

English Tests Used for Admissions Decisions: An Opportunity for ESL Professionals

Misty Wilson

Follow this and additional works at: https://surface.syr.edu/englishusa_journal

Recommended Citation

Wilson, Misty () "English Tests Used for Admissions Decisions: An Opportunity for ESL Professionals," *EnglishUSA Journal*: Vol. 5: Iss. 1, Article 5.

Available at: https://surface.syr.edu/englishusa_journal/vol5/iss1/5

This Article is brought to you for free and open access by SURFACE at Syracuse University. It has been accepted for inclusion in EnglishUSA Journal by an authorized editor of SURFACE at Syracuse University. For more information, please contact surface@syr.edu.



International students studying in colleges and universities in the US are often asked to provide proof of English proficiency as part of the application process. Proficiency can be proven through various ways: completion of an English as a second language program, reported country of origin, proof of previous English-medium school attendance, or through English proficiency test scores. Not all English proficiency tests, however, are the same. Some measure a test taker's knowledge of the language while others measure a test taker's ability to use the language. Some tests measure all four skills (listening, reading, writing, and speaking) while others might only fully measure only one or two skills. Some were designed to be efficient while others were designed to accurately measure English proficiency. Which English proficiency tests to accept and what score to set is not a one-size-fits-all answer, and it could be argued that the testing environment has become even more complex with the number of new tests appearing on the market and the most recent move toward digitally based at-home tests, accelerated by COVID-19.

In the United States, decisions regarding which tests to accept and what score to set have increasingly moved to international enrollment managers who, among other university test users, may lack the language assessment literacy needed to make such decisions (Baker, 2016; Taylor, 2013). While language assessment experts are best placed to assist in determining whether a test is fit for purpose, that is whether it is appropriate for

use as an admission requirement, they are not often sought during the decision-making process. This presents an opportunity for ESL professionals to use their expertise to offer guidance on these decisions, and recent research suggests ESL professionals can be influential in an admissions officer's decision to accept or reject English proficiency tests (Wilson, 2021). In order to be a part of the decision-making process, however, it is important that ESL professionals have a foundational understanding of testing principles so they can effectively contribute to decisions regarding choosing tests and setting scores.

Testing Principles

Many ESL professionals have heard of the five general principles of language assessment that should be considered when designing an English test: practicality, reliability, validity, authenticity, and washback (Brown & Abeywickrama, 2010). My goal in this section is to introduce principles that may be used when evaluating an English test for admission purposes and provide enough information on each to aide in the decision-making process. There are five principles, seen in the Test Evaluation Model in Figure 1, that can serve as a springboard to evaluate language tests for use as an academic admission requirement: validity, reliability, security, practicality, and impact.

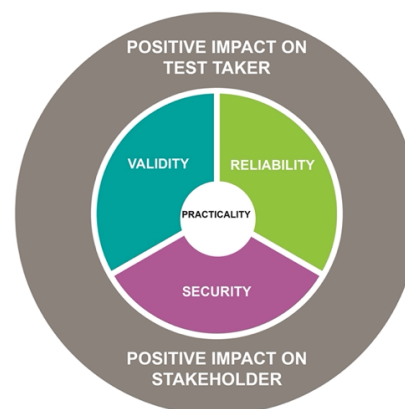


Fig.1 Test Evaluation Model

Validity

In general, test validity refers to a test's ability to measure what it says it will measure. A test may appear to be valid to an untrained eye (this is called face validity), but does it work? This question can really only be answered through validation studies, which provide evidence that the test measures what it says it will measure. When investigating the validity of a test, there are, however, a few elements you can look for. For a test to be used to measure English for admission into an academic program, the test should adequately measure all four language skills (listening, reading, writing, and speaking), and provide a clear score for each skill. This is important because it is not uncommon for students to have unbalanced skill levels, and some academic programs to have higher expectations for speaking skills than perhaps listening skills. To get a complete picture, you need a test that can provide information on each skill. Secondly, you should consider the task types included in the test. Using English for academic purposes is different from using language for interpersonal communication, so you need a test that assesses the test taker's ability to use the language for academic purposes. For example, does the test measure reading or writing in a way that replicates the type of reading and writing they will be expected to do in class?

Reliability

Test reliability refers to a test's ability to produce stable and consistent results. For example, if the same test taker takes the test on different days, will they receive the same score? Or will a test taker receive the same score if marked by different examiners? Another threat to test reliability could be cultural bias. For example, if the test taker is asked to describe an image of a man grilling, but has no previous experience with this

activity, their inability to describe the image may have nothing to do with their language proficiency. Some items to look for when examining test reliability include clear instructions at the beginning of the test, practice materials that familiarize test takers with all task types, and the testing of materials before being released into a live environment.

A good test needs both validity and reliability. Consider a scale that is 3lbs off. You may weigh yourself every day and get a consistent weight, but if the scale is off, then the measurement is not reliable. Maintaining validity and reliability requires knowledge and evidence-based assessment design, in addition to rigorous ongoing procedures to support the test in question.

