

ISSUES IN ENGLISH LANGUAGE ASSESSMENT IN SOUTH PUNJAB: DEVELOPMENT OF A TRAINING MODULE

Fria Khan, Dr. Muhammad Anwar Farooq, Waheed Shahzad, Hafsa Saboor, Muhammad Muneebullah

Institute of Humanities and Arts, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan fariazulfiqar573@gmail.com

Institute of Humanities and Arts, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan anwar.farooq@kfueit.edu.pk

Institute of Humanities and Arts, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan waheed.shahzad@kfueit.edu.pk

ORCID ID: 0000-0002-2237-8237

Institute of Humanities and Arts, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan hafsa.saboor17@gmail.com

Institute of Humanities and Arts, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan muneebullahryk786@gmail.com

Corresponding author: waheed.shahzad@kfueit.edu.pk

Abstract

This research explores major concerns in assessing English language among university teachers of English in South Punjab, Pakistan, and aims at coming up with and piloting a context-sensitive training module to enhance the Language Assessment Literacy (LAL). The study applies mixed-method research design, involving both data collection based on the Likert-scale questionnaire of 50 teachers as well as the semi-structured interview/focus group data of ten to twelve teachers; quantitative analysis will be done through the descriptive statistics method whereas the qualitative analysis will be conducted using the manual thematic analysis technique. The results have shown that the assessment practices are still more traditional with a number of teachers being content to depend on written tests and moderate with respect to oral assessment with weak and inconsistent rubric use- which implies low levels of standardization and transparency in grading. The report by teachers cites significant operational limitations particularly in creating equitable evaluation in large groups and scarcity of professional development opportunities which support systemic rather than individual failures. The qualitative themes also point to the gaps in assessment literacy such as not having enough before service preparation, lack of confidence in creating a valid and fair assessment (particularly speaking/writing), and not having a clear picture of formative purposes of assessment. According to these requirements, the study suggests a specific training module on the principles of the sound assessment, the use of formative and summative in big classrooms, the rubric design and use, the effectiveness of feedback, and the attainment of culturally responsive assessment based on the multilingual assessment, which will be measured using the pre/after formats and feedback data.

Keywords: *Language Assessment Literacy (LAL), English language assessment, South Punjab, higher education, teacher training module, formative assessment, contextualized assessment*

1: Introduction

English has acquired significant significance in the subcontinent especially as the language of international communication and modern education since the British rule in the 18 th century. Language helps people to communicate their thoughts by speaking, listening, reading, and writing

(Rauf and Saeed, 2019; Al-Mansoor, 2014). The Pakistani higher education students are experiencing the same difficulties as students in other developing nations, particularly when they are in giant classes of English (Nino and Paez, 2018). Efficient teaching needs quality instructions, assessment literacy, and learning conducive conditions (Chishti, 2011; Inbar-Lourie, 2008). Because of economical limitations, downsizing of the class is not feasible; thus, the present research is grounded on the participatory action research to enhance teaching and learning in big classes of English language in Pakistan in order to enhance equity, cooperation and quality of education.

The teaching of English language in South Punjab is experiencing severe difficulties such as cultural bias, poor standardisation, low teacher preparation and failure to match with the learners needs. To cope with these problems, contextualised evaluation tools, standard frameworks, and thorough teacher training should be implemented to guarantee effective assessment fairly.

Research Objectives

- To investigate the concerns in evaluation English language among English teachers in South Punjab.
- To find out certain problems and gaps that teachers experience when implementing effective assessment practice.
- In order to evaluate training to test the developed training module and test its efficiency using pre and post training evaluation.

The study will also seek to build a holistic insight into the issues encountered by teacher of English language in South Punjab in the context of assessment. It is hoped that the results will detect deficiencies in the assessment competency of the teachers, and use it to design specific professional development programs (Guskey, 2002). The resulting training module is expected to enhance the assessments literacy of instructors, resulting in the more reliable, valid, and effective classroom assessment practices, as well as evaluation of student learning outcomes.

2. Background of Study

The growing professional and academic requirements combined with globalisation have ensured that English language assessment has become a major concern in South Punjab in Pakistan. Since English is the most important language in higher education and job, there is a need to have proper measurement of language proficiency. Nevertheless, evaluation in the area is challenged, which leads to unequal learning outcomes and poor communicative skills. All these problems are as a result of insufficient teacher training, the use of the old evaluation techniques, the socio-economic differences, and the imbalance between the goals of the curriculum and assessment activities.

