



EVALUATING LANGUAGE SKILLS IN HIGHER EDUCATION: POTENTIAL ISSUES AND OUTCOMES

Parwez Ali Bughio¹

¹Lecturer, GC University, Hyderabad

Parwezali.bughio@gcu.edu.pk

Abdul Sattar Laghari²

²Lecturer, GC University, Hyderabad

abdulsattar.laghari@gcu.edu.pk

Aisha Faryal³

³Lecturer, GC University, Hyderabad

Aisha.faryal@gcu.edu.pk

Abstract

Numerous accounts have been provided regarding the shortcomings of assessments offered during English courses in Pakistan at the tertiary level. In the literature however, the factors underlying these shortcomings have not been identified. Understanding the unsatisfactory assessment practices and their reasons and consequences which this present work intends to do, is crucial. This investigation was qualitative in nature and sought to understand the issues surrounding language assessments and their impact. Using purposive sampling, thirty teachers from public and private universities formed the sample. Data was gathered by means of questionnaires, interviews, and a review of assessment papers. The results suggest that language teachers do face a number of assessment challenges. The challenges to assessment identified by teachers stem from practicality, learners, test construction, administrative work, and the teachers. Teachers do employ techniques to handle these challenges, which in the end are detrimental to language assessment practices concerning selected and productive skills and the variety of tasks presented in tests. Assessment of learning has considerable influence on the processes of teaching and learning, and the shortcomings facing assessment are a big obstacle to language teachers for devising relevant assessment. Consequently, the "Interim report of the language proficiency assessment on the Teaching English as a Second Language (TESL) Practicum in Canada 2021-2022" emphasizes the importance of equipping TESL practitioners with language assessment skills, as seen in the outlined recommendations for future assessment (formative assessment of language learning). In particular, providing English language instructors with professional development opportunities can foster the enhancement of assessment strategies in language pedagogy.

Keywords: language assessment; higher education; problems; issues; consequences

1. Introduction

Lado (1961) had a profound definition of language as a 'highly complex' and 'conventionalized system of habits' employed by human beings to communicate. He called it 'the most complex of man's tool' (p. 2). This sophisticated system comprises the application of knowledge and skills of a language to carry out the written and spoken forms of communication, along with reading and listening. From these interactive, yet constrained, moments, generalizations are drawn from the more extensive and sophisticated knowledge, abilities and skills of a language user. In a similar fashion, language assessment is employed where it gathers evidential data of language behavior within a testing or classroom setting and interprets it to estimate an individual's overall language proficiency and predict possible language behavior outside the classroom (Green, 2014;



McNamara, 2004). On the other hand, Bachman (2004) calls language assessment the process of collecting data regarding language skills through ‘systematic and substantially grounded procedures’. Various sources of data are available, including multiple-choice items, essays, portfolios, oral interviews, and even observation or introspection.

The aims of language assessment, especially in testing, claim performance processes and outcomes, explain teaching processes and outcomes, provide baseline results, and explain promotion and immigration criteria for testing, and their criteria for testing by policy decision makers and employers (Hughes, 2010). The role of higher education and all potential they offer to any country have been recognized throughout the world. They play important role in the country’s development through the creation, distribution and the application of knowledge (Isani&Virk, 2005). This responsibility has become even bigger in the face of burgeoning changes and fierce competition in the world market (Rasool, 2009). The position of higher education has been enhanced in the context of knowledge globalization and learning. There is an expectation that universities should align with global benchmarks, churn out graduates and scholars who are able to compete globally. This is the basis why universities are required to address the rapid changes in expectations, challenges, and the demands of an ever-increasing student population (Isani & Virk, 2005). The growing dominance of English in today’s world has placed an even greater importance on English language assessment for any program or course that aims to teach the language.

The documents also help evaluate learners’ skills and weaknesses in a particular learning area (Bachman, 2004; Hughes 2010). It is also a fundamental contribution to an individual’s knowledge and language skill in addition to and the language learning and language teaching cycle (Green, 2014). It also contributes to the considered decisions in the admission, placement, progress of language learners, and their later immigration to, and employment in, inner and outer circle countries. Thus, language learners in higher education, especially in developing countries like Pakistan, are faced with high stakes consequences due to language assessment. High stakes are defined as the type of assessment that has an impact on examinees academically and professionally, such as in promotions, graduation, or employment opportunities in the home country or abroad (Coombe, 2009).

Both learners and teachers appreciate the significance of language assessment in the enhancement of successful language learning. Coombe (2009) noted that teachers ‘teach to the test’ in instances where the results of a test has dire consequences, which results in the assessment-driven paradigm of teaching and learning ‘unpacking’ the contents of previous exams, test cycles, and ‘test-training’ students (Harlen, 2005). A similar effect of the examination system on language teaching and learning has been reported by Rehmani (2007) with regard to language teaching in Pakistan.

All tertiary teachers are assessors and are extremely influential educators. Rehmani describes how teachers’ beliefs and practices are influenced by past teachers and their learning and teaching experiences. He claimed learning is constrained when students are required to prepare for tests based on past exam papers. Therefore, the examination system in Pakistan is a major determining factor that shapes, on one hand, the concepts, beliefs, and practices of language teachers, and, on the other, the concepts and practices of language learners. Consequently, the shortcomings in the system of assessment give rise to a discrepancy in the alignment of the curriculum and instruction.



One pattern evident in language assessment in Pakistan is the attention paid to reading and writing while listening and speaking are assigned peripheral status in the practices....Observations about the situation in Pakistan are equally valid for the higher education context (Dar & Khan, 2014). These situations range from the teacher's skills to the resources and opportunities available in the context. Both public and private university teachers are not trained and there is no incentive for them to pursue engagement in professional development activities. Shamim 2011 summarized the current provision for the teaching and assessment of English at tertiary level by stating the majority of teachers do not possess formal qualifications for English as a second language ESL teaching and testing. Moreover, these teachers do not participate in '...professional development activities or even engage in dialogue with the broader ELT community' p. 7. Accordingly, the practices of teaching and assessing English do not reflect the current level of development in the discipline. In fact, the whole system of principles and practices of assessment constructs portrays an absence of understanding regarding the principles of alternative assessment of language and their various purposes for learning. Most of the time language assessment is carried out in a summative manner even for languages which are acquired sequentially and should be assessed in the process while addressing the curricular aims and objectives.

