



RESEARCH PAPER

Critical Evaluation of two English Language Tests, Lahore Board of Intermediate and Secondary Education (BISE) and University of Cambridge's First Certificate English (FCE)

Dr. Humaira Irfan

Assistant Professor, Department of English, University of Education Lahore, Punjab, Pakistan

PAPER INFO	ABSTRACT
Received: April 19, 2020	This paper is a critical evaluation of two English language tests, Lahore Board of Intermediate and Secondary Education (BISE) and University of Cambridge's First Certificate English (FCE). These examinations are compared and contrasted and their strengths and weaknesses are evaluated in terms of specific objectives, specifications, scoring criteria, practicality, and wash back. The researcher used the qualitative methods such as, content analysis method and focus group interview to collect the data. The findings reveal that Lahore (BISE) assesses the memorization potential of the students rather than their creativity and critical thinking. It is found out that Lahore (BISE) has negative wash back effects on classroom teaching and learning. It is recommended that authentic tests should be constructed that would evaluate students' proficiency in English language.
Accepted: June 15, 2020	
Online: June 30, 2020	
Keywords: BISE, Critical Thinking, FCE, Memorization, Proficiency, Wash Back	
Corresponding Author: humaira.irfan@ue.edu.pk	

Introduction

Lahore Board of Intermediate and Secondary Education (BISE) is one of the largest examining bodies of Pakistan. It was established in 1954. In 2019, more than 181, 587 (BISE, 2019) students had taken English examination at Intermediate level, a crucial stage leading to admission in professional institutions such as, medicine, engineering, commerce etc. It is equivalent to international A Level examination. The First Certificate English (FCE) was set up in 1939 by the University of Cambridge. About 2 million people every year take Cambridge ESOL examination in 140 countries of the world. It is an upper-intermediate examination, at level B2 in accordance with the Common Framework of European languages. The pupils take the FCE Examination because they intend to apply for work or study abroad or to develop a career which requires English.

Literature Review

It is known that testing of students is an indispensable element of education across the globe. It strongly impacts both teachers and learners if not administered constructively (Davies, 1990). Schellekens (2007) and Hughes (2003) view that it is absurd to implement the rational educational decisions without having adequate information about the achievement of the students. Weir (2004) rightly perceives that testing involves testers and markers. He asserts that English teachers are in need of intensive training to construct effective tests and mark students' papers. The construction of English language tests and assessing student papers are highly crucial aspects of the programs and courses. The teachers have to be unbiased and fair with regards to assessment as their decisions about grades can have serious impact on students' future lives.

Hughes (2003) has classified English language tests. An English test conducted annually is described as an achievement test. Proficiency test is intended to determine candidates' potential in the usage of a language. Placement test is obligatory for placing the students at different stages of education programs and levels. In addition, a diagnostic test is deployed to diagnose the strengths and weaknesses of the learners in a language. In Pakistan, board and university examinations are known as achievement tests. On the other hand, IELTS, TOEFL and University of Cambridge ESOL examinations are known as proficiency tests. These tests are essential for students who are in quest of foreign education and satisfactory employment opportunities.

Puppin (2007) believes that traditional tests tend to be de-contextualized and inauthentic as these are heavily dependent on prescribed textbooks. These tests are based on subjective grading and correction criteria, predictably for that reason, lead to negative wash back. All annual board and university examinations in Pakistan are thought to be highly traditional and stereotypical. Contrastingly, performance tests are considered more reliable and valid as they consist of contextualized tasks and standardized scoring criteria.

It is noteworthy that Pakistani Board Examinations fail to measure the students' proficiency in English. National Education Policy (2001, p.38) states, '*the public examinations in Pakistan are invalid and unreliable as they encourage cramming*'. McNamara (1996) also reinforces that board examinations do not enhance students' learning positively. Hence, Bailey (1998) implies that English tests whether traditional or performance based must be designed thoughtfully to reflect authenticity and ensure the natural progress in a language. The tests are supposed to include real life tasks to assess learners' competence and sociolinguistic ability in English.

Siddiqui (2007, p.164) precisely expresses his view that students memorize the short stories, essays, dialogues, letters, poems etc. from the prescribed textbooks and model papers and reproduce these items in board examinations. They are encouraged to learn the questions by heart because it is the demand of examinations in Pakistan.

