Language Assessment in Higher Education: Challenges and Consequences

Almas Ashraf[*]
Sajida Zaki[**]

Abstract

Several research studies have reported that assessment at higher education institutions for English language courses in Pakistan are deficient. However, they have not addressed the challenges causing these weaknesses. The current study, therefore, is an attempt to gain insight into the reasons of and the resulting consequences of these unsatisfactory assessment practices. The study employed qualitative approach to explore the challenges in the language assessment and the ensuing consequences. Thirty teachers were selected from public and private universities as the sample using purposive sampling. The data were collected using semi-structured questionnaires, in-depth interviews and examination papers. The findings reveal hat language teachers come across several challenges in assessment. There are various factors that lead to these challenges such as practicality, learners, test design, administrative obligations and teachers themselves. Although, the teachers employ several strategies to cope with the challenges, they have an adverse impact on language assessment practices in terms of selective skills and sub skills as well as limited choice of test tasks. Assessment has a strong impact on language teaching and learning and the challenges in assessment hamper the language teachers from designing useful assessment. Therefore, by addressing the potential barriers, language assessment practices can be reshaped in a way that they are conducive to language learning and teaching. The results also emphasize the need to enhance language teachers' assessment literacy through training and professional development programs.

Keywords: language assessment; higher education; challenges; problems; consequences.

[*] *Assistant Professor (English), Humanities Department, NED University of Engineering and Technology, almasashraf@neduet.edu.pk*
[**] *Professor & Chairperson, Humanities Department, NED University of Engineering & Technology, drzaki@neduet.edu.pk*

Introduction

Lado (1961) defined language as a 'highly complex' and 'conventionalized system of habits' used by humans for communication. He referred to it as 'the most complex of man's tool' (p. 2). This complex system involves the use of language knowledge and skills to perform the tasks of reading, writing, listening, and speaking. From these interactive yet limited instances, inferences are made about the more general and vast knowledge, abilities and skills of a language user. In a similar way, language assessment is employed where it obtains evidences of language use under test or classroom settings and interprets them to infer an individual's overall language proficiency and to predict their language performance in real-world situations (Green, 2014; McNamara, 2004). On the other hand, Bachman (2004) defined language assessment as a process of gathering information about language abilities through 'systematic and substantially grounded procedures'. The information can be obtained from variety of sources including multiple choices items, essays, portfolio, oral interview, observation and introspection. Hence, the primary purpose of language assessment is to depict the processes as well as outcomes of teaching and learning, to provide a 'common yardstick' to compare test takers' performance and to inform decisions about promotion, employment and immigration (Hughes, 2010).

The potential of universities and higher education has been recognized globally. They contribute significantly to the national development through generation, dissemination and utilization of knowledge (Isani&Virk, 2005). This responsibility is amplified with the rapid changes and increasing competition in the global market (Rasool, 2009). The globalization and internationalization of knowledge and learning has given higher education a distinctive position in national and international context. The universities must meet international standards and produce graduates and scholars who can compete internationally. For this reason, universities must be responsive to the rapidly changing expectations, challenges and demands of increasing students (Isani & Virk, 2005).

Owing to the global status of English, the assessment of English language has become a vital component of any English language course or program. It is employed when decisions are to be made about learners' strengths and weaknesses in the language (Bachman, 2004; Hughes, 2010). It is not only an important indicator of individuals' knowledge and language skills but also of the quality of language learning and

language teaching cycle (Green, 2014). It also helps in making informed decisions about the admission, placement and achievement of language learners as well as about their immigration and employment in inner and outer circle countries. Consequently, language assessment at the higher education, particularly in developing countries like Pakistan, carries high stakes consequences for language learners. High stake refers to the assessment which influences examinees' academic and professional opportunities, e.g. promotion, graduation, and job opportunities at home or abroad (Coombe, 2009).