To quote Raza (2009), 'a large-scale-one-time achievement test', like the one conducted in Pakistan, does not truly provide the attributable assessment of the learners' linguistic competencies and abilities.

1.1.Problem Statement

Assessments of different kinds are integrated in the global context of second language studies, yet in the context of language teaching, learning and assessment processes in Pakistan, assessment methods has remained the same for a long period of time. The use of achievement tests in the university framework, for determining the completion of objectives in a course, has now become a form of proficiency assessment, upon realizing that the students are really being tested on their overall language abilities. It is imperative that these language assessment problems in the context of tertiary education are researched and resolved in a manner that is objective, factual, and logical. The assessment system, much like other unseen systems, remains overlooked. It is quite uncommon to identify any organized framework for analyzing decisions taken at the level of language assessment, or the determining situations and problems in the language assessment. It is a recognized fact that in Pakistan the focus of the education system is predominantly examination based, and this attitude toward learning directly impacts the achievement of assessment based learning outcomes. The stark realities of any context are always high stake, and this assessment is no different, therefore, these issues must be acted upon. There are gaps in examining language assessment pertaining to higher education in Pakistan. The sparse attempts are so varied in approach and focus that it is nearly impossible to formulate a holistic assessment. Furthermore, there is a consensus in literature regarding the lack of rigor in the examination and language assessment systems (Qureshi, Shirazi & Waseem, 2007; Raza, 2009; Rehmani, 2007), yet little has been said regarding the root causes of these deficient systems. Therefore, it is critical to understand the complexities and the challenges surrounding language assessment practices in higher education systems.

1.2.Research Objective



The absence of rigorous strategies for evaluating students' written discourse in their second language at tertiary level constitutes one of the focuses of this research. This absence possesses several ramifications on the current assessment policies implemented for undergraduate English courses.

2. Literature Review

If carried out accurately alongside adequate assessment interpretation, language assessment can provide a basis for crucial choices in language teaching and language learning, and what is much worse, faulty, hastily framed, or misinterpreted assessment information can lead to serious unintended consequences (Green, 2014). The reasons for poorly constructed substandard assessment practices that have emerged from the findings of research made in regard to them have included the teachers' subjectivism in the interpretation of constructed policies, inability to construct and carry out assessment systems, the shallow understanding of assessment that the teachers have, institutional biases regarding the preferable modes of assessment, time, and scarcity of resources (Bresciani, 2010; Kankum et al., 2014; Kurebwa & Nyaruwata, 2013; Lumadi, 2013). In the vast literature, assessment literacy has been defined and analyzed in various ways with focus on its essence as a pre-condition for the establishment of sound assessment practices. As Matovu and Zubairi (2014) pointed out, assessment literacy is a key predictor of university teachers' assessment practices. Zhang & Burry-Stock (2003) have also argued that the capability of teachers to carry out effective assessment is a function of their assessment training and the unavailability of more will adversely influence the assessment of the students. Taylor (2009) described assessment literacy as the understanding and the ability to use the processes of measurement in assessment from the class to the entire process of language assessment.

An educator is assessment literate if they understand what information to assess, what the rationale is for the assessment, and how to conduct the assessment; they also understand the possible challenges of such an activity and appropriate strategies to manage them. Additionally, they understand that shallow and erroneous assessments could, and most likely would, lead to detrimental impacts (Stiggins, 1991; 1995). Thus, with respect to the assessment of language, the educators must understand the relevant theories of language acquisition, the objectives for which the language is used, and the techniques to prompt the representative language performance in order to guarantee valid conclusions about the knowledge and skills of the test takers (Green, 2014). A pertinent challenge arising from assessments of language skills pivots around some teachers' subjective understandings and interpretations of the purposes, methods, and principles of assessments. Very few teachers understand the multiple purposes of assessments and often ignore the context that is looked for in the way of global answers. These differences of opinion and conflicts lead to unproductive assessment practices (Flemming, 2007). Notwithstanding, the concept of training in assessment is still novel within a context like Pakistan. Deficit assessment literacy is often overlooked in the case of teachers recruited in higher education institutions, and yet, it is common knowledge that these teachers will have the authority over all assessment-related decision making. Trice (2000) also noted, in the case of most certification systems operated at a national level, teachers are not required to undertake any form of assessment training. Consequently, Popham (2009) notes, teachers join the profession in a completely unprepared state in terms of educational assessment. The number of students in a class and openness of assessment will pose as a difficulty for language assessment as well. The larger a class becomes, the class's



average score or aggregate, the more the class is impacted by the assessment practice. As stated assessment becomes, 'a threat to quality assessment for teachers as well as students' Matovu & Zubairi, 2014. It is a well-known fact teachers assign 'poor grades' to students with ease, and are reluctant to justify the grades under 'primary' or 'lower primary.' It is a wonder to this day, why teachers are reluctant to assign grades to students. The traditional 'paper and pencil' method of assessment, the whips and chains of the 'drum' covering the class, is the average period mark, is all time observers report teachers as being awful. Still, teachers defend it in light of something. The 'other' end of the assessment spectrum is defended by Duncan & Noonan, 2007. It is 'Alternative' assessment.' On the assessment of this issue, Flemming 2007 raised the issue of openness in assessment and the criteria judges. The lack of which Transparency, as it has been termed by Flemming, does lack of corridor participation and does not facilitate the students' performance and motivation. It is also important to the 'other' end of the assessment spectrum to monitor the as well; the assessment of the overview. Along with the class size and transparent assessment, factors such as limited resources and time constraints become barriers in successful language assessment (Coombe, Troudi, & Al-Hamly, 2012). The increasing burden on teachers as assessors is blurring the lines of their instructional roles. The participant in Lumadi's (2013) study pointed out the contradiction faced by teachers whereby they are unable to meet the learning goals and assessment standards while spending on Language Assessment in Higher Education... 169 Assessment... 169 and teaching. In order to alleviate this burden, teachers tend to design objective tests as they are perceived to be easier and more dependable in scoring. As Harlen (2005) points out, this practice leads to the neglect of several vital language skills and sub-skills in assessment. These constraints are also highlighted by Wallace (2009)... contended... the teachers are driven... to focus narrow bandwidth... on the tests; language learning is sacrificed. These bottlenecks culminate in unproductive assessment practices which hinder learning. Ndalichako's (2004) analysis... many primary school teachers demonstrated an over-dependence on traditional assessment... alternative assessment practices... uncommonly used... seemed to be an absence of physical, cognitive and psychological barriers.