Board examinations do not assess students' critical thinking and creativity but scarcely any efforts are being carried out to improve the system. .

A critical evaluation of past papers of Lahore (BISE) discloses that the topics are repeatedly taken from prescribed textbooks. Siddiqui (2007, p.103) believes that textbooks have occupied a central position in Pakistani assessment system and classroom teaching is strongly dependent on assessment as it revolves around the textbooks. The amazing fact is that for the past 10 years, the syllabus at intermediate level has not been revised and updated (Past papers, 2020). The textbooks have model essays and letters which students memorize for tests. The questions come from the textbooks that do not require reflection, imagination and criticality on the part of the learners. More importantly, students have to remember the logical sequence of paragraphs or events as presented in the textbooks, otherwise, they might lose their score because the examiners keep in sight the model test papers or textbooks for marking the papers (Irfan, 2018).

Siddiqui (2007, p. 189) believes that Pakistani examination bodies lack creativity and critical thinking. Appropriately, Mustafa (2005) reflects that board examinations insist on rote learning that dictates specific classroom teaching. Testing's influence on teaching and learning is known as wash back (Hughes, 2003). Bachman and Palmer (1996, p.27) also reiterate that wash back is *'an aspect of impact on processes of learning and instruction.'* The wash back can be harmful or beneficial (Hughes, 2003). A test has positive wash back that assesses the needs of the learners but carries negative wash back if testing techniques are at variance with the objectives of the course. It is perceived that board examinations have negative wash back on classroom teaching and learning processes. Aptly, it is stated in National Education Policy (1992, p. 69), *'we are caught in a vicious circle; the cycle begins at a badly constructed syllabi and ends at a rag bag system called examination'*. Further, Siddiqui (2007, p.189) argues, *'implicit wash back effect is the teacher's own view of teaching which gets contaminated by the hanging sword of memory-gearred tests'*.

A good English test comprises multiple characteristics, such as, validity, reliability, authenticity, interactiveness, impact and practicality for fostering creativity and autonomous learning (Brown and Pickford, 2006; Bachman, 1990). Validity covers interpretation and meaningfulness of the achieved scores, reliability shows how consistently scores are achieved, authenticity reveals the co-relation between the test tasks and actual use of the target language, interactiveness includes traits of the test taker, impact is associated with the effects of tests on individuals and society and finally practicality means the test administration and availability of resources (Bachman and Palmer, 1996, p. 19-26).

It is a dilemma that Lahore BISE excludes these good qualities. In contrast, ESOL Examination has relevance to needs of learners, provides accurate and consistent assessment of each skill and relates assessment to the teaching curriculum to achieve the positive impact. In addition, it has authenticity, construct validity and interactiveness. The aim of examination is to assess pupils' comprehension, creativity

and communicative ability. The scoring criteria are explicitly explained. There are two marking schemes regarding writing, i.e. examiner's overall impression and requirements of a particular task. But Lahore (BISE) does not provide comprehensive scoring criteria. The marking of examiners is holistic and students complain about the impressionistic marking scheme (Mustafa, 2005). Bailey (1998) believes that score reporting should be detailed and diagnostic and not collapsed into one grade.

Material and Methods

The researcher invites 6 students to participate in the focus group interview. In this study, the participants are students of a large scale public university who are currently studying for degree in Masters in English. The age of all students ranges from 25 to 30. All students are in-service school teachers, teaching in different schools of Lahore district. The study is qualitative and has used content analysis method and recorded a focus group interview of 6 in-service female and male teachers (Table 1). The questions of focus group interview revolve about teachers' age, gender, teaching experience and their views about Board Examination. The purpose of a focus group interview is to supplement and reinforce the data collected through the content analysis of past papers of BISE and FCE.

Table 1
Participants of focus group interview

Teachers	Total
Female	3
Male	3
Total	6

The researcher has appropriately followed the research ethics. She explained the purpose of the study in plain language statement. She read out the plain language statement for the participants who voluntarily agreed to take part in the audio recording of focus group interview. The participants filled in the consent forms before participating in focus group interview. The anonymity of the participants was carefully ensured.