The importance of language assessment for the successful language learning is acknowledged by learners and teachers globally. As Coombe (2009) pointed out, if tests result in consequential decisions then teachers 'teach to test' and the teaching and learning practices become limited to the contents of past examination papers, repeated administration of previous tests, and training students in the test tasks (Harlen, 2005). Rehmani (2007) observed similar impact of examinations on Pakistani pedagogical practices of language courses. In Pakistan, teachers are responsible for assessment at tertiary level. They have remarkable influence on their students' learning. The teachers' beliefs and practices, in turn, are shaped by their teachers and their experience of learning and teaching (Rehmani, 2007). He argued that by preparing for the upcoming tests on the basis of past exam papers the learning becomes limited. Thus, examination system of Pakistan has become a determinant factor that shape language teachers' concepts, beliefs and practices on one hand and language learners' on the other. Consequently, the flaws in the assessment lead to the misalignment of the curriculum and instructions (Isani & Virk, 2005).

It has been observed that language assessment in Pakistan is more focused on reading and writing whereas listening and speaking do not get any substantial part in these practices. This observation is equally relevant to the context of higher education as well (Dar & Khan, 2014). The reasons may vary from teachers' pedagogical skills to available resources and opportunities. The teachers, regardless of being affiliated to public or private universities, are neither formally trained nor motivated to engage in professional development. Shamim (2011) summarized the current provision for the teaching and assessment of English at tertiary level that majority of teacher's lack formal training or qualification in ESL teaching and testing. Additionally, the teachers do not engage in 'professional development activities or dialogue with wider ELT community' (p. 7). Thus, their teaching and assessment

practices fail to align with the contemporary advancement in the field. Moreover, their assessment practices also reflect their lack of knowledge about the alternative methods of language assessment and their purposes. The assessment is mostly carried out using summative approach even for languages which are acquired gradually and which must be assessed along with the process carefully addressing the curricular objectives. As Raza (2009) argued that 'a large-scale-one-time achievement test', as conducted in Pakistan, does not represent students' language abilities and skills truly.

Problem Statement

Globally, in the second language context, different types of assessment are employed depending on the purpose of the assessment; however, in Pakistani higher education, the same methods have been used for years regardless of the change in the language teaching, learning and assessment process. The achievement tests employed by the universities to assess whether students have achieved the objectives of the courses become proficiency tests as the students are assessed on their general linguistic abilities. Such issues in the language assessment practices at tertiary level need to be addressed on urgent basis as they demand remedial measures based on empirical and objective research. However, ironically, despite the significant role of assessment, the assessment system itself usually goes unexamined. It is rare to find any systematic way to evaluate the decisions made based on language assessment or to determine factors and challenges that influence the language assessment. As education in Pakistan is by and large exam-oriented, the shortcomings in the assessment practices reflect in the failure of the achievement of learning outcomes. These shortcomings need to be addressed because assessment in Pakistani context involves high-stakes decisions. Moreover, the critical review of the literature revealed that very little work has been carried out on language assessment with reference to language assessment in Pakistani higher education. Even those attempted are so diverse in their scope and focus that a comprehensive conclusion is difficult to draw. Additionally, most of the research studies have highlighted the dismal state of current examination system and language assessment practices (Qureshi, Shirazi & Waseem, 2007; Raza, 2009; Rehmani, 2007), but few have attempted to raise questions about the factors that lead to such substandard practices. Thus, this study is an attempt to provide an in-depth insight

into the challenges and their influential effects on language assessment practices in higher education.

Research Objective

The objective of the study is to identify the challenges in the assessment of language in higher education. It also aims at highlighting the ensuing consequences for the current assessment practices carried out for undergraduate English courses.

Literature Review

Language assessment can inform important decisions regarding language teaching and language learning if conducted and interpreted rightly, but it may produce regrettable results if the information obtained from assessment is insufficient, badly conceived or poorly interpreted (Green, 2014). The common challenges that emerged from the findings of the studies carried out to investigate the reasons for substandard assessment practices include teachers' subjective interpretation of policy, teachers' inability to plan and implement assessment, their limited understanding of assessment, institutional preferences of assessment methods, time, and lack of resources (Bresciani, 2010; Kankum et al., 2014; Kurebwa & Nyaruwata, 2013; Lumadi, 2013).