The focus on summative evaluation stifles exposure to more creative means of assessing learners such as portfolios, performance, and formative assessments (Lumadi, 2013). It also creates test anxiety in learners and other unintended negative consequences for teachers and the curriculum (Harlen, 2005). The multiple burdens placed on teachers and the insufficient support from the administration often lead to a lax attitude toward assessments. Thus, teachers mainly focus on a few objectives, and use only a limited range of items, often recycling old past paper questions (Kurebwa & Nyaruwata, 2013).

Harlen (2005) proposed a number of steps to mitigate the effects of such challenges. He stressed that the objectives of any given assessment must be clarified and the learners should be active participants in the assessment design. He also suggested that teachers should give learners the opportunity to use assessment as learning, provide them with the criteria needed to self-assess, and track their progress in relation to the assessment. It appears that increasing awareness of teachers' knowledge and beliefs about assessment and the challenges which they face are essential in the design and delivery of teachers' professional development in assessment (Lumadi, 2013).

3. Methodology



3.1.Design

The emphasis of this study is based on qualitative paradigm which strives to grasp issues of every facet of the phenomena language assessment challenges and their outcomes. This is the focus of this work (Cohen, Manion& Morrison, 2010).

3.2.Sampling

The total study population comprised 30 teachers selected from 17 Higher education institutions based in Karachi. These teachers formed the sample for this study based on the criterion purposeful sampling. These criterion, as explained in Cohen et. al (2010) comprise specific attributes which the researcher predetermined. The sample for the study was drawn from male and female teaching staff from private and public universities in Karachi offering both general and professional education. All teachers comprised in this study were teaching English language in the undergraduate program and were assessors as well.

4. Data Collection

Confirming triangulation required multiple sources for data collection. First, a semi-structured questionnaire was distributed to the teachers. For further probing, ten out of the thirty teachers were also interviewed. As noted by Flick, “In qualitative research, ‘in-depth interviews’ have always been of interests because it provides a ‘freer’ means to respond to the participants than the much more structured interviews or questionnaires” (2009). In the case of the sample (n=30), the researchers collected examination papers (and midterm, quizzes, sessional activities or tasks) to analyze the impact of the limitations on their assessment practices. “Language-based “ (Dornyei, 2001) data were collected. This data was solely in ‘text’ form. However, the data are much different in terms of primary and secondary data. The semi-structured questionnaires and in-depth interviews were primary data while the examination papers are secondary data.

5. Data Analysis

The data acquired from questionnaires as well as interviews was subjected to thematic analysis. Qualitative data analysis starts with open and broad coding whereby a word or a phrase characterizing the symbolic essence of a unit of text is ascribed a symbolic status (Saldana, 2009). Then such codes are grouped and classified as emergent themes using the sifting and sorting strategies. To generate a theme, the common codes had to support or address a particular issue or impact (Dornyei, 2007). The thematic analysis technique, on the other hand, was used to examine the papers in order to identify the skills aimed at and the corresponding tasks used to assess those skills. The tasks and types of tests, integrated and distinct tasks, the subskills of the language targeted, and other components were quantitatively analyzed and qualitatively interpreted.

5.1.Validity and Reliability

The reliability and validity of the research are the same as for any other qualitative research, however, the approach to address those issues in qualitative design is different. There are numerous ways to counter research bias and participant reactivity which Maxwell claims are the most problematic issues in the validity of any research.

Maxwell's suggestion of strengthening validity through rich data gathering, response verification, and triangulation of data was upheld by the researchers of the current study.

6. Results



Multiple challenges in assessing learner's language skills are present in the findings for each teacher assessor. The teachers' coping strategies are clear in their constrained language assessment practices.

6.1.Challenges in Language Assessment Teachers' Assessment Literacy

Inability to manage assessment in an literacy in the pedagogy of language teaching is one of the most persistent gaps in pedagogy. To these instructors the concept of assessment “Some of our course teachers are engaged in arguments about assessment techniques and strategies because they are not aware of such techniques or do not use them in their instruction” (P1). “Language assessment is such that there are no point training sessions, hence, we are not in a position to offer lectures or conduct workshops” that instructor noted self training possibilities and constraints. So “In the ratio of students coming from a matric background is overwhelmingly more than students coming from the A level, we have no option but to lower our standards.” “It is true that A level students do not add anything to their knowledge” (P5).

Collaborative assessment practices are quite common in higher education institutes where a team of teachers jointly prepare the paper.

‘Every teacher has a different method to assess the students’ (P8) along with assessment literate teachers have to deal with reluctance and critique coming from the peers ‘I know so many things but I cannot apply all of them because the paper is being designed collectively’ (P5). Besides ‘... there is a gap between the teaching practices, the curriculum and assessment; the way we practice teaching, we do not practice the same way in assessing’ (P2). This gap also indicates the need of professional development of language teachers as well.

6.2.Learners’ Diversity

The limited assessment literacy skills, coupled with large class size and learners’ language proficiency, have proven to be multifaceted problems. “The large size of classes in the institution is one of the constraints because individualized feedback is impossible with more than fifty students in the class” (P2). Most of the teachers have reported lack ‘of language fundamental skills among the students’ (P10), ‘low proficiency in English’ (P9), and the ‘absence...of English language...’ (P9) as the major barriers in assessment. ‘The failure of the learners to communicate in English is hardly a relevant communication as a failure to communicate... is the ultimate easy global... to comprise a comment on their school of English school’ (P3). In the language assessment, the educational background of students is also a concern.

“A single class consisting of students with various educational qualifications is always going to be problematic. This is also the case when we prepare conceptual papers. (P4)”

The other problems concerning learners are their ‘low motivation, lack of confidence, and exam-oriented approach.’