Results and Discussion

First of all, I would like to discuss, are these English language tests devised with specific purposes, learners' needs and their intended outcomes. As Hughes (2003) argues that the purpose of the test is to provide information about the achievement of learners without which logical execution of educational decisions cannot be taken. Lahore (BISE) is the final achievement test based on a syllabus which comprises literary texts such as poems, plays, essays and novels whereas the basic objectives of learning English at intermediate level are to acquire higher education and to get good jobs (Siddiqui, 2003.) It does not assess the true potential of the students as it indirectly measures their language ability. On the other hand, First Certificate English (FCE) is a proficiency test that aims at testing learners' communicative ability in English

language. FCE is designed in accordance with the needs of the learners to use English language in multiple sociolinguistic situations. It comprises diverse contexts, regular assessment of each language skill at the appropriate level and accuracy. Further, examination is related to teaching curriculum to produce and promote a positive learning experience.

Analyzing the specifications of Lahore (BISE) English test indicates that the content is dependent on prescribed textbooks. The operations in a paper involve the ability to choose suitable responses, answer short questions, sentence construction, translation and essay writing. Reading comprehension is assessed by short questions. The structure of the test reveals that it is divided into two sections with objective and subjective type questions, medium is paper and pencil and the total time for the test is three hours. The board examinations do not frequently change content and format. Lahore (BISE) does not include a reading comprehension passage, listening and speaking skills; and topics for essay writing are taken from their course books, thus encourages rote learning. It is perceived that such type of tests make use of de-contextualized test items that merely assess students' potential to produce the memorized answers. This testing undermines the quality of instruction in an education system.

On the contrary, the course materials of FCE reflect the content and format of the examination. It is divided in five parts; listening, reading, writing, speaking and the use of language. The tasks have interactiveness (involvement of test taker's individual characteristics), authenticity (degree of correspondence between given task and target language use) and variety such as, multiple choice questions, gapped text, matching, fill in the blanks and making words; writing ability is tested by various genres, such as, an essay, an article, a report, a review, a letter of application, an e-mail or a short story. FCE has constructed validity (pertains to meaningfulness and interpretations of test scores; Bachman & Palmer, 1996). The construct of reading skill is defined as '*candidates are expected to show understanding of specific information, text organization features, and tone and text structure*' (FCE Handbook for Teachers, 2008). Taylor (2006) believes, the focus of 'can-do' in Cambridge ESOL Examination is a legacy of the steady shift away from a focus on knowledge and form in language teaching and learning towards a focus on function and communication.

Unlike FCE, Lahore (BISE) does not provide comprehensive scoring criteria. It neither provides descriptors (criteria for correctness) for marking and scoring purposes nor defines the construct (ability) what the candidates are expected to do. Therefore, students complain about the impressionistic and subjective marking. The questions set in the Lahore (BISE) paper demand rote learning (Mustafa, 2005). McNamara (2000) says '*language testing is crucially dependent on definitions of the test construct.*' The examiners mark each question in the test to assign grades to candidates. Total marks are 100. A stands for excellent (80-100), B very good (70-79), C good (60-59), D satisfactory (50-49), E weak (49-40). Like the FCE examination, for objective type items a scoring key is constructed so that marking could be done by the clerical staff. The score of a candidate is norm-referenced (referring to other candidates' performance).

Bailey (1996) asserts that score reporting should be detailed and diagnostic and not collapsed into 1 score.

In the FCE, a candidate's overall grade is based on the total score gained by the candidate in all five papers. Regarding writing, for instance, candidates' answers are assessed with reference to two marking schemes: one based on examiner's overall impression, the other on the requirements of a particular task. The criteria for writing performance is not just accuracy including spelling and punctuation but also content, organization, cohesion, range of structures, vocabulary, register, format and effect on the target reader (Taylor, 2006). All the papers are equally weighted each contributing 40 marks to the examination's overall 200 marks. A is exceptional (80-100 marks), B is very good (75-79), C is good (60-74), D is borderline (55-59) and E is weak (54 or below) (FCE Handbook for teachers, 2020).