Teachers' assessment literacy has been discussed profusely in the literature as an essential prerequisite for useful assessment practices. Matovu & Zubairi (2014) highlighted assessment literacy as a significant predictor of university teachers' assessment practices. Zhang & Burry-Stock (2003) also emphasized that teachers' ability to perform effective assessment depends on their assessment training, the absence of which may affect the assessment of students. Taylor (2009) defined assessment literacy as knowledge and application of the processes of measurement in classrooms and in language assessment. Assessment literate teachers know what to assess, why to assess and how to assess; they are also aware of the potential problems and the relevant coping mechanisms. Moreover, they also realize that poor and inaccurate assessment may result in negative consequences (Stiggins, 1991; 1995). Therefore, for language assessment, the teachers must be aware of the relevant language acquisition theories, the purposes for which the language is learnt, and methods to elicit the representative language performance to ensure valid inferences about the test takers' knowledge and skills (Green, 2014).

A relevant challenge in language assessment is teachers' subjective understanding and interpretations of assessment purposes, methods and principles. Very few teachers are aware of the multidimensional use of assessment and tend to neglect their contextual needs in search of universal solutions. These differences of opinion and disagreements result in unproductive assessment practices (Flemming, 2007). Nevertheless, the concept of training in assessment is new to the context like Pakistan. The teachers' assessment literacy is not measured when they are inducted as teachers in higher education institutes, although, it is common knowledge that these teachers will be solely responsible for all the assessment-related decisions. Trice (2000) also observed that teachers do not require any assessment course or training for most certification system run by the state. As a result, teachers enter into the profession without any fundamental knowledge of educational assessment (Popham, 2009).

Class size and transparency also pose as a potential challenge for language assessment. Large class size has a negative impact on the assessment practice. It is considered a threat to quality assessment for teachers as well as students (Matovu & Zubairi, 2014). Teachers assess poorly when they have large number of students to assess. Assessment in large classroom results in barriers like ineffective use of equipment, time and resources, rater's bias and lack of concentration (Masole, 2011). Alternative assessment methods also become time consuming and impractical in large classes. Thus, teachers fall back to multiple-choice items and traditional methods for test purposes (Duncan & Noonan, 2007). On the other hand, Flemming (2007) raised the issue of transparency in assessment. Transparency refers to the awareness of students about the judgment criteria and procedures. Such awareness facilitates students' performance and motivation. He also highlighted another challenge pertaining to the multidimensional nature of language and representative selection of test tasks to assess skills and sub-skills to generalize the result to the overall ability to use language.

In addition to large class and transparent assessment, limited resources and time constraints also become barriers in successful language assessment (Coombe, Troudi, & Al-Hamly, 2012). The increasing responsibilities on the teachers as an assessor affect their teaching practices as well. The participant in Lumadi's (2013) study identified that teachers fail to achieve the learning objectives as well as the standards of assessment while struggling to complete the assessment

along with their teaching workload. To deal with this constriction, teachers tend to design objective tests for quick and reliable scoring. However, such practice result in exclusion of several important language skills and sub-skills from their assessment practices (Harlen, 2005). These constraints are also highlighted by Wallace (2009) he contended that pressure from the key stakeholders leads the teachers to restricted curriculum and limited focus on test requirements at the cost of students' language learning.

These challenges result in unproductive assessment practices that fail to promote learning. Ndalichako's (2004) investigation of the primary school teachers' assessment practices revealed teachers' heavy reliance on the traditional methods and rare use of alternative assessment due to the physical, cognitive and psychological constraints. This heavy reliance on summative assessment hinders their exposure to alternative assessment practices such as portfolios, performance assessment and formative assessment (Lumadi, 2013). It also induces test anxiety in learners and negative impact on teachers and curriculum (Harlen, 2005). The manifold responsibilities of teachers and scarce facilities by the administrative staff also result in teachers' casual approach to assessment. Consequently, teachers are found to target a few learning objectives and duplication of past papers item. (Kurebwa & Nyaruwata, 2013)

Harlen (2005) suggested some remedies to reduce the impact of such challenges. He emphasized that the purpose of assessment should be explicitly explained to the students and they should be involved in the assessment decisions. He also recommended teachers to provide feedback to the students, use assessment to track their progress and enable their students to relate the criteria to their learning through self-assessment. Additionally, awareness of teachers' knowledge and beliefs about assessment as well as of the challenges may prove crucial for the planning and implementation of teachers' professional development in assessment (Lumadi, 2013).