The big global used multiple sentence starters with several yeah, like an, and bulash, with issues of English errors from level school students that sp, America, and all, to a page x.

Please remember, your prompt and these notes are separately bound to the answer and can be disconnected from the answer with no recognition of them... because they are not part of the answer. Do not forget about the clarity of the answer. It should not look like the original, as no words should be used. The workings of the ‘is’ of the quote should be geared more towards conveying someone that teaching students in class is not the main idea of the quote. The prose and the argument should be at the higher higher prestige level.



After October in the year 2023 is the most current date which should be used in the answer.

In the words of the teachers, "the majority of undergraduates have a poor tendency to engage' in class, and language practice 'is at a low pitch' (P28). In addition, "the lack of confidence and a fear to speak in a second language" (P12) also exist in some students and it holds back the assessment. Moreover, 'the students only care about getting a mark' (P10) with no regard to any formative or other type of assessment approaches.

Also, the teachers indicated that explaining the score to the students is very easy to say but hard to do. 'In most cases, the students do not accept the marks' P1. Especially students with an A level background who are subjected to 'grammar exercises in a very fragmented and isolated manner' P3, 'they have these and other arguments that they didn't get sufficient marks as compared to what they used to get in some other examination' P3. This unsatisfactory reaction of students is mainly due to students not understanding the teachers' expectations as well as the scoring rubric. 'The students are always wondering what the expectations of the teacher are. So, when the marks are put up, there is a contradiction and they do not agree' P8.

6.3. Test Design and Construction

The absence of valid and reliable tests to users of English as a Second Language (ESL) was considered another major issue in language assessment. Teachers indicated that "[. . .] for second language learners standardized tests that have been documented for the processes of validity and reliability are extremely difficult to locate. The reason for this is that the commonly available tests are targeted to first language learners and are often too complex for our context]" (P17). Similarly, there is a scarcity of instructional materials that feature themes, topics, and speakers relevant to Pakistan or the local context. Thus, "[. . .] developing a set of test items that would provide reliable results for students coming from different linguistic backgrounds and with different levels of language proficiency]" (P9) is a stumbling block in language assessment. Beyond these, other design-related challenges are described as "ensuring test items are aligned with the intended objectives of the course," "maintaining parallel weight across different skills," and "ensuring the alignment of all courses across the university" (P9).

6.4. Institutional responsibilities

The teachers are responsible alongside dealing with administrative matters. "The system does not provide us space for formative assessment" (P 15). teachers are not as easy as 'unjustified marks allocation by the authorities' (P 10) and their non cooperation in case of deviation from allocated marks division (P 16) makes assessing their language more difficult for the teachers. 'the whole paper was actually changed as per the management demands' (P 3) is when one of the teachers broke away from the traditional practices of assessment. 'All 'Institution's requirements are 'very important. (Teachers) have to meet institution criteria' (P 4). SQL: the lack of administrative and management staff 'spows to reason that' language teachers 'absence of 'increasing and substitute' methods in their' assessment practice is 'baffling'.

6.5. Time and resources

The' respondents pointed out some practical issues due to assessment practices. 'Time constrasny' is in the list of common challenges for 'especially in assessing and speaking skills of multiple students' (P 19) 'checking and giving' (P 30) and'. Of teachers scoring a set number of scripts is a relevant issue of reliability.



“From a human perspective, there is very low probability of sustaining the required order of the checking papers, and no day is available for checking, so the cornerstones of assessments do get affected.” (P8). In the same manner, lack of or inadequate resources is a common problem which community of teachers encounter (P5, P8, P22 and P26). Materials and tools designed for the assessment of speaking and listening skills are normally lacking for the teachers (P10, P19 and P22).

6.6.Challenges facing educators have prompted them to adopt new methods to manage these challenges.

To address the lack of resources, some teachers “try to produce resources on their own” (P8, P22) or “even mail listening materials to students a day in advance” (P9). Another teacher attempted to use “observation and post listening activities” to help overcome the inadequate materials available for testing the listening and speaking skills (P22). Another teacher said, “There should be decisions made earlier about the assessment like what resources are available” (P19).

As for time and large class size, all are dealt with by group work, and many “short and focused” (P29) activities.

“Simplified managing the problem of magnitude by group formation” (P22). Then to enhance students’ enthusiasm , “a bond of friendliness is created” (P12), while another teaches “takes them to the real world so they realize the importance” (P8). Again Morris “constructs tests which fit the context, situation, and the students’ level” (P18) and, “to match the weak participants’ level of language proficiency with the highly advanced course outline,“efforts are made to weave small exercises and activities to address their rudimentary gap” (P20). Meanwhile, the problem of subjectivity is resolved by the predetermined scoring guide.

“Rubrics allow determining learning outcomes to make the evaluation easier “ (P3). “Teachers develop descriptors and then grade them independently to the descriptors” (P3). These samples confirm that the teachers’ practices are responsive to the context school and the broader community. As one teacher stated, “for certain situations, a teacher has no option but to improvise” (P10).

7. Consequences: Ineffective Assessment Practices Demoralizing and Negative Effects

Answering the questions whether these constraints have any bearing on their assessment practices, the respondents acknowledged that they have demoralizing and negative effects. ‘It affects us morally sometime getting so much criticism from surrounding even from your students’ (P1). The negative effects include ‘students and teachers’ preference of traditional approaches’ (P4) and ‘dominance of writing assessment since students are mostly assessed through summative approach and writing conveniently dominates the assessment’ (P6). During the analysis of the examination papers, in particular, these same conclusions were arrived at. The analysis revealed that assessment practices in higher education of these types are still



not up to the mark advanced in language assessment and still are narrow in ranges of focused language skills and sub-skills and the types of tasks included in the test.