One of the issues is a lack of assessment literacy of teachers. Numerous English teachers are not professionally trained in language testing and, instead, use the traditional strategies, i.e., rote memorisation and grammar-translation, instead of communicative and formative assessment that involves the measurement of real use of language. Lack of pre-service and in-service training is long term supports bad practices and the distance between the theory of assessment and practice in classrooms. As a result, tests do not accurately represent the real language skills of students resulting in a mistaken choice of instructions. These problems are enhanced by socio-economic inequality. The urban institutions generally possess higher resources and exposure to English and the rural schools can have little access to skilled teachers, digital tools, and new evaluation systems. Another cause of grading disparities and misunderstanding between the teacher and the learner is curriculum-assessment misfit and the absence of standardised structures in the various institutions.

Such challenges need to be taken care of through systematic reforms such as in-service teacher training, alignment of charts, and frameworks of standardised assessments. To make sure that there is a valid, reliable, and equitable English language assessment in South Punjab, the assessment literacy needs to be strengthened through a continuous professional development.

2.1 Teacher Assessment Literacy

Literacy teacher assessment plays a critical role in proper evaluation of the English language. In South Punjab, informal assessment is an inconsistent and unreliable practice due to a lack of formal training in assessment by many teachers there (Rozilat and Matore, 2023). The absence of theoretical understanding of validity and reliability makes educators use old approaches thus impacting negatively on fairness, accuracy and learning outcomes of students. Jeong (2013) discovered that teachers who have undergone assessment training have better evaluation practices and instructors who have not been trained tend to ignore the principles of test-design, thus measuring language ability inaccurately. Herein lies this necessity to have the Language Assessment Literacy (LAL) training in South Punjab. Research indicates that teachers with LAL incorporate the use of formative, performance, and self-assessment modes, replacing rote testing and allowing a more objective and fair evaluation of the communicative competence of the students (Herrera and Macias, 2015; Giraldo and Quintero, 2019b). In South Punjab, English teachers are not well assessed with regard to literacy because specialised assessment training programs are not available. Therefore, they use old-fashioned and inconsistent assessment practices that do not influence the formative assessment and new communicative learning requirements. The problem of assessment in South Punjab is closely connected with the level of economy. The scarcity of resources is a limitation to teacher training and access to assessment resources, and rural students have little exposure to English. Standardised tests tend to place underprivileged learners in a disadvantaged position, which creates the necessity to adopt context-sensitive assessment practices (Khan and Khan, 2016).

One of the effective methods to enhance assessment literacy is to integrate the formative assessment practices into teacher training. Peer review, reflective journals, and classroom discussions are the techniques that assist teachers to track the learning and improve instruction (Clark, 2012). The assessment identity of teachers also depends on the professional development that promotes alignment to the contemporary language pedagogy (Looney et al., 2017). In South Punjab, wholesale evaluation literacy teaching must be integrated in the pre-service and in-service training, which will involve validity, reliability, fairness, and formative and summative approaches. Ongoing training, technology, collaboration, and context-sensitive education are mandatory to provide equitable, reliable, and effective evaluation of English language.

2.2 Challenges in English Language Assessment

The economic conditions play a great role in determining the quality of education in South Punjab which means that assessment outcomes are unequal. Economically disadvantaged rural students are unable to achieve high marks in English tests because they have less to offer and do not have enough qualified teachers and do not get to interact with English outside the school (Habib, 2013). Standardised tests tend to favor such learners as giving them an opportunity to perform on the assumption that they have been exposed to languages they lack, further enhancing the performance differences between the rural and urban populations. The gaps in assessment in South Punjab are associated with the insufficient resources in the rural institutions such as textbook and digital resources as well as trained educators. The inability to assess communicative skills through paper-based testing and deficiency in technology are obstacles to the disadvantaged students (Tayyaba,

2012). Opportunities of English learning are influenced by socioeconomic factors. Wealthy students have more exposure with the help of private schools and media, whereas poor-income students have to rely on the limited school education and this results in their unfair performance on standardized English testing (D'Elia et al., 2023; Siddiqui and Gorard, 2017). Standardised English testing uses the same measures of proficiency, which disregards the existence of socio-economic and linguistic variations. The tests also benefit the urban students who have a higher exposure to English at the expense of the rural learners. In South Punjab, both native Saraiki and Punjabi speakers have some extra obstacles because the evaluations do not consider realities of the local language, and the actual linguistic skills of students are underestimated (Muhammad et al., 2024).