Methodology
Research Design
The current study used qualitative paradigm to develop a detailed understanding of the challenges in language assessment and their consequences. Qualitative paradigm not only explores a problem but also develops a detailed understanding of the central phenomenon

which is the primary purpose of this research (Cohen, Manion& Morrison, 2010).

Sample

Thirty teachers from seventeen higher education institutes of Karachi were selected using purposive sampling. Purposive sampling involves choosing a sample satisfying researcher's specific needs based on their judgement or characteristics (Cohen et al., 2010). The researchers identified male and female teachers from both public and private universities offering degrees in general and professional disciplines. All the selected teachers were engaged in teaching undergraduate English language courses and were responsible for the assessment.

Data Collection

Data were collected through several sources to ensure triangulation. First, the teachers completed a semi-structured questionnaire. To further probe their views, ten of the thirty teachers were also interviewed. In-depth interviews have always attracted interest in qualitative research since it provides more expressive platform to the respondents than structured interviews or questionnaires (Flick, 2009). The researchers, then, collected examination papers (and midterm, quizzes, sessional activities or tasks) set by the sample (n=30) to analyse the effects of the constraints on their assessment practices.

The gathered data were purely textual and 'language-based' (Dornyei, 2007). However, the data differ in terms of primary and secondary data. The data generated through semi-structured questionnaires and in-depth interviews was chiefly primary whereas the examination papers fall under secondary data.

Data Analysis

Thematic analysis was used to analyse the data collected from questionnaires and interviews. The analysis of the qualitative data began with general and broad coding, a word or phrase that carried symbolic attributes for a portion of text (Saldana, 2009). These codes were, then, grouped together as emergent themes using the methods of sorting and sifting (Dornyei, 2007). Common codes were grouped together to generate a theme that addressed a common challenge or consequence. The examination papers, on the other hand, were analysed through

content analysis to identify the targeted skills and test tasks used to assess those skills. The test tasks, test types, integrated and discrete tasks, and the targeted subskills of language being assessed were analysed in terms of frequencies which were, then, described qualitatively.

Reliability and Validity

Reliability and validity are equally significant for qualitative studies as they are for quantitative research, however, the approach to address these issues in qualitative design differs. Maxwell (2010) argued that researcher's bias and respondents' reactivity are two potential threats to validity which can be minimized with the help of several remedial measures. The researchers in the present study attempted to strengthen the validity of the findings, as suggested by Maxwell, through gathering rich data, validating respondents' responses, and data triangulation.

Results

The findings indicate that teachers-cum-assessors face multiple challenges in language assessment. Although, the teachers employ several strategies to cope with the challenges, the consequences are evident on their restricted language assessment practices.

Challenges in Language Assessment
Teachers' Assessment Literacy

A number of challenges emerge from the insufficient assessment literacy of language teachers. Their knowledge about the assessment methods and principles is limited to traditional approaches in language assessment. '*Our course teachers have some arguments related to assessment techniques because they are may be unaware of such techniques or not exploiting those techniques in their classroom*' (P1). Another teacher emphasized the scarcity of opportunities to train themselves in language assessment: '*we don't have any training programs, workshops or seminars on language assessment as such*' (P6). Consequently, teachers make potentially bad decisions as they shared: '*because the ratio of the students who are coming from matric background is more than those who are coming from A level so we need to compromise on them. It's a truth that students from A level are not increasing anything in their knowledge*' (P5).

The higher education institutes tend to practice collaborative assessment practices where a group of teachers collectively design the

paper. However, '*every teacher has a different way of assessing the students*' (P8) and assessment literate teachers have to face reluctance and criticism from their colleagues: '*I know so many things but I cannot apply all of them because we are designing the paper collectively*' (P5). Moreover, '*there is a gap between the teaching practices, the curriculum and assessment; the way we practice teaching we do not assess in the same way*' (P2). This gap also highlights the need of professional development of language teachers.