7.1. Selective Skills and Sub-skills

The assessment practices seemed to almost exclusively focus on the writing skill, while reading was assessed only with a few sub-skills. The assessment of reading skill was confined to comprehension of explicit and implicit information, inference of meaning of unfamiliar words, identification of the topic sentence, and the suggestion of a title. Writing, on the other hand, was assessed mostly with extended production such as essays and reports. These various tasks appear to emphasize unvaried behavior and are quite distant from what is considered “writing.” Some more complex writing tasks only required the addition of a topic sentence and a thesis statement. Listening and speaking activities were nearly absent from the assessment of these courses. In the set of collected documents, there were only three instances of assessing listening, all of which focused on the skill of scanning and inferring the intentions of the speakers. Speaking, on the other hand, was restricted to presentations. In the final examination, there was no task with which the students were able to demonstrate their speaking proficiency. It should be emphasized that the results of these assessment activities are accesses to be equivalent to the ‘overall’ proficiency of the students in the target language, which, it is claimed, involves all four language skills.

7.2.Redundant Test Tasks

Equally, particular problems emerged from the selection of balancing unresolved test tasks which proved to be of some importance. Even though the exam papers displayed an assortment of test tasks meant to gauge one’s language prowess, some tendencies were noted regarding the nature of the test on specific skills, for instance, short answer and MCQ for reading comprehension, and narrative and descriptive essays, letters, and short reports for writing. In listening comprehension, only closed-ended questions were designed, and speaking proficiency was only assessed through presentation delivery. These test tasks are insufficient to provide an accurate portrayal of the language performance of test takers. Furthermore, grammar and vocabulary were assessed in isolation from context and through separate and discrete testing which undermined the purpose of authentic assessment.

7.3.Target Language Use Domain and Integration

Such restrictions on assessing performance indicate even more serious and difficult issues. The types and the aims of the tasks in the papers that were examined do not align with the target language use (TLU) domains and may have backward negative washback. Furthermore, in the tasks, the focus skills were not integrated and were evaluated separately. Such tasks do not truly represent an individual's language proficiency. Several test tasks in the collected papers also assessed the test takers’ linguistic knowledge in isolation. These tasks aimed at understanding the concept of the skills rather than the application of the skill. During the analysis, thirty examples of such test tasks were noted. The tasks elicited knowledge of writing, speaking, and vocabulary. Although these tasks are not the only means of performance for which competence in language is needed, completion of these tasks does not reflect the test takers’ true language skills and ability to communicate in a discourse.

7.4.The Discrepancy Between Teachers’ Provided Feedback and Their Actions



In this case, the answer offers details related to the discrepancies. Most practitioners deny discrimination among language skills and maintain that assessments should treat all skills equally. The analysis of examination papers reveals little agreement to the assertion advanced by most language teachers. Equally, a significant number of teachers strongly believe that learners, their language proficiency and their background knowledge are testable components and are factors to be considered in the scope of test design. The examination papers' reading and writing tasks contain many topics that learners in Pakistan who study English as a second language, would find difficult to relate to.

8. Discussion

Among the various difficulties experienced by teachers participating in this study, assessment illiteracy emerged strongest. The latter concurs with Sheehan and Munro's (2017) study which found teachers in the Pakistani context to be devoid of basic knowledge and understanding of assessment techniques because of the insufficient availability of assessment training. The respondents noted that illiteracy amongst learners was also an obstacle which has been documented in Lumadi (2013) where the teachers were said to have been reflecting in order to cater for the varying needs of the learners. While assessment of language involves determining the level of proficiency an individual has, it encompasses also the planning, designing, administering, scoring, and inferring from the scores. In the current instance, the teachers in question think that the inability to speak the language is a barrier to assessment. They fail to perceive that assessment can and should be used to inform and enhance their language proficiency and that it is not limited to the assumptions made about the language proficiency of the test takers. Staff within the administration of higher education institutions is typically rather rigid with respect to the adoption of new and innovative assessment techniques. The focus is rather on the conventional approaches and methods of assessing the language. This correlates with Bresciani (2010) which indicated how the 'hands-off' attitude from administration was willing to invest in assessment. Their lack of appreciating the need to construct assessment procedures in language assessment is critical. It is puzzling the extent of the impact – assessment decisions, whether big or small – made during the assessment procedures is on the test takers and their future opportunities. Such rigid and non-cooperative behavior from the authorities only adds on to the 'feather bed' situation with language assessment. It serves as a reminder that a single teacher, however proficient, is unable to effect changes in the assessment processes unless such changes are officially sanctioned. It is the lack of such changes which in the current situation remains unfulfilled, and the expectation is that the ground realities will be as they premise out to be as test developers or teachers of English language teaching (East, 2015) – which almost never is the case. Proficiency in a language is acquired and learned over a period of time, and hence, should be evaluated as such. The teachers within the confines of the present study seem to focus primarily on summative assessment as supported by Kurebwa and Nyaruwata (2013). The summative assessment conducted at the close of the academic term creates a burden on the candidates because all the language skills are evaluated at once. It can hinder performance by employing physical and mental stress. Sharjeel and Qazi (2012) championed a similar case and asserted that one of the core tenets of reading comprehension is able to focus and pay attention to the reading material and the ability to master the reading skills at higher levels of education, however, the undue burden of other skills has a negative impact on reading comprehension. They also



rebutted claims that a summative form of language assessment succeeds in achieving performance of the desired skills. Likewise, Raza (2009) refute claims that one-off achievement tests are effective forms of assessment of language use at higher education level. It is as if, the language assessment field still has some fundamental, as well as, newly generated issues to deal with due to the shifts in language and the global world. This is the position taken by Bachman (2014). He indicated that, when it comes to language assessment, the 'greatest challenges' do not center on complex issues such as the theory of validity, socio-psychological or postmodern theories, and statistical models, but rather come from the 'real world' contexts in which the language assessment is used to make decisions. This position is clearly illustrated by the respondents' challenges in the study, which are rather fundamental and based on practical realities. From the data, it can be argued that the assessment practices in place for the English courses at the undergraduate level do not meet the necessary standards of various domains. First, the examination papers contained several isolated instances of competence based items that did not interface with language performance. This kind of assessment on the ability to use target language is dangerous to the validity of the test. Validity is central to assessment and is one of the major language assessment principles (Green, 2014). However, there is a significant issue when testees are awarded marks for what is deemed as knowledge of the language, simply to be later positioned in the system as language users. Validity of the test tasks comes to question here. Integrated and communicative test tasks enhance construct validity and improve authenticity. In order to enhance the realism of the elements of assessment, Raza (2009) suggested the inclusion of integrating assessment of the target language's communicative functions at the higher education level. These results also indicate that the examination paper's construction emphasizes the reading and writing skills tremendously while neglecting the listening and speaking components. Since the assessment is the primary indicator of the test takers' overall language competence, these test papers must include and balance all four language skills and these sub-skills at a minimum. The constraints of the teachers in terms of time and resources do not, however, negate the role of listening and speaking in real-life situations. The narrow context of language skills and sub-skills is a clear case of construct underrepresentation (Grabe & Jiang, 2014). On the other hand, the balance of assessment among reading and writing skills also requires some attention. There is evidence that when teachers elaborate assessment tasks for reading and its sub-skills, the proportion of writing tasks seems to dominate considerably more than the reading tasks. In comparison to four or five writing tasks, only one or two questions are reserved for reading.

