

Vol. IV, No. I (Spring 2019) | Page: 55 – 59

p-ISSN: 2663-3299

e-ISSN: 2663-3841

L-ISSN: 2663-3299

DOI: 10.31703/glr.2019(IV-I).07

URL: [http://dx.doi.org/10.31703/glr.2019\(IV-I\).07](http://dx.doi.org/10.31703/glr.2019(IV-I).07)

Achieving the Validity of Essay Questions in the Subject of English at B. A. Level Examinations

Abstract:

The writing skill is one of the important language skills. This skill is especially useful for students of higher classes. In the Pakistani institutes of education, it is a matter of common observation that the writing skill gets more attention than the other language skills. Essay questions form an important and essential part of the assessment of the writing skill at the college level in Pakistan. The present study focuses on the Essay-type questions set in the subject of English (a compulsory subject) in the B.A. examinations for the students of government colleges. The examination is conducted by the Islamia University of Bahawalpur which is a public sector university. The study points out that the essay questions for the said examination lack validity and that is why they fail to measure accurately the intended objectives as well as skills. The study suggests ways to make the essay questions valid.

Key Words:

Assessment, Essay Questions, Validity, Writing Skill

Introduction

The concept of validity is one of the most important ideas that educationists utilize when they use tests. This is a concept that holds pivotal significance in educational measurement. We need to test the test's validity before we can accept the conclusions drawn from the test results. A test is not valid or invalid in general. It is valid or invalid for a specific standard, in a specific situation, with a specific group, for a specific interpretation, and for a specific purpose. It is not the test itself that has validity but rather the interpretations of the results. If a test does what it is intended to do for a particular group, then correct interpretations are possible. If a test is invalid, it is by definition biased. While bias may result from improper test administration, scoring, or interpretation, the main cause of bias is the test itself. A well-developed test should be reasonably valid for most test-takers. In the preparation of any test, two questions must always be considered:

- (i) Which skills exactly does it measure? and,
- (ii) How effectively does it measure them?

Efforts should be made to make each test as valid as possible. Each test is conducted to obtain a true measure of some specific skill/s. If it measures skills outside its scope of objectives, it cannot be said to be a valid test. For instance, the test item given below taken from an examination will be invalid if the examiners intend only to measure the writing skill:

Q- Discuss whether photography, in your opinion, is an art or a science? Such a test may not be valid because it requires specific kind of knowledge, therefore, it will favor some students who possess such knowledge.

Literature Review

In educational assessment, according to W.J.Popham (2006,p.94), validity “refers to the accuracy of the inferences or interpretations that are made based on students’ performances on measurement devices”. Because inaccurate inferences will result in unsound judgment, therefore validity is regarded as the most significant factor to judge the quality of an educational test. Validity involves the degree to which we can infer from a test score whether or not an individual possesses some hypothetical or expected ability, skill, or quality. P.W Airasian (2001,p.252) suggests that validity enables the teacher to make a correct decision about pupils’ learning. For G. Sax (1997,p.304), validity is “the extent to which measurements are useful in making decisions and providing explanations relevant to a given purpose”. Even if they are reliable, tests are useless if they are not valid for some designed objective. Each test must be validated to determine what skill it measures. L.R.Gay (1991,p.157) writes: “The most simplistic definition of validity is that it is the degree to which a test measures what it is supposed to measure.”^e

Such a test may not be valid because it requires a specific kind of knowledge, therefore, it will favor some students who possess such knowledge.

* Assistant Professor, Department of English, The Islamia University of Bahawalpur, Bahawalpur, Punjab, Pakistan.
Email: drasifkhan39@gmail.com

Types of Validity

Face Validity

It simply means the way a test looks to the students, test administrators, educators, etc. If a test looks as if it measures what it intends to measure, it is said to have face validity. Face validity is concerned with what teachers and students think of the test. Does it seem to be an appropriate way of assessing the students' performance, or does it appear very ordinary, or too difficult, or unrealistic? A test which intends to measure the speaking skill but which did not require the candidates to speak might lack face validity. An example might be an English reading comprehension test designed for American children who are given to adult foreign learners of English just because the two groups are thought to have a similar degree of proficiency in the language. Thus it is a test maker's duty always to keep face validity in mind though the other qualities of a good test should never be compromised only to satisfy public opinion. Luckily, it is quite possible to make tests that are technically sound and yet do not offend those who take them.

Content Validity

If a test is intended to measure mastery of a specific skill or the content of a particular course of study, it should be based upon a careful analysis of the skill or an outline of the course. For example, if a test is designed to measure students' mastery of English sentence structure, an analysis must first be made of syntax itself and decisions made on which aspects need to be tested and in what proportions. Therefore, a test maker should provide information about the specific materials or skills being tested, and the basis for their selection, while constructing a test. According to Arthur Hughes (1996,p.22): "A test is said to have content validity if its content constitutes a representative sample of the language skills, structures, etc. with which it is meant to be concerned."