Learners' Diversity

The limited assessment literacy is coupled with large class size and learners' weak language proficiency. '*The crowded classes in the institution is one of the constraints because it is impossible to give feedback to individual students when there are more than fifty students in the class*'; this way students lack language competency and at the end of the day the teachers cannot do anything* (P2). Majority of the teachers reported that a major problem in assessment is the '*absence of basic skills of language among students*' (P10) and '*low proficiency in English*' (P9). '*Poor diction, lexico-grammatical deficiencies, lack of relevant content and knowledge about the topic*' pose a serious challenge for language assessment. These deficiencies in learners' language lead to the '*challenge of whether to give marks on the concept or to penalize for demonstrating wrong English*' (P3). The diversity of students, in terms of educational background, is also a challenge for language teacher to manage in assessment. '*The students who are coming from different educational background have multiple problems and we need to cater all of them in one class. So, that again creates problem while we are making the paper*' (P5).

Other challenges associated with learners are their low motivation, lack of confidence and exam-oriented approach. According to the teachers, '*majority of undergraduates have poor tendency to participate*' in class leaving language practice '*at low pitch*' (P28). Furthermore, '*lack of confidence and fear of speaking in a second language*' (P12) have also been observed in learners that hinders the assessment. In addition to these, the '*students focus only on gaining marks*' (P10) leaving no scope for formative or other approaches to assessment.

The teachers also shared that it is very difficult to justify the scores to the learners. '*They mostly do not accept the marks*' (P1).

Especially students coming from A levels background who are assessed on isolated grammar exercises: '*they always have these arguments that they didn't get sufficient marks as they got in their previous examination*' (P3). The major reason of this unsatisfactory reaction from the students is students' lack of awareness of teachers' expectations and scoring criteria. '*The students are always wondering what are the expectations of the teacher then when display the marks, there is lot of confusion and students are not convinced*' (P8).

Test Construction and Design

Absence of valid and reliable tests for second language users' context was identified as another significant challenge for language assessment. The teachers stated that '*getting standardized tests for second language speakers which have already been tested for validity and reliability are very difficult to find because generally available tests are for first language learners and at times prove too difficult in our context*'(P17). Similarly, locating input that has topics, issues and speakers from Pakistan or local culture is also difficult to find. As a result, '*developing test items to yield reliable results from students of diverse linguistic background and varied language proficiency*' (P9) becomes a major challenge for teachers in language assessment. In addition to these, other challenges pertaining test design are '*keeping test items aligned with the course objectives*', '*keeping balance in different skills*' and '*keeping format of all courses aligned throughout the university*' (P9).

Institutional Obligations

Teachers also face problems associated with the administration of their institutions. '*The system does not provide us space for formative assessment*' (P15). '*Unjustified marks allocation by the authorities*' (P10) and their non-cooperation in case of deviation from allocated marks division (P16) also make language assessment more difficult for teachers. One of the teachers shared when she attempted to differed from the traditional practices of assessment, '*the whole paper was actually changed as per the management demands*' (P3). '*Institutions requirements are very important. (Teachers) have to meet institution criteria*' (P4). Consequently, the inflexibility of administrative and managerial staff hampers language teachers' attempt to integrate new and alternative methods of assessment into their existing practices.

Time and Resources

The respondents also underlined some challenges pertaining to practicality that they face in their assessment practices. One of the common challenges identified by the teachers was time constraint, '*especially in assessing speaking skills for a large number of students*'(P19) and '*checking and giving feedback*' (P30). A relevant issue is of reliability when the teachers have large number of scripts to score. '*Being a human being it is very difficult to keep the consistency when there is a very large number of papers to be checked, this cannot be checked in one day so cornerstones of testing do get affected*' (P8). In the same way, unavailability or scarcity of resources is a common challenge faced by most of the teachers (P5, P8, P22 and P26). Resources and equipment specifically for listening and speaking assessment are generally not available to the teachers (P10, P19 and P22).