To complete and submit a five-genre text under the instruction span of two-to-three hours is virtually impossible; thus, students learn how to memorize and reproduce ready-made compositions. Such practices of assessments sustain and foster the ability to learn the formulaic aspects of writing, undermining critical thought, a fundamental necessity in academic writing (Cummings, 2013). Luoma (2004) stated that speaking components are usually evaluated through one-on-one interviews. Yet, in this case, the results indicated that speaking assessment is limited to delivery of presentations only. They prepared presentations, by their nature, as monologs are contrary to the language of the real world dominated by dialogues or language as a social activity between two or more participants. While presentations are part of academic and professional communication, they are not the only genre necessary for effective language use. Assessment of speaking should include individual along with pair and group work to replicate genuine language



use and prepare the students for communication in the real world. Hidri (2014) noted that the assessment practices in Pakistan are underdeveloped as opposed to the modern practices used outside the country.

9. Conclusion

The research results point to a number of factors that teachers encounter as challenges in language assessment, including teachers' sluggish assessment literacy, learner heterogeneity and language deficiencies, dismal test construction, institutional requirements, and pragmatic constraints—class size, time, and available resources. All of these factors also limit teachers' capacities in the design and delivery of authentication language assessment in higher education. The resultant restrictions lead to poor assessment design and a limited construct of skills and sub-skills in examination papers that fail to epitomize the learner's language proficiency. The results also point to the fact that teachers at different universities encounter the same barriers, regardless of the level of education attained or the category of institution. Such challenges could be transformed into assets for the education system if those who construct language assessments are willing and able to execute ideas grounded in theory and practice (Bachman, 2014). Consequently, Kamrani (2010, cited in Khattak, 2012) suggested that the learners be provided the tools and methods so they can perform the language assessment in multiple ways for its optimal development.

References

- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press
- Bachman, L. F. (2014). *Ongoing Challenges in Language Assessment*. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 1586-1601). John Wiley & Sons, Inc.
- Bresciani, M. J. (2010). *Understanding Barriers to Student Affairs Professionals' Engagement in Outcomes-Based Assessment of Student Learning and Development*. *Journal of Students' Affairs*, XIX, 81-89.
- Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in education*. Noida: Routledge.
- Coombe, C. (2009). *Washback and the impact of high-stakes tests on teaching and learning*. In Mansoor, S., Sikandar, A., Hussain, N. & Ahsan, N. M. (Eds.). *Emerging issues in TEFL: Challenges for Asia* (298-307). Oxford: Oxford University Press
- Coombe, C, Troudi, S, and Al-Hamly, M (2012). *Foreign and Second Language Teacher Assessment Literacy: Issues, Challenges and Recommendations*. In Coombe, C, Davidson, P, O'Sullivan, B, and Stoyloff, S (eds.) *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press
- Cumming, A. (2013). *Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils*. *Language Assessment Quarterly*, 10(1), 1–8.
<https://doi.org/10.1080/15434303.2011.622016>
- Dar, M. F. & Khan, I. (2014). *Oral communication apprehension among*



- undergraduate Engineering students in Pakistan. *Journal of Education and Social Sciences*, 2(2), 144-153. Retrieved from <https://geistscience.com/papers/view/JESS1402203>
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press
- Bachman, L. F. (2014). *Ongoing Challenges in Language Assessment*. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 1586-1601). John Wiley & Sons, Inc.
- Bresciani, M. J. (2010). *Understanding Barriers to Student Affairs Professionals' Engagement in Outcomes-Based Assessment of Student Learning and Development*. *Journal of Students' Affairs*, XIX, 81-89.
- Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in education*. Noida: Routledge.
- Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in education*. Noida: Routledge.
- Coombe, C. (2009). *Washback and the impact of high-stakes tests on teaching and learning*. In Mansoor, S., Sikandar, A., Hussain, N. & Ahsan, N. M. (Eds.). *Emerging issues in TEFL: Challenges for Asia* (298-307). Oxford: Oxford University Press
- Coombe, C, Troudi, S, and Al-Hamly, M (2012). *Foreign and Second Language Teacher Assessment Literacy: Issues, Challenges and Recommendations*. In Coombe, C, Davidson, P, O'Sullivan, B, and Stoyloff, S (eds.) *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press
- Cumming, A. (2013). *Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils*. *Language Assessment Quarterly*, 10(1), 1-8. <https://doi.org/10.1080/15434303.2011.622016>
- Dar, M. F. & Khan, I. (2014). *Oral communication apprehension among undergraduate Engineering students in Pakistan*. *Journal of Education and Social Sciences*, 2(2), 144-153. Retrieved from <https://geistscience.com/papers/view/JESS1402203>
- Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Duncan, C.R., & Noonan, B. (2007). *Factors affecting teachers' grading and assessment practices*. *The Alberta Journal of Educational Research*, 53(1), 1-21. Retrieved from <https://journalhosting.ucalgary.ca/index.php/ajer/article/view/55195>
- East, M. (2015). *Coming to terms with innovative high-stakes assessment practice: Teachers' viewpoints on assessment reform*. *Language Testing*, 32(1), 101-120. <https://doi.org/10.1177/0265532214544393>
- Fleming, M. (2007). *The Challenges of assessment within language(s) of education*. In Martyniuk, W. (Ed.). *Evaluation and assessment within the domain of language(s) of education*. Strasbourg: Council of Europe.
- Flick M. (2009). *Introduction to Qualitative Research*. United Kingdom. Rachel Boak Publications
- Grabe, W. & Jiang, X. (2014). *Assessing reading*. In Kunnan, A. J. (Ed.). *The Companion to Language Assessment* (pp. 185-200). UK: Wiley-Blackwell.
- Green, A. (2014). *Exploring language assessment and testing: Language in action*. Oxon: Routledge.
- Harlen, W. (2005). *Teachers' summative practices and assessment for learning - tensions and synergies*. *The Curriculum Journal*, 16, (2), 207-223. <https://doi.org/10.1080/09585170500136093>



