Scientific
Research
Publishing

Can There Be an Objective Language Tests?

Mohammed Almalki¹, Ajwaad Aljohani²

¹Department of English Language, University of Jeddah, Jeddah, KSA
²English Language Institute, King Abdulaziz University, Jeddah, KSA
Email: 2100683@uj.edu.sa

Abstract

This paper critically reviews the notion of objectivity in language assessment. There is a tendency in language testing and assessment narratives to view objectivity and subjectivity from the perspective of testing instruments and items, ignoring the broader context. Therefore, this paper re-reads the objectivity of language tests from a broader perspective. This paper recaps the factors that could impact language objectivity; namely, 1) factors related to assessment tools, which include validity and reliability parameters, items, 2) factors related to teachers to include un-biasness, fairness, and preventing discrimination of learners on the basis of race, ethnicity, or physical disability, or even the level of learning aptitude, and 3) organizational accessibility which can affect objectivity since students' cultures should be considered and they should have an equal opportunity to get access to learning tools and assessment resources. This paper recommends exploring the perception of ELT teachers and assessment designers about the factors and constraints of achieving the highest degree of objectivity.

Keywords

Objectivity, Reliability, Subjectivity, Language Testing, Validity

1. Introduction

It is a central requirement of good language assessment to be objective. As far as defining objectivity is concerned, there is no clear-cut concept. An objective assessment involves avoiding personal viewpoints and judgments when applying tests. Subjective tests, on the other hand, are those in which personal judgment intervenes with the assessment measures. Additionally, specific language areas are tested with objective tests, and overall proficiency is determined by subjective tests (Shaban, 2014). With this intention, objective and subjective assessments should be clearly understood by educators. Thus, the process of language

assessment ranges from objectivity to subjectivity depending on underlying factors that influence the degree of objectivity. This objectivity-subjectivity continuum can be minimized or maximized in light of the extent of alignment to language assessment measures. Objectivity and subjectivity have been associated with both the terms testing and assessment. According to Sharma (2020), testing is a large part of teaching which measures learner's knowledge for the purpose of determining the developmental stages of the learning process; as for assessment, it is an ongoing part of the learning and teaching process in which learning is not only monitored, but it is the degree of student performances which match their abilities. Therefore, it could be said that language testing/assessment without objectivity does not describe the effectiveness of the teaching-learning process.

From a psychological perspective, Bloom's taxonomy of learning domains, Objective-based tests also evaluate the entire human development in three domains: cognitive, affective, and psychomotor domains. As the word itself connotes, they are based on a particular objective of teaching and evaluating. They provide proper direction and thus streamline the whole process of evaluation.

Generally, the objectivity of language test is viewed from two perspectives: to evaluate language learner's progress towards achieving specific learning goal(s), or to measure language learning without reference to a specific course which is assessing language learner's proficiency (Fulcher, 2010; Giraldo, 2019). Thus, the objectivity of language test should meet these goals: facilitating the learning process and executing accountability (Archer, 2017). Therefore, it could be said that test objectivity should serve as part of the planning, development, and implementation process in educational assessment (James, 2010).

The objectivity of language test is generally discussed upon the following levels: 1) Scoring-instrument objectivity is related to scoring and grading concerns, including validity and reliability of the test. 2) Teacher-related objectivity refers to those issues pertaining to the teacher ability to be fair and unbiased. 3) Administrative or organizational objectivity that discusses other issues related to the accessibility of learner to tools and resources that help test performance.

2. Statement of Problem

As mentioned earlier, a language test objectivity has been evaluated from different dimensions. For better understanding, this paper evaluates lists of different factors influencing the objectivity of language test and to see the degree of objectivity-subjectivity affected by the factors. Specifically, this paper attempts to provide a critical review of factors which might have an influence on the objectivity of language tests. Especially that tests in the language domain differ from those in the other areas of content; thus, the type of language tests used by language educators should be carefully considered (Brown & Hudson, 1998) to meet the test objectives. In addition, the field of language assessment in general might need further consideration on the combination of factors influencing objectivity.

3. Literature Review

3.1. Instruments: Related Objectivity

If we explore the three levels that categorize the objectivity in language assessment, the first line the teachers are familiar with is the concepts of validity and reliability in relation to traditional forms of assessments. Both concepts (validity and reliability) are essential in the context of the classroom and in the context of high-stake testing for all learners. In other words, the objectivity of a measuring instrument is meant for the degree to which equally competent users get the same results. Accordingly, this presupposes subjective test instrument could minimize the level of assessment objectivity.