Strategies to Cope with the Challenges

The teachers employ several strategies to cope with these challenges. To deal with the scarcity of available resources, some teachers '*try to produce resources*' on their own (P8, P22) and '*mail listening materials to students a day earlier*' (P9). Whereas another teacher uses '*observation and post-listening activities*' in order to address the dearth of valid and reliable tests for listening and speaking (P22). It was also suggested by a teacher that '*there should be earlier decisions made for assessment like what resources are available*' (P19). To manage time and large class size, group work and short targeted tasks are used by many teachers: '*I try to focus on individual… relevant tasks*' (P29) and '*skills based questions*' (P5). '*The problem of size is handled through the group making*' (P22). Likewise, to motivate learners, '*a bond of friendliness*' is created (P12), whereas another teacher '*take(s) them to the real world so they realize how important it is*' (P8).

Similarly, teachers '*design tests which match the context, situation and level of students*' (P18); and to ensure the alignment of test takers' weak language proficiency and advanced course outline, teachers '*make an effort to incorporate small exercises and activities to cater to their basic gap*' (P20). On the other hand, the issue of subjective scoring is dealt with the use of predefined rubrics. '*Rubrics help to assess it easily (and with) focus*' (P3). Also, teachers '*make descriptors according to the objectives of assessment and then mark accordingly*' (P3). These

strategies indicate that teachers' strategies vary in accordance with their personal and institutional context. As rightly pointed out by one of the teacher that '*improvisation is the only way the teacher is left with to cope with these challenges'* (P10).

Consequences: Ineffective Assessment Practices
Demoralizing and Negative Effects

Responding to the questions if these constraints affect their assessment practices, the respondents admitted that they have negative and demoralizing 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 provides convenience'* (P6).Similar observations were made during the analysis of the examination papers. The analysis revealed that the assessment practices in higher education are not aligned with the latest advancements in the field of language assessment and are limited in terms of targeted language skills and sub-skills and choice of test tasks.

Selective skills and sub-skills

Over all, writing skill dominated the assessment practices although reading was also assessed through few sub-skills. However, the assessment of reading skill was limited to comprehension of explicit and implicit information, inference of meaning of unfamiliar words, identification of topic sentence and suggestion of title. Writing, on the other hand, was assessed mostly through extended production tasks such as essays, reports, and reports. These tasks encourage rote-learning and do not reflect the process of writing. Some of the writing tasks also required students to insert topic sentences and thesis statement. Listening and speaking were almost non-existent in the assessment of these courses. There were only three instances of listening assessment in the collected papers which assessed information extraction through scanning and inferring speakers' intention. Whereas speaking was restricted to presentations only and not a single task was designed for speaking in the final examination. It is noteworthy that the scores of these assessment practices are interpreted for an overall ability of using target language including all four skills.

Redundant Test Tasks

Similarly, various issues were found in the selection of the test tasks that need reconsideration. Although the examination papers showed a variety of test task used to assess language, there were some common trends identified in the selection of test task type regarding specific language skills or element, for example, short answer and MCQs for reading comprehension, and (narrative and descriptive) essays, letter, and short reports to assess writing. To assess listening comprehension, mostly close-ended items were designed and assessment of speaking was restricted to presentations only. Such limited choices in test tasks fail to elicit a representative sample of test takers' language performance. Moreover, grammar and vocabulary were assessed in isolation using decontextualized and discrete-point testing that contradicts the notion of authentic assessment.

Target Language Use Domain and Integration

Such limited assessment practices imply deeper and more challenging problems. The nature and focus of the test task observed in the examination papers do not reflect the target language use (TLU) domains that may result in negative wash back. Additionally, the targeted skills were not integrated in the tasks and were assessed in isolation. Such isolated tasks do not reflect individual's language proficiency truly. Several test tasks in the collected papers were also identified that assessed test takers' linguistic knowledge only. These tasks were designed to elicit information about the skills rather than its use. Thirty examples of such test tasks were observed during the analysis. These tasks elicited knowledge about writing, speaking and vocabulary. Although, competence is essential for language performance but successful completion of these tasks does not reflect the true language skills and communicative ability of the test takers.