A specification of the objectives or skills should be made at an early stage in the designing of the test. After having considered the purposes of the assessment, the content of a test should be drawn up in the form of a list called a 'content specification'. It is a list that shows the content of a test according to the purposes of the assessment. It should not be expected that every item in the specification list will appear in the test. There may be too many things and all of them may not appear in a single test. The content specification provides the test maker with the basis for making an authentic choice of elements to be included in the test. A test in which major skill/s are insufficiently reflected is not likely to be accurate.

Empirical Validity

Empirical or statistical validity can be achieved by comparing the test results with those of some independent outside standard as for example marks given at the end of a course or examiners' ratings. This particular study is concerned with determining whether the essay items constructed for the B.A students carry Face Validity and Content Validity or not and how to increase the validity of these essay items by constructing them in a proper manner.

The Main Educational Purposes of Language Tests

Determining the content validity of a test involves an awareness of the purposes of the test. Before planning a language test, the examiner must determine its purpose or function. According to purpose, there are three general types of tests which may be defined in the following manner:

A **General Proficiency Test** indicates what an individual is capable of doing now as a result of his overall learning experiences.

An **Aptitude Test** indicates an individual's potential for acquiring specific skills.

An **Achievement Test** indicates the extent to which an individual has mastered the specific skills acquired in a formal learning situation.

The teacher is in the best position to know which tests are appropriate for his students. The appropriateness of a test is largely determined by its purpose: why is a test needed at a particular stage? The concern in this study will entirely be with measures of achievement.

Achievement Test

An achievement test is concerned with long-term rather than short-term objectives. Decisions have to be made about what should be included in the test, and whether assessing one thing can be assumed to include another. For example, if the student has learned to write a business letter, can it be assumed that he can cope with personal letters too?

The Composition as a Language Testing Technique

Composition tests are a favorite form of measurement because they are an easy type of test to construct and measure certain high-level abilities better than do the objective techniques. Compositions have always been an important part of the writing skill. The students are required to write different kinds of compositions: essays, stories, letters, dialogues, and directed writing, etc. The

objective-type tests are universally considered to be the best and most authentic type of assessment because it is easier in an objective-type test to achieve a high degree of validity and reliability. The objective-type tests are easier to mark and they give reliable results. Furthermore, the objective type tests help a lot in focusing on the specific components of the writing skill. But the objective-type tests are more suitable for the students who are in an elementary stage of learning the foreign language. For the intermediate or advanced stages fluency, communicative competence, continuity and development of ideas, and expression are more important which can better be tested through composition rather than an objective-type test. That is why compositions are an integral part of tests at the intermediate and advanced levels. But there is the one big problem of making the composition tests valid and making the composition scores reliable. Enough work has been done to ensure that both validity and reliability are possible to achieve in the composition tests.

Discussion and Analysis

The present study is concerned with the assessment of the writing skill at the B.A level. English is taught as a compulsory subject at the B.A level. The students have to pass a hundred marks examination each in Part I and Part II. The focus in this study is on part II examination which contains a fifteen marks compulsory question on essay writing. The students are usually given five essay topics. They have to choose any one topic to write an essay on. The data for the present study consists of the essay topics taken from past examination papers. Given below is a list of the essay topics taken from the past B.A. Part II examinations under the Islamia University of Bahawalpur: 2014 examination: Energy crisis in Pakistan, Lack of good governance in Pakistan, Overpopulation, Technical education; 2015 examination: National resolve, My Hero, Our political leaders, Difference between government and private colleges; Literature and society, Education for women, Democracy, War against terrorism, My dreams; 2016 examination: Poverty causes failure, The role of men in our society, Ignorance and human sufferings; 2017 examination: Corruption, Energy crisis in Pakistan, Terrorism; 2018 examination: Higher education; 2019 examination: Information technology, How we can make Pakistan strong, Necessity of dams, Life in a college hostel.

The data for the present study were also collected through questionnaires to ascertain how far the teachers and the students considered the essay topics to be valid, interesting, and really motivating, whether the essay topics presented students with certain specific tasks. The participants in the present study included students and teachers. Two government colleges for boys in the city of Bahawalpur were selected; ten teachers and one hundred students were included in the sample. Both the students and teachers were part of the B.A. program. The questions designed for B.A (Part II) students were meant to know:

- i. Whether the students found the essay topics given to them in the examination really interesting and motivating to write on.
- ii. Whether the students reproduced the crammed material in the examination or they were encouraged to display their communication skills.

The questions designed for the teachers were intended to gauge their degree of satisfaction with the essay topics. The intent of the study is to know, whether, in teachers' opinion, the essay topics measured the students' communicative ability or they tested the students' ability to memorize ready-made material.

The results show that only 13% of the students often find the essay topics interesting and motivating while only 19% of teachers feel that the essay topics are often motivating (Figure 1). A clear majority of the students (60%) often reproduce the memorized material while writing the essay in the examination. Even a greater number of teachers (68%) think that the essay topics often test the ability of the students to memorize the crammed material (Figure 2).