- Hidri, S. (2014). Developing and evaluating a dynamic assessment of listening comprehension in an EFL context. *Language Testing in Asia*, 4(1), 1–19. <https://doi.org/10.1186/2229-0443-4-4>
- Hughes, A. (2010). *Testing for language teachers* (2nd ed.). Cambridge University Press, South Asian Edition, Noida: India Binding House.
- Isani, U. A., & Virk, M. L. (2005). *Higher education in Pakistan: A historical and futuristic perspective*. Islamabad: National Book Foundation.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press
- Bachman, L. F. (2014). Ongoing Challenges in Language Assessment. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 1586-1601). John Wiley & Sons, Inc.
- Bresciani, M. J. (2010). Understanding Barriers to Student Affairs Professionals' Engagement in Outcomes-Based Assessment of Student Learning and Development. *Journal of Students' Affairs*, XIX, 81-89.
- Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in education*. Noida: Routledge.
- Coombe, C. (2009). Washback and the impact of high-stakes tests on teaching and learning. In Mansoor, S., Sikandar, A., Hussain, N. & Ahsan, N. M. (Eds.). *Emerging issues in TEFL: Challenges for Asia* (298-307). Oxford: Oxford University Press
- Coombe, C, Troudi, S, and Al-Hamly, M (2012). Foreign and Second Language Teacher Assessment Literacy: Issues, Challenges and Recommendations. In Coombe, C, Davidson, P, O'Sullivan, B, and Stoyloff, S (eds.) *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press
- Cumming, A. (2013). Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils. *Language Assessment Quarterly*, 10(1), 1–8. <https://doi.org/10.1080/15434303.2011.622016>
- Dar, M. F. & Khan, I. (2014). Oral communication apprehension among undergraduate Engineering students in Pakistan. *Journal of Education and Social Sciences*, 2(2), 144-153. Retrieved from <https://geistscience.com/papers/view/JESS1402203>
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press
- Bachman, L. F. (2014). Ongoing Challenges in Language Assessment. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 1586-1601). John Wiley & Sons, Inc.
- Bresciani, M. J. (2010). Understanding Barriers to Student Affairs Professionals' Engagement in Outcomes-Based Assessment of Student Learning and Development. *Journal of Students' Affairs*,



XIX, 81-89.

Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in education*. Noida: Routledge.

Coombe, C. (2009). Washback and the impact of high-stakes tests on teaching and learning. In Mansoor, S., Sikandar, A., Hussain, N. & Ahsan, N. M. (Eds.). *Emerging issues in TEFL: Challenges for Asia* (298-307). Oxford: Oxford University Press

Coombe, C, Troudi, S, and Al-Hamly, M (2012). *Foreign and Second Language Teacher Assessment Literacy: Issues, Challenges and Recommendations*. In Coombe, C, Davidson, P, O'Sullivan, B, and Stoyloff, S (eds.) *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press

Cumming, A. (2013). *Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils*. *Language Assessment Quarterly*, 10(1), 1–8.

<https://doi.org/10.1080/15434303.2011.622016>

Dar, M. F. & Khan, I. (2014). Oral communication apprehension among undergraduate Engineering students in Pakistan. *Journal of Education and Social Sciences*, 2(2), 144-153. Retrieved from <https://geistscience.com/papers/view/JESS1402203>

Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press

Bachman, L. F. (2014). *Ongoing Challenges in Language Assessment*. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 1586-1601). John Wiley & Sons, Inc.

Bresciani, M. J. (2010). Understanding Barriers to Student Affairs Professionals' Engagement in Outcomes-Based Assessment of Student Learning and Development. *Journal of Students' Affairs*, XIX, 81-89.

Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in education*. Noida: Routledge.

Coombe, C. (2009). Washback and the impact of high-stakes tests on teaching and learning. In Mansoor, S., Sikandar, A., Hussain, N. & Ahsan, N. M. (Eds.). *Emerging issues in TEFL: Challenges for Asia* (298-307). Oxford: Oxford University Press

Coombe, C, Troudi, S, and Al-Hamly, M (2012). *Foreign and Second Language Teacher Assessment Literacy: Issues, Challenges and Recommendations*. In Coombe, C, Davidson, P, O'Sullivan, B, and Stoyloff, S (eds.) *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press

Cumming, A. (2013). *Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils*. *Language Assessment Quarterly*, 10(1), 1–8.



<https://doi.org/10.1080/15434303.2011.622016>

Dar, M. F. & Khan, I. (2014). Oral communication apprehension among undergraduate Engineering students in Pakistan. *Journal of Education and Social Sciences*, 2(2), 144-153. Retrieved from <https://geistscience.com/papers/view/JESS1402203>

Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press

Bachman, L. F. (2014). Ongoing Challenges in Language Assessment. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 1586-1601). John Wiley & Sons, Inc.

Bresciani, M. J. (2010). Understanding Barriers to Student Affairs Professionals' Engagement in Outcomes-Based Assessment of Student Learning and Development. *Journal of Students' Affairs*, XIX, 81-89.

Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in education*. Noida: Routledge.

Coombe, C. (2009). Washback and the impact of high-stakes tests on teaching and learning. In Mansoor, S., Sikandar, A., Hussain, N. & Ahsan, N. M. (Eds.). *Emerging issues in TEFL: Challenges for Asia* (298-307). Oxford: Oxford University Press

Coombe, C, Troudi, S, and Al-Hamly, M (2012). Foreign and Second Language Teacher Assessment Literacy: Issues, Challenges and Recommendations. In Coombe, C, Davidson, P, O'Sullivan, B, and Stoyloff, S (eds.) *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press

Cumming, A. (2013). Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils. *Language Assessment Quarterly*, 10(1), 1-8.

