONOLOGIK INTERFERENSIYA OMILLARI. *Инновационные исследования в современном мире: теория и практика*, 4(5), 17–20. извлечено от <https://in-academy.uz/index.php/zdit/article/view/44676>
10. Green, A. (2014). *Exploring Language Assessment and Testing*. NY. Routledge.
11. Hughes, A. (2003). *Testing grammar and vocabulary*. In *Testing for Language Teachers*. (Cambridge Language Teaching Library, pp. 172-185).