A close analysis of the data leads to the following findings:

- i. It can easily be judged that the essay topics given to the students in the examinations are not very interesting. They hardly motivate the students to write on them. Topics like Literature and Life, Information Technology, The Role of Motor-way, etc. can hardly be motivating for many students.
- ii. Again the essay topics are often very general and quite broad in their scope of discussion. They seldom ask the students to write on specific themes having specific objectives. Just have a look at topics like My Hero, My Dreams, Corruption, Democracy, The Problem of World Peace, Friendship, etc. Topics like these do not focus on any particular theme. As a result, different students will choose to write on different areas of discussion.
- iii. Many topics require the students to have knowledge of specific fields or to have personal experience which many of the students may not be having, for example, Technical education, Information Technology, Difference between government and private colleges, Life in a college hostel, Literature and Society, Hajj, Students and Politics, Zakat, etc., need special kind of knowledge to write on them. Often the same topics reappear in subsequent examinations. Topics like Literature and Society and Democracy have appeared more than once in the examinations. All the problems mentioned above tend to make the tests less valid.

Data Obtained from Students' and Teachers' Questionnaires

Conclusion and Suggestions

The analysis of the past examination essay topics and the remarks of the students, as well as teachers (obtained through questionnaires), show that the essay topics lack validity, especially 'face validity'. Many of the essay topics are not interesting and motivating for the students. Sometimes, the topics need the students to have specific knowledge. This encourages the tendency among students to reproduce ready-made and crammed materials. This situation seriously hinders the students' ability to develop their communicative skills effectively.

As is clear from the findings, the essay topics need to be improved seriously. Conventional and stereotyped topics should be abolished. The type of essay topics that require the students to have specific kinds of knowledge should be avoided. The essay topics should be of general interest. They should motivate the students and produce in them a real need to express their ideas. They must be of immediate concern to the students and must pose some problem to the students whose solution then they may try to find. The topics must not be too general or broad in their range and scope of discussion rather they must point out clearly to specific areas of discussion. An essay once given in the examination must never be repeated again. While choosing the topics, all types of essays, i.e., descriptive, narrative, argumentative, factual, reflective, humorous, etc., should be given representation.

Here are some examples of how the essay topics can be made more specific in their themes. The specimens of more general topics are taken from past examinations:

General Topic: Hobbies

Specified

- i. Describe some hobbies which consume less time and are, at the same time, healthy;
- ii. Describe how a person's hobbies change as he grows older;
- iii. Tell how you were once influenced by someone else's hobbies or how you influenced someone else's hobbies;
- iv. In this age of materialism and competition, it is difficult to spare time for hobbies. How far do you agree?
- v. Were you ever unable to adopt some hobby because you could not afford it?

General Topic: Democracy

Specified

What, in your opinion, are the factors which have disturbed the growth of democracy in Pakistan? Discuss some of them.

General Topic: Students and Politics

Specified

- i. Is politics a useful activity for the students? Give reasons for your answer;
- ii. How can students contribute positively to politics? Discuss.
- iii. Discuss some of the disadvantages of students' involvement in politics.

The suggestions given above will help in designing essay topics that set specific tasks for the students in accordance with the learning objectives and thus may lead to the essay tests that are valid.

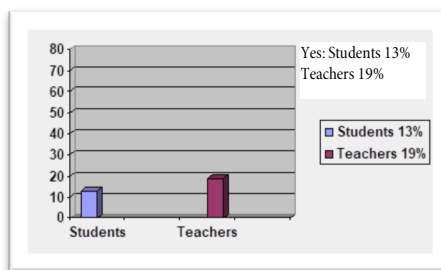


Figure 1. Do essay topics motivate?

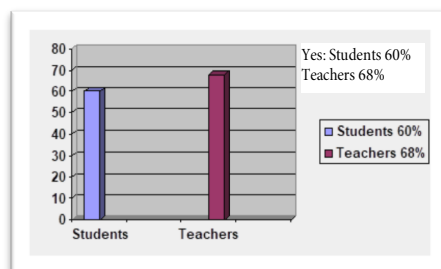


Figure 2. Do essay topics test the ability to reproduce memorized material?

References

- Airasian, P.W. (2001). *Classroom Assessment: Concepts and Applications*. 4th Ed. New York: McGraw-Hill.
- Gay, L.R. (1991). *Educational Evaluation And Measurement: Competencies for Analysis and Application*. 2nd Ed. New York: Macmillan Publishing Company.
- Hughes, A. (1996). *Testing for Language Teachers*. Great Britain: Cambridge UP.
- The Islamia University of Bahawal Pur, The. (2003-2007). Examination Question Papers, B. A., Part II, English.
- Popham, W.J. (2006). *Assessment for Educational Leaders*. USA: Pearson Education, Inc.
- Sax, G. (1997). *Principles of Educational and Psychological Measurement and Evaluation*. 4th Ed. USA: Wadsworth Publishing Company.