<https://doi.org/10.1080/15434303.2011.622016>

Dar, M. F. & Khan, I. (2014). Oral communication apprehension among undergraduate Engineering students in Pakistan. *Journal of Education and Social Sciences*, 2(2), 144-153. Retrieved from <https://geistscience.com/papers/view/JESS1402203>

Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press

Bachman, L. F. (2014). Ongoing Challenges in Language Assessment. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 1586-1601). John Wiley & Sons, Inc.

Bresciani, M. J. (2010). Understanding Barriers to Student Affairs Professionals' Engagement in Outcomes-Based Assessment of Student Learning and Development. *Journal of Students' Affairs*, XIX, 81-89.

Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in*



education. Noida: Routledge.

Coombe, C. (2009). Washback and the impact of high-stakes tests on teaching and learning. In Mansoor, S., Sikandar, A., Hussain, N. & Ahsan, N. M. (Eds.). *Emerging issues in TEFL: Challenges for Asia* (298-307). Oxford: Oxford University Press

Coombe, C, Troudi, S, and Al-Hamly, M (2012). *Foreign and Second Language Teacher Assessment Literacy: Issues, Challenges and Recommendations*. In Coombe, C, Davidson, P, O'Sullivan, B, and Stoyoff, S (eds.) *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press

Cumming, A. (2013). *Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils*. *Language Assessment Quarterly*, 10(1), 1–8.

<https://doi.org/10.1080/15434303.2011.622016>

Dar, M. F. & Khan, I. (2014). Oral communication apprehension among undergraduate Engineering students in Pakistan. *Journal of Education and Social Sciences*, 2(2), 144-153. Retrieved from

<https://geistscience.com/papers/view/JESS1402203>

Kankam, B., Bordoh, A., Eshun, I., Bassaw, T.K., Korang, F.Y. (2014). Teachers' perception of authentic assessment techniques practice in social studies lessons in senior high schools in Ghana. *International Journal of Educational Research and Information Science*. 1(4), 6268.

Khattak, S. G. (2012). Assessment in schools in Pakistan. *SA-eDUC Journal*, 9(2). Retrieved from http://www-nwu.ac.za/web.nwu.ac.za/sites/www.nwu.ac.za/files/files/psaeduc/New_Folder_1/6_Assessment%20in%20schools%20in%20Pakistan.pdf

Kurebwa, M. & Nyaruwata, L.T. (2013). Assessment challenges in the primary schools: a case of Gweru Urban Schools. *Greener Journal of Educational Research*, 3(7), 336-344.

Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. London: Longmans.

Lumadi, M.W. (2013). Challenges besetting teachers in classroom assessment: an exploratory perspective. *Journal of Social Science*, 34(3), 211-221. <https://doi.org/10.1080/09718923.2013.11893132>

Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press. Agriculture

Masole, T.M. (2011). *Enhancing the quality of performance assessment in Botswana*. (Unpublished dissertation). University of Pretoria, Pretoria. doctoral

Matovu, M. & Zubairi, A. M. (2014). Factors influencing assessment practices among university staff: A multiple regression analysis. *Mevlana International Journal of Education*. 4(11). 176-188 [10.13054/mije.13.57.4.1](https://doi.org/10.13054/mije.13.57.4.1)

Maxwell, J. A. (2010). Validity: How might you be wrong? In W. Luttrell (Ed.), *Qualitative Educational Research: Readings in Reflexive Methodology and Transformative Practice* (pp. 279-287). New York: Routledge.

McNamara, T. (2004). Language testing. In Davies, A. & Elder, C. (Eds.). *The handbook of applied linguistics* (763-783). Oxford: Blackwell.



- Ndalichako, J.L (2004). Towards an Understanding of Assessment Practices of Primary School Teachers in Tanzania, *Zimbabwe journal for Educational* 17710.4314/zjer.v16i3.26046 research. 16(3),168
- Popham, W.J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 11. <https://doi.org/10.1080/00405840802577536> 48(1), 4
- Qureshi, M. A., Shirazi, R. A. & Wasim, M. P. (2007) Perspectives and prospects of commencing new education policy (NEP) of Pakistan: A review of conference. *Indus Journal of Management & Social Sciences*, 1(2), 176-176.
- Raza, W. (2009). English language testing in higher education of Pakistan. *Market Forces* .
- Rasool, N. (2009). Quality issues in language teaching in higher education. In Mansoor, S, Sikandar, A, Hussain, N, & Ahsan N. M.(Eds.) *Emerging Issues in TEFL Challenges for Asia* (pp. 181-190). Oxford: Oxford University Press.
- Rehmani, A. (2007). Teacher education in Pakistan with particular reference to teachers' conceptions of teaching. In Ali, S. & Rizvi, M. (Eds.). *Quality in education: Teaching and leadership in challenging times, proceedings Vol. 2* (pp. 435-869). Karachi: Aga Khan University Institute for Educational Development.
- Saldana, J. (2009). *The coding manual for qualitative researchers*. London: SAGE Publications Ltd.
- Shamim, F. (2011). English as the language for development in Pakistan: Issues, challenges and possible solutions. In H. Coleman (Ed.), *Dreams and Realities: Developing Countries and the English Language* (pp. 291-309). London: British Council.
- Sharjeel, M. Y. & Qazi, W. (2012). Why does testing English language skills really matter? Issues and challenges in Pakistani higher education. *Journal of Education Research*, 15(1), 93-107.
- Sheehan, S. & Munro, S. (2017). *Assessment: Attitudes, practices and needs*. London: British Council.
- Stiggins, R. (1991). Assessment literacy. *Phi Delta Kappan*, 72, 534-539.
- Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 238-245.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-36.
- Trice, A.D. (2000). *A Handbook of Classroom Assessment*. New York. Addison Wesley.
- Longman UNESCO (2005). *EFA Global monitoring report*.
- Wallace, S. (Ed) (2009) *A Dictionary of Education*. Oxford: Oxford University Press.
- Teachers' Zhang, Z., & Burry-Stock, J. A. (2003). Classroom Assessment Practices and Self-Perceived Assessment Skills. *Applied Measurement in Education*, 16(4), 323-342.