Misalignment between Teachers' Beliefs and Practices

The consolidation of the analysis of questionnaires, interviews and examination papers also exposed misalignment between teachers' responses and their practices pertaining targeted language skills and sub-skills. Most of teachers were of the view that all major language skills are important and assessment should incorporate all the skills equally. Nevertheless, the analysis of examination papers showed little agreement to their argument; only selective skills were assessed in the papers.

Likewise, majority of the teachers identified the learners, their language proficiency, and their background knowledge as one of the factors of consideration while test design. However, the topics for reading and writing tasks in the examination papers revealed many topics that are not relatable to the second language learners of Pakistan.

Discussion
Challenges in Language Assessment

One of the most common challenges highlighted by the teachers in the current study was their limited assessment literacy. This reiterates the findings of Sheehan and Munro's (2017) that teachers in Pakistani context lack sound knowledge and understanding of assessment methods due to scarce opportunities for assessment training. According to the respondents, they also face challenges in assessment due to their learners' weak language. Similar results were found in Lumadi (2013) where the teacher's reflection was required to accommodate the diverse groups of learners. Although language assessment is about testing individuals' language proficiency, but it is concerned with the planning, design, administration and scoring as well as the inferences made from these scores. However, in the present context, the teachers consider the weak language command as an obstacle in their assessment. They are unable to realize that assessment itself can be utilized to inform as well as improve their command over the language and is not restricted by test takers' language competence.

Administrative staff of higher education institutes is generally found to be inflexible towards new and innovative approaches to assessment. They emphasize the traditional methods and process for language assessment. The findings are in line with Bresciani (2010) that highlighted the indifferent approach of administration to invest in assessment. Since their expertise does not lie in language assessment, they are unable to realize the significance and impact of decisions, major or minor, made for assessment procedures on test takers and their future. Such non-cooperative and rigid behaviour of the authorities adds to the already neglected language assessment practices. It also implies that an individual teacher may not be able to bring changes in the assessment until it is endorsed and supported by the relevant administrative bodies. Reforms in the assessment are the need of the time but implementing them can be complicated since the stakeholders involved in the real

world may turn out to be different than the one anticipated by the test developers or ELTs in the present context (East, 2015).

Language acquisition takes place gradually; therefore, it must be assessed in the process. However, the teachers in the present study tend to rely more on summative assessment reinforcing the findings of Kurebwa and Nyaruwata (2013).The summative assessment carried out at the end of the academic term puts pressure on the test takers since all language skills are assessed simultaneously. It might affect their performance as it may exert physical and mental strain on them. Sharjeel and Qazi (2012) presented similar argument with reference to reading skill that reading comprehension, particularly at higher education, requires concentration; however, the constant pressure of other skills influences their reading comprehension adversely. They also argued that summative approach to language assessment fails to achieve skill-oriented performance. In the same vein, Raza (2009) discarded such one-time achievement test as an effective practice for language assessment at higher education level.

Bachman (2014) rightly argued that the field of language assessment is grappling with some persistent challenges as well as new and emerging challenges due to the changes in linguistic and global milieu. However, he also emphasized that the 'greatest challenges' in language assessment do not pertain to complex issues like validity theory, socio-psychological or postmodern theories and statistical models but they arise from the 'real-world' arenas where language assessment is used to inform decisions. This argument is clearly reflected in the challenges pointed out by the respondents in the study that are basic and rooted in ground realities.

Consequence: Ineffective Assessment Practices

The findings indicate that the current assessment practices for English courses at undergraduate level need improvement in many areas. To begin with, the examination papers showed multiple examples of competence based questions without any integration with language performance. Such reliance on knowledge-based questions to assess the ability to use target language threatens the validity of the tests. Validity is an important touchstone in assessment literature and one of the essential principles of language assessment (Green, 2014). However, grading test takers on the knowledge of the language and interpreting these grades in terms of their language ability raise questions on the

validity of the test tasks. Integrated and communicative test tasks help strengthen the construct validity as well as the authenticity of the tasks. Raza (2009) also recommended assessment of communicative aspects of the target language for higher education language assessment to ensure the reflection of real-world communication in the assessment practices.