Security

A valid and reliable test is nothing if it is not secure. Whether the test is given in-person or online, security measures must be in place to authenticate the identity of the test taker, assure material security and accountability, and to ensure the test results have not been influenced by malicious behavior. With the advent of online language tests being used for admission purposes, the question of security becomes even more important, and online tests which claim to be secure should be examined closely. There also needs to be an understanding that accepting these tests may come with inherent risks, and institutions should be prepared with policies in place that allow them to respond to various situations.

Practicality

If the test's validity, reliability, and security are the three main pieces of the test puzzle, then practicality is the central point connecting all three. Understanding how a more practical test might impact validity and reliability is important. A longer test, though less practical, could be more reliable if it

provides more opportunities for a test taker to showcase what they know. For example, consider the role of time in a writing task. An extended writing test would take longer for the test taker, but it would provide greater evidence of the range of vocabulary and grammar; in contrast, a shorter writing test would take less time but provide less information on the ability of a test taker.

Impact

Impact is often called washback in language assessment research, and it refers to the impact preparing for or taking the test has on the learner or the curriculum used to prepare the learner. In other words, does preparing for the test build proficiency and the targeted skills needed to succeed after the test? In questioning the impact of a test, take what is known about the test items or tasks, and consider what test preparation might look like. If a test taker was to prepare to do well on those task types, would they increase in proficiency and their ability to use the language? If so, then this would be positive washback. Preparation for the test should lead to better readiness for college and social experiences in an English language setting.

English proficiency tests can also have a profound impact on colleges and universities which trust that the test accepted is valid, reliable, and secure. If it isn't, there can be dire consequences to student retention and the academic achievement of students.

How to Choose Tests and Test Scores

When asked to weigh in on decisions regarding English proficiency tests used for admissions purposes, start by reviewing samples of the test and consider the five testing principles discussed above. Then review test taker sample responses and evaluator commentary. This is important so you can determine whether the test is

assessing language in a way that reflects how students will use it in their academic programs. Additionally, and perhaps most importantly, all tests should be evaluated through a critical and independent lens, so you should look for 3rd party peer-reviewed empirical studies that support the test's claims of validity, reliability, and security.

While some programs rely on institutional benchmarking, trusting that peer institutions have done their due diligence in evaluating a test, programs should not rely on this alone. It is not always clear by looking at a list of institutions that accept a test whether an institution is accepting a test as a supplemental piece of evidence or whether they have institutional safeguards, such as placement tests, language support services, or academic support programs as a safety net. Furthermore, when accepting a new test, especially one that has only recently been widely accepted and is lacking a substantial amount of peer-reviewed validation studies, an institution should be cognizant of the risks inherent in that decision and set up policies that allow for retesting of students or placement into an ESL program if needed.

Setting Scores

Similarly, when setting scores for tests, colleges and universities need to go beyond choosing a score based on what peer institutions have set and even beyond concordances touted by test creators. Score setting is the process of establishing a cutoff score for an assigned purpose such as admission into a graduate program. Setting scores too high may lead to rejecting qualified applicants while setting scores too low may lead to accepting students who do not have the needed English proficiency to succeed.

Concordances may be used to equate two different tests that measure the same construct; however, several studies have

shown that linking two tests that have different constructs (e.g., measuring speaking ability through the reading of a sentence vs. an interview task) is a flawed methodology (e.g., Jones & Xu, 2020; Yannakoudakis & Cummins, 2015). Just because a concordance table has been created by a test provider does not mean a test is comparable to another one or that the scores from the two tests can be used interchangeably.

Though there are several methods for determining what score should be set, these decisions should only be made with a clear understanding of what a test score represents in terms of what a student is able to do at that level, the academic rigor of the institution, and the language support available to students. Once a score has been chosen, qualitative and quantitative data should be regularly collected and examined to determine whether the score has been appropriately set.

Conclusion

Admitted students should have full access to the academic and social offerings of the program they are enrolled in, but this starts with ensuring the tool used to measure their language proficiency works. As many schools move to expand how a student can provide proof of English proficiency, there will be continued opportunities for ESL professionals to advocate for the needs of multilingual learners and contribute to the decisions regarding which tests to accept. Determining whether a new English test should be accepted as proof of English language proficiency for international students is a complicated task that should be undertaken with great care. There is too much at stake for both admitted students and colleges and universities to make quick, uninformed decisions.

References

Baker, B. A. (2016). Language assessment literacy as professional competence: The case of Canadian admissions decision makers. *Canadian Journal of Applied Linguistics*, 19(1), 63–83.

Brown, H.D. & Abeywickrama, P. (2010). *Language assessment, principles and classroom practices* (2nd ed.). White Plains, NY: Pearson Education, Inc.

Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412.

Wilson, Misty E. (2021). When professional roles take center stage: how international enrollment managers make sense of their decision to accept Duolingo English Test. Northeastern University.
<http://hdl.handle.net/2047/D20406239>

Misty Wilson, Ed.D., currently uses her TESOL expertise to support teacher training, test taker resource development, and research initiatives at IELTS USA. Previously, she spent 15 years in higher education, initially teaching English and later directing a large ESL program.