3.1.1. Validity

Validity corresponds to the extent to which inferences can be correctly drawn from the results of a particular assessment tool, process, and individuals assessed (Fraenkel et al., 2012). It measures what it proposes to measure; and offers useful, meaningful information about the person's abilities. Before 1989, validity was considered as the capacity of an instrument to assess what it was supposed to assess and nothing else (Brown & Abeywickrama, 2010; Lado, 1961). Right after 1989, Messick's (1989) perception of validity replaced this old perspective and is now highly embraced: The interpretations made of scores in assessment should be clear and substantially justified; if this is the case, there is relative validity in score interpretations. For interpretations to be valid, naturally, assessments need to activate students' language ability as the primary construct (Bachman & Palmer, 2010). Choosing a specific Scoring instrument is closely related to validity and thus the degree of test instrument objectivity.

3.1.2. Reliability

Teachers often know that reliability refers to the extent to which a particular assessment (i.e., test) generates consistent results. Billing (1973) notes that "in objective tests, scoring is objective and reliable: the name objective implies that everyone who scores the response to the item will arrive at the same mark". Hence, the reliability of the measure is enhanced, especially if it is based on analytic and specific rubrics (Jonsson & Svingby, 2007). Another necessary strength of an objective test is its adequacy at sampling the subject matter and instructional objectives of the course on which the test is based and designed. The larger number of questions set enables an adequate sampling of the student's actual performance. Therefore, the large number of questions would increase the reliability and the validity of the test. Besides, the knowledge that samples are a wide variety of questions encourages students to read wide and broad. Objective tests are basically made up of several items that have only one answer, for instance, missing words, incomplete sentences, true/false statements, multiple choices questions, matching pairs of statements and responses and alike. Such tests are called objective tests because the items that compose the test instrument must have precisely predetermined correct responses. Consequently, no matter what

form of it takes educational objective it assesses.

3.1.3. Questions Form and Objectivity

By looking at the most significant questions/items underpinning the test objectivity and allowing limited subjective judgment made by rater, we have to differentiate these types upon the probability or potentiality to orientate the teacher in designing objective assessment.

1) Selected-Response Items

Are composed of questions to which there is a single most suitable answer. They usually are referred to as objective tests (Suskie, 2018). Some of the most commonly used selected-response tests cover multiple-choice, fill-in-the-blank, true-false, and matching of statements. Selected-response items are easier to score, quickly answered, and enable the test designer to cover a broader range of curriculum in a short time. Yet, it constrains students to a single appropriate answer, encourages learners to learn by recognition and easy to guess the correct answer.

2) Constructed-Response Items

Are deemed as the less objective form of questions. They include a question like a short answer essay. Even though they allow students to demonstrate a complex, in-depth understanding, less likelihood of guessing the correct answer motivates students to learn and organize the information, principles, and application of knowledge. At the same time, it is more time-consuming to score and more time-consuming to answer.

3) Utilizing Grading Rubrics (Holistic vs. Analytic)

Grading rubrics are practical and efficient tools that allow for objective and consistent assessment of a range of performances, assignments, and activities. Rubrics of language assessment can assist in defining rater expectations scoring standard and show students how to meet those descriptors or standards, making them responsible for their performance by following easy-to-follow rubrics. The feedback that learners receive through a grading rubric can help them to improve their performance on subsequent assessments. Rubrics can help to rationalize scores when students inquire about the teacher's approach to assessment. Rubrics as well allow for more consistency and Objectivity in the grading process for teachers who instruct the same course. The rubrics can be categorized into two types; holistic and analytic. White (1985) suggested that holistic scoring draws attention to the strength limits of learner performances rather than the weakness points. Whereas Weigle (2002) mentioned that analytic scoring rubrics provide more information about a test taker's performance than the single score of a holistic rating, furthermore, it demonstrates a brief profile of the rated areas of language ability.

3.2. Teacher-Related Objectivity

3.2.1. (Fairness and Unbiasness)

Since the 70s, there have been great paces in educational assessment develop-

ment practices to ensure assessment measures are as fair as possible to the largest possible population of students. Fairness in testing is fundamental because it is self-evident; hence all students deserve an equal opportunity to demonstrate what they actually understand, know and can do. The teacher should consider the ethical issue in assessment; the teacher can be biased in evaluating the student from several aspects; culture, age, nationality, ethnicity, and gender. Therefore, the evaluator must be well qualified, have the knowledge and ethics necessary to give an objective assessment based on the best assessment practices. The teacher has to adhere to the assessment code of ethics that ban bias of any form and the assessment should support the learning process (Seden & Svaricek, 2018).

An educational assessment sometimes demonstrates cultural sensitivity respects the diversity of learners` ethnic, religious, cultural backgrounds which should not play a role in assessment (Fulcher, 2010). The test has to address content that a student from any part of the country and from any socio-economic class would have access to understanding and equal opportunity to apply the skill.