To overcome socio-economic gaps, alternative models of assessment should be used in context. Oral, portfolios and contextualised assignments are non-standardised methods that are better indicators of English proficiency. Bilingual tests also provide inclusive assessment and learning since students can show their knowledge in both languages, English and their native language (Awad and Adwan, 2024). An important role of teachers is to decrease the gaps in socio-economic assessment. Context-responsive assessment options through targeted training, formative assessment, customized feedback, and government-sponsored access to digital assessment tools allow fairer underprivileged assessment practices (Shahzad, et al., 2020; Munir et al., 2023; Mahara, 2024).

2.3 The Role of Training Modules in Improving Assessment Practices

The effectiveness of the English language assessment in South Punjab is very much related to the competencies of the teachers. Studies indicate that systematic training packages contribute to the development of assessment skills in educators, which results in better performance among students. Professional assessment needs specialised training as a way of boosting the aspect of validity, reliability and relevance with the learning objectives (Sainsbury and Walker, 2008). Measuring training effectiveness entails testing professional development effectiveness, effect on student performance, and obstacles to training implementation.

In South Punjab, teachers tend to use rote and grammar exams since they are not adequately trained (Cheema, Siddiqui and Masood, 2022). It is demonstrated that systematic, specific training enhances assessment skills, confidence, quality of feedback, and correspondence with communication language goals (Furtak et al., 2016). Organized training courses enhance the assessment literacy of teachers by instilling the knowledge of validity, reliability, and equity (Popham, 2011). Without training, teachers are guided by the traditional testing that is guided by personal beliefs (Adie, 2012). Studies indicate that training enhances formative assessment practices such as peer and self-assessment, which enhances the quality of evaluation and student participation (Andersson and Palm, 2018). The training on assessment enhances the performance of students by enhancing evaluations by teachers. Increasing language learning is achieved by having the trained teachers design curriculum-related assessments that have significance in providing meaningful feedback (Brookhart, 2011). Literature indicates that structured instruction results in the promotion of alternative assessments and feedback-oriented classrooms, enhanced student engagement, comprehension, and motivation and the transformation of assessment as a learning tool (Wiliam and Thompson, 2017; Cisterna et al., 2025).

Structured assessment training modules despite their advantages have significant impediments in South Punjab. Access is limited due to low levels of institutional support and funding partly in rural schools. Teacher education programs focus on general pedagogy as opposed to assessment

literacy, and evaluation skills get underdeveloped (Naz, Jabeen and Rashid, 2020). Change also poses a barrier to adoption because teachers who are used to the traditional approach view the alternative assessment as cumbersome or time-consuming (Mansory, 2020). To overcome the assessment training challenges in South Punjab, the assessment literacy must be considered in teacher training, long-term governmental financing, multifamily cooperation, and digital training modules and constant professionalization of teachers through seminars, refresher courses, and incentive-based certification are the keys to the successful and long-term solution of the assessment practice problems.

2.4 Large Class Sizes and Assessment Constraints

High classes lead to negative effects on the quality of instruction and validity of assessments all over the world (Glass et al., 1982; Zyngier, 2014). In South Punjab, English classes are overcrowded, and they do not give the teachers enough time to interact with the students and assess them. This paper discusses how the rising number of students in classes puts a burden on assessment procedures, raises the workload for teachers, lessens student engagement, and warrants a more advanced assessment procedures and teacher training. In the big classes, educators use written tests that are grammar-based, and they do not focus on oral and communicative abilities (Khan and Khan, 2019; Akram, 2017). Speaking tests are obstructed by time, which restricts language growth. It has been found that interactive oral assessment can increase proficiency thus, it is important to have managed, structured ways of assessment in overcrowded classrooms (Stoyhoff, 2008; Younas et al., 2024).