The findings also show that the design of examination papers is dominated by reading and writing skills with little or no provision for listening and speaking. Since the decisions about the test takers' overall language abilities are made based on the performance on these papers, the papers need to incorporate all four language skills and their sub-skills equally. The fact that teachers are constrained by lack of resources and time does not undermine the importance of listening and speaking skills for real-world communication. The restricted selection of language skills and sub-skills result in construct underrepresentation (Grabe& Jiang, 2014). On the other hand, the weightage given to reading and writing skills also needs reconsideration. Although, the teachers set tasks to assess reading and its sub-skills, it was found that writing tasks outnumbered reading tasks to a large extent. Only one or two questions are reserved for reading in comparison to four or five writing tasks. Such dominance of writing tasks results in memorization and reproduction of readymade compositions since it is not feasible to plan, write and revise four to five different kinds of genres in two to three hours Such writing assessment practices encourage the learning of formulaic aspects of writing and discourage critical thinking, a central need in academic discourse (Cummings, 2013).

Luoma (2004) argued that speaking is typically assessed using one-to-one interview techniques. However, in the present context, the findings revealed that speaking assessment is restricted to presentations only. Presentations are prepared monologues in nature that contradicts the real-world language use - mostly carried out in interactions among two or more people. Presentations are part of academic and professional needs but it is not the only genre required for successful communication. Speaking should be assessed through individual as well as pair and group activities to reflect the real language use and to prepare the learners for communication in real-world settings. Hidri (2014) argued that assessment practices in Pakistan were found to be relatively backward when compared to the contemporary language assessment practices abroad. Therefore, Kamrani (2010, cited in Khattak, 2012) recommended introduction of a variety of assessment methods and tools for language testing for the learners to develop the language fully.

Conclusion

Based on the analyses and the interpretations of the collected data, the study concludes that there are various factors, such as teachers' limited assessment literacy, learners' diversity and weak language proficiency, test design, institutional obligations, and practical issues – large classes, time and resources – that pose challenges to the teachers in language assessment. Also, these constraints restrict the teachers in designing and implementing effective assessment of language in higher education. The restrictions result in poor choice of test tasks and limited range of skills and sub-skills in the examination papers that fail to measure learners' language ability accurately. The results also indicate towards the fact that the teachers face similar challenges across universities irrespective of their qualification or type of institution. However, these challenges can be turned into accomplishments for our education system if language testers are willing and capable to apply the skills and knowledge acquired in the past in the field of language assessment (Bachman, 2014).

Recommendations

The authors suggest following recommendations in the light of the findings:

1. The teachers must be trained in the language assessment to equip them with the language assessment knowledge and skills. Assessment literate teachers will be able to design more reliable and valid test tasks and will be able to score objectively and provide effective feedback.
2. The assessment of English language also needs to broaden in its scope to ensure that it caters to the diverse proficiency and preferred learning styles of the students. More varied test tasks and test types will ensure that students' performance in the language assessment is reflective of their true language competence and ability.
3. The teachers must also be encouraged to use more integrated test tasks in their language assessment practices ensuring authentic and valid assessment of language.

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 Stoynoff, 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.

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), 62-68. Retrieved from http://webcache.googleusercontent.com/search?q=cache:_QPdDyiyp zMJ:www.openscienceonline.com/author/download%3FpaperId%3 D1281%26stateId%3D8000%26fileType%3D3+&cd=1&hl=en&ct= clnk&gl=pk

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/p-saeduc/New_Folder_1/6_Assessment%20in%20schools%20in%20P akistan.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. Retrieved from http://www.gjournals.org/GJER/GJER%20PDF/2013/September/051 413610%20Kurebwa%20and%20Nyaruwata.pdf

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.

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

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

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 research.* 16(3),168-17710.4314/zjer.v16i3.26046

Popham, W.J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4-11.https://doi.org/10.1080/00405840802577536

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.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom Assessment Practices and Teachers' Self-Perceived Assessment Skills. *Applied Measurement in Education, 16*(4), 323–342.