Cultural sensitivity is more about including content, contexts, and scenarios relevant to people from all sorts of different backgrounds. Thus, language assessments are culturally inclusive when accommodating to all learners' cultural backgrounds.

3.2.2. Discrimination of Learner Level and Intelligence

An objective test should discriminate different levels of attainment. It should indicate differences in the students' performance in terms of cognitive ability and educational attainment. If the students are tested, and their results are the same, the test fails to discriminate the students' different abilities.

3.3. Accessibility: The Organizational Objectivity

Accessibility in educational assessment refers to the tools, educational resources, devices, facilities, and accommodations that are provided to learners so that students either can use them to practice or perform the test as the peers do, or have an equivalent assessment experience to achieve the principle of the objective language test. According to Kunnan (2000), the accessibility of assessment to test-taker involves many variables that include financial, geographical, personal, the familiarity of test equipment.

At the classroom level, raters are fully aware when issues of accessibility due to linguistic, physical, cognitive, or emotional capabilities arise. In a school context, there ought to be teams of support providers, including classroom and special education teachers, tutors, and social services personnel, who focus on guaranteeing that students possess equal access to the same educational opportunities or resources as their or classmates' peers.

Language assessment has to ground on guidelines for assessing Special physical disabilities or those who have cognitive needs. In addition, assessment should

be accessible to students for whom English is not the first language.

4. Discussion

As we have seen in the above-mentioned factors that affect the degree of objectivity, we can realize that the boundaries between objectivity and subjectivity in language assessment are not totally clear-cut. Priority can come first to the instruments designed to evaluate student and should be valid and reliable. In addition, it is important to highlight that tests made by teachers, especially if they are untrained might have an effect on the objectivity-subjectivity of the test. For example, in writing/essay-task tests, students with poor writing will be suffered from untrained teachers scoring, as essay-type tasks mainly lack objectivity (Educational Stuff). According to Bachman and Palmer (2010), writing/summary tasks are subjective tasks, the reader should be able to determine whether a text corresponds to the main idea(s) of the task in terms of organizing thoughts and structure before evaluating. Designing the assessment based on the in which one select-response questions are used, e.g., MCQs and reviewing the scoring through grading holistic rubrics, would maximize the extent of assessment objectivity. In this regards, Alderson (2000) identifies two methods of scoring to meet the objectivity of such assessments: objective and semi-objective. In objective tasks, there is only one correct answer, and they can therefore be objectively scored, such as multiple-choice questions, true/false questions, and matching exercises. As for semi-objective tasks, they also have limited answers, but there could be more than one correct answer. Despite having multiple answers, semi-objective tasks are still limited production, e.g., sentence completion. Though, it could be difficult to evaluate answers with constructive responses objectively, e.g., gap filling as sometimes one-word answer may be specified in an answer key regardless to acceptable alternative answers (University of Minnesota, 2019). In other words, the term objective indicates objectivity in the scoring of the tests. A test is objective for eliminating the scorer's personal opinion and bias judgment; thus, avoiding ethnic, cultural and religious bias, hence, this could ensure the validity of the test as fairness is highly associated with validity when it comes to scoring on knowledge or skills (Lantolf & Poehner, 2013). As well rater should consider the variation of the ability of examinees whether on intelligence, attainment or physical levels. For example, in designing language assessment, learners' age, styles, and needs should be considered as they improve the learning process (Giraldo, 2019). In addition, the objective of assessment should be related to direct influences in the teaching environment, such as content, length, scoring allocation and test weights (Al Hinai & Al Jardani, 2021). The recognition of the quality of objectivity in a test has been largely the responsibility of the organization that design the assessment (school, university, testing services etc.) for the development guidelines that regulate all issues to objective type tests, by considering three main components to make unbiased assessment which are: theoretical ideas for language assessment, professional design, and

school/university polices (Giraldo, 2019).

Based on all, the quality of assessment objectivity is largely based on the skills of those who construct and administer it and they should reflect and consider the learners' learning process, theoretical and technical considerations, e.g., test design to arrive at reliable, valid interpretations of language tests.

5. Conclusion and Recommendation

To sum up, the term objectivity in language assessment is a fuzzy concept, but we can follow the measures that correlate the assessment more closely to objective tools. In order for achieving an acceptable degree of objectivity, we have to have a balanced combination of all factors influencing objectivity. These factors can be divided into three; instrument-scoring objectivity, teacher objectivity, regulation or organizational objectivity. We could say that we have an objective assessment if that assessment was valid, reliable, fair, unbiased, and well-structured. For eliciting a deeper understanding of the objectivity in language assessment, further research might explore teachers/test designers' perception on the challenges and considerations they face to meet the objectivity of language tests.