The high number of students in classes is a major contributor to the workload of the teacher and denies them the opportunity to give detailed feedback and formative assessment. This leads to teachers using standardized, summative examinations that limit individual learning (Makhambetova et al., 2021). It has been found that in overcrowded classrooms, syllabus is done instead of assessed (Bennett, 2011; Ali et al., 2022; Sarwat et al., 2024a). This is further aggravated by assessment literacy training scarcity especially in the rural areas in Pakistan where the instructors have not undergone formal training on language testing (Akram et al., 2020; Sarwat et al., 2024b). The teachers, therefore, rely on rote-based techniques, which result in inconsistent grading and poor diagnostic capability (Iqbal et al., 2020). The development of professional life is a continuous process that should foster valid, formative, and equitable assessment practices (Saleem et al., 2021). Large classes have a negative impact on the engagement and learning results of students because they decrease the interaction between teachers and students and individual feedback (Finn et al., 2003). Congestion promotes passive assessment, which restricts the ability to assess language in an interactive manner, which facilitates language development (Li and Gao, 2015). These limitations in South Punjab encourage memorization, test anxiety, and limited feedback in time, which ultimately undermine fluency and comprehension of English among students, as well as their general proficiency in English (Segool et al., 2013).

The South Punjab is the cause of large classes due to the low resources, lack of trained teachers, and funds to develop professionals especially in rural public schools (Likuru and Mwila, 2022). Lack of strong standardisation and cultural beliefs based on exams result in variation in assessment practices, which believe in grades yet do not focus on skill development (Al-Tameemi et al., 2023). These obstacles are exacerbated by the lack of technological integration since most schools do not have the digital infrastructure, which promotes the use of the old-fashioned paper-based assessment and restricts the innovative evaluation approach.

2.5 Curriculum Gaps and Resource Limitations

The problem of effective assessment of English language is hampered by outdated curricula and limited resources in South Punjab. Existing courses lack the focus on modern pedagogy and technology, which limits ways of evaluation and outcomes of learning (Razzaq, 2023). This is necessary to address the gaps in the curriculum and allocation of resources needed to enhance the quality of assessment and effectiveness of education. According to Razzaq (2023), there are fundamental flaws in the English language teaching, which also applies to the educational strategy of South Punjab. The existing curricula focus on grammar and memorisation over communicative competency, and students are not ready to take any practical speaking and listening evaluation (Khan and Khan, 2019). The methods of assessment are conservative, as they are based on paper-based assessments that demotivate the formative, peer-based, and technology-assisted evaluation (Thankachean, 2024). The rural learners are disadvantaged by the lack of exposure to different test formats at the tertiary level in the high stake's exams like IELTS and TOEFL (Desalegn et al., 2023). It is hence necessary to have curriculum reform that incorporates communicative skills and current assessment strategies.

The scarcity of resources in South Punjab increases the gaps in curriculum by narrowing the opportunities to use audiovisual resources and interactive technologies that are needed to develop the listening and speaking skills. The use of old-fashioned textbooks and lectures undermine the communicative proficiency and test validity as students are not ready to face new language assessment (Arain et al., 2021; Heikonen et al., 2017). South Punjab has a bad socio-economic status that makes assessment hard especially to poor students. The low literacy rates of parents and poor exposure to English make people less proficient in the language and less effective in assessments (Ali et al., 2022; Ali et al., 2009). Teacher training on the current assessment techniques is also constrained by the budget, which causes the use of traditional testing that does not take into account all the learner needs (Gallardo, 2020). To eliminate the existing curriculum and resource gaps in South Punjab, there should be overall reforms. These are updating the English curricula using communicative language teaching, incorporating the use of audiovisual and digital testing instruments, and enhancing the literacy of teacher assessment (Brookhart, 2011). Standardised but formative assessment systems must also be consistent (Popham, 2011) and socio-economic support schemes as well as parental involvement programs should also be used to mitigate inequality and enhance the results of language proficiency (Segool et al., 2013; Finn et al., 2003).

3: Research Methodology

The research is based on a mixed-method design in order to investigate English language assessment practices of university teachers in South Punjab. Likert-scale questionnaires yield quantitative data, which reveals trends and issues, and qualitative interviews and focus groups give the data the contextual depth. Combining the two methods sequentially will make the findings statistically sound, contextual and representative of the lived assessment experiences of teachers.