It is observed that Lahore (BISE) carries negative wash back impact as compared with Cambridge ESOL Examination. FCE testers undertake to explore the impact in an organized and systematic mode to ascertain that the test leads to productive classroom learning and teaching rather than merely a test preparation activity. FCE has validity, reliability, impact and practicality. According to National Education Policy (2001, p.38) the public examinations in Pakistan maintain focus on cramming, thus are invalid and unreliable. Further, learners and teachers are reluctant to accept any changes in the curriculum (Jenkins, 2006).

The regular updating of FCE has allowed the examination to keep pace with changes in language teaching and testing. About the examination, the feedback is collected from the students at the end of the examination. Survey questionnaires are sent to candidates, teachers, oral examiners and examination administrators. The collected data assists to develop examination specifications including the development of test construct and content, editing and trialing of draft task types and materials, assessment criteria and research into validity and reliability of material (for testers and examiners) and assessment procedures. Regarding public examinations, National Education Policy (2008-2012, p.39) states '*Pakistan shall make efforts to develop international level academic assessments and research and development cells will be established in each board to improve the system.*'

The participants are discontented with BISE examination (Table 2). Board Examinations have negative impact on classroom teaching and assessment.

Table 2
English Teachers' Perceptions of Lahore BISE

Teacher:1	<i>"Honestly speaking, Lahore Board Examination encourages cramming". "Lahore Board Examination does not test creativity and critical thinking".</i>
Teacher:2	<i>"It can't be called a proficiency test because questions are based on textbooks. The textbooks give even situations for writing letters and applications". "Pupils memorise essays and stories to pass the examination".</i>

Teacher:3 *“Lahore Board Examination does not assess students’ English competencies and*

Teacher:4 *proficiency”.*

Teacher:5 *“I agree with others that such type of examination kills students creativity and fail to develop their English language skills”.*

Teacher:6

Conclusion

In short, Lahore (BISE) could be an internationally competitive examination like University of Cambridge FCE if it takes pains to devise the tests based on the needs analysis of learners and includes authentic and interactive tasks to develop the communicative ability in language, comprehensive scoring criteria and beneficial wash back. It is believed that examination boards are capable of embracing positive changes in the tests.

References

- Board of Intermediate and Secondary Education (BISE), (2020). <http://www.BISE.com>
- Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L.B & Palmer, A. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Bailey, K.M. (1998). *Learning about language assessment: Dilemmas, decisions, and directions*. Boston: Heinle and Heinle.
- Brown, S. and Pickford, R. (2006). *Assessing skills and Practice*. London: Routledge.
- Cambridge ESOL Examinations. (2020). <http://www.cambridgeesol.org>
- Council of Europe. (2001). *Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press.
- Davies, A. (1990). *Principles of Language Testing*. Oxford: Blackwell.
- FCE Handbook for Teachers (2020). <http://www.cambridgeesol.org>
- Hughes, A. (2003). *Testing for Language Teachers. 2nd edition*. Cambridge: Cambridge University Press.
- Irfan, H. (2018). *Testing Creative Writing in Pakistan*. Newcastle: Cambridge Scholars Press.
- Jenkins, J. (2006). The spread of EIL: a testing time for testers' *ELT Journal* 60(1) 42-50
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Mustafa, S. (2005). The importance of an examining board'. <http://www.education@dawn.com>
- National Education Policy (2008-2012). Islamabad: Ministry of Education Press.
- National Education Policy (2001). Islamabad: Ministry of Education Press.
- National Education Policy (1992). Islamabad: Ministry of Education Press.
- Past English Papers of Intermediate (2020). Lahore: Urdu Bazar.
- Puppin, L. (2007). A Paradigm Shift: From Paper and Pencil Tests to Performance Based Assessment. *English Language Teaching Forum*, 45(4), 10-16.
- Schellekens, P. (2007). *The Oxford ESOL Handbook*. Oxford: Oxford University Press.

Siddiqui, S. (2003). Using Literature to teach Language <http://www.education@dawn.com>

Taylor, L. (2006). The Changing Landscape of English: implications for language assessment'. *ELT Journal*, 60(1) 51-60

Weir, C.J. (2004). *Language Testing and Validation: An Evidence-based Approach*. London: Palgrave MacMillan.