Acknowledgements

Acknowledgements to all who facilitated getting access to research data.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Al Hinai, M. K., & Al Jardani, K. S. (2021). Washback in Language Testing: An Exploration with a Focus on a Specific EFL Context in Oman. *International Journal of English Linguistics, 11,* 68-74. https://doi.org/10.5539/ijel.v11n1p68

Alderson, C. J. (2000). *Assessing Reading*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732935

Archer, E. (2017). The Assessment Purpose Triangle: Balancing the Purposes of Educational Assessment. *Frontiers in Education, 2,* Article No. 41. https://doi.org/10.3389/feduc.2017.00041

Bachman, L., & Palmer, A. (2010). *Language Testing in Practice*. Oxford University Press.

Billing, D. E. (1973). Objective Tests in Tertiary Science Courses. In B. S. Furniss (Ed.), *Aims, Methods and Assessment in Advanced Science Education* (pp. 131-148). Heyden.

Brown, H. D., & Abeywickrama, P. (2010). *Language Assessment: Principles and Classroom Practices*. Pearson Education.

Brown, J. D., & Hudson, T. (1998). The Alternatives in Language Assessment. *TESOL Quarterly, 32,* 653-675. https://doi.org/10.2307/3587999 http://links.jstor.org/sici?sici=0039-8322%28199824%2932%3A4%3C653%3ATAILA%3E2.0.CO%3B2-3

Educational Stuff (n.d.). *Objectivity: A Characteristic of a Good Test*.

https://acadstuff.blogspot.com/2017/06/objectivity-characteristic-of-good-test.html

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to Design and Evaluate Research in Education*. McGraw-Hill.

Fulcher, G. (2010). *Practical Language Testing*. Hodder Education.

Giraldo, F. (2019). Designing Language Assessments in Context: Theoretical, Technical, and Institutional Considerations. *How, 26,* 123-143.
https://doi.org/10.19183/how.26.2.512

James, M. (2010). Educational Assessment: Overview. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (Vol. 3, pp. 161-171). Elsevier.
https://doi.org/10.1016/B978-0-08-044894-7.01713-9

Jonsson, A., & Svingby, G. (2007). The Use of Scoring Rubrics: Reliability, Validity and Educational Consequences. *Educational Research Review, 2,* 130-144.
https://www.sciencedirect.com/science/article/abs/pii/S1747938X07000188
https://doi.org/10.1016/j.edurev.2007.05.002

Kunnan, A. J. (2000). *Fairness and Validation in Language Assessment: Selected Papers from the 19th Language Testing Research Colloquium, Orlando, Florida* (Vol. 9). Cambridge University Press.

Lado, R. (1961). *Language Testing: The Construction and Use of Foreign Language Tests. A Teacher's Book*. McGraw-Hill.

Lantolf, J. P., & Poehner, M. E. (2013). The Unfairness of Equal Treatment: Objectivity in L2 Testing and Dynamic Assessment. *Educational Research and Evaluation, 19,* 141-157.
https://doi.org/10.1080/13803611.2013.767616

Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher, 18,* 5-11.
https://doi.org/10.3102/0013189X018002005

Seden, K., & Svaricek, R. (2018). Teacher Subjectivity Regarding Assessment: Exploring English as a Foreign Language Teachers' Conceptions of Assessment Theories That Influence Student Learning. *CEPS Journal, 8,* 119-139.
https://ojs.cepsj.si/index.php/cepsj/article/view/500
https://doi.org/10.26529/cepsj.500

Shaban, A. M. S. (2014). A Comparison between Objective and Subjective Tests. *Journal of the College of Languages (JCL), No. 30,* 44-52.
https://jcolang.uobaghdad.edu.iq/index.php/JCL/article/view/169

Sharma, M. R. (2020). Perceptions on Language Testing and Assessment: A Case Study of B. Ed Students' in Nepal. *Journal of Advances in Humanities and Social Sciences, 6,* 27-33. https://ideas.repec.org/a/apb/jahsss/2020p27-33.html
https://doi.org/10.20474/jahss-6.1.4

Suskie, L. (2018). *Assessing Student Learning: A Common Sense Guide*. John Wiley & Sons.

University of Minnesota (2019). *Objectivity and Subjectivity in Evaluation*. Center for advanced Research on Language Acquisition (CARLA).
https://carla.umn.edu/assessment/vac/improvement/p_1.html

Weigle, S. C. (2002). *Assessing Writing*. Cambridge University Press.
https://doi.org/10.1017/CBO9780511732997

White, E. M. (1985). *Teaching and Assessing Writing: Recent Advances in Understanding, Evaluating, and Improving Student Performance*. Jossey-Bass Publishers.