3.1 Population and Sampling

The research aims at the sampling of English teachers in South Punjab (purposive sampling). Questionnaires were filled in by 50 teachers and 10-12 teachers were interviewed. The choice of this context was shaping the limits of resources, the variety of proficiency and assessment difficulties, which guarantee abundant practice-based information on higher-education English assessment.

3.2 Data Collection Instruments

In order to investigate the concerns of English language assessment and in a bid to justify the construction of a specific training module, the research uses two main data collection methods: a structured questionnaire as a quantitative data collected, and a semi-structured interview and focus group as a qualitative data. The reasons behind the choice of these tools were to have a broad summary alongside a comprehensive view of the experiences and views of teachers.

The quantitative tool is a structured questionnaire made up of closed-ended questions.

Quantitative data was collected by use of a structured questionnaire that was administered to university Teachers of English. It investigated the practices, challenges and professional training requirements of assessment using 15-20 five-point Likert-scale items. Demographics, existing assessment practices, institutional constraints and perceptions of assessment literacy development were covered. Semi-structured interviews and focus group discussions will be used to supplement the numerical results with a chosen group of 10 to 12 instructors. These instruments are especially useful when the issue under discussion is to be considered rather thoroughly so that the participants could share their experiences, concerns, and recommendations concerning the language assessment. The interview guide will contain open ended questions which will coincide with the themes that will come out in the questionnaire. Qualitative sessions will be tape recorded with the consent of the participants and transcribed by listening. The communicative nature of these tools will make the participants free to talk and think, providing information which will not be recorded in the fixed-answer questionnaires.

3.3 Data Analysis Procedures

All the information gathered using the structured questionnaire and the semi structured interviews focus groups will be analysed in systematic manner in order to make the findings valid, meaningful and in accordance to the study objectives. The analysis will be performed in two directions, but complementary, i.e., quantitative analysis of the data collected in the questionnaire and qualitative analysis of the data gathered in the interview and focus group transcripts.

The data received with the help of the Likert-scale-based questionnaire will be analyzed by means of the descriptive statistical tools. These include:

Frequencies (how many participants chose each option),

Percentages (distribution of responses across the scale),

Means (average ratings for each item), and

Standard deviations (to understand variability in responses).

Qualitative Data Analysis (Manual Thematic Analysis)

The data gathered through interviews and focus group discussions will be analyzed using manual thematic analysis, based on the six-phase approach outlined by Braun and Clarke (2006). These phases include:

Familiarizing with the data by reading transcripts multiple times,

Generating initial codes from significant words or phrases,

Searching for themes among the codes,

Reviewing and refining themes for coherence,

Defining and naming themes, and

Producing the final narrative report.

3.4 Training Module Development and Evaluation

This paper tries to define the issue of assessment and suggest a training module that can be used to overcome this problem. Based on quantitative and qualitative data, the module aims at filling

gaps in assessment literacy including fair test construction and rubric application to guarantee context-specific and needs-based professional growth of university English instructors in South Punjab.

This module will discuss major areas, which include:

Knowing the laws of sound and appropriate assessment.

Formative and summative assessment to large classrooms.

Formulating and implementing rubrics towards equitable assessment.

Giving valuable and practical feedback.

Multilingual classroom assessment: Culturally Responsive.

The interactive activities will be applied in the training module, which will be piloted with the help of selected instructors. The effectiveness of it will be measured using pre and post-training Likert-scale questionnaires and open-ended feedback, which will assess the change in assessment confidence, knowledge, and classroom application. It is a combined assessment that makes it practical and contextual.

4: Findings and Analysis

In this chapter, the author gives quantitative questionnaire data and qualitative thematic data of interviews and focus groups emphasizing the major patterns, issues, and views about English language assessment practice among university teachers in South Punjab.

Table 4.0

Demographic Profile of Respondents

Category	Gender		Age Group		Years of Teaching Experience		Type of Institution		Assessment Training	
	Freq. (n)	%	Freq. (n)	%	Freq. (n)	%	Freq. (n)	%	Freq. (n)	%
Option 1	Male: 28	56%	25–30: 10	20%	1–5 years: 14	28%	Public: 34	68%	Yes: 20	40%
Option 2	Female: 22	44%	31–40: 22	44%	6–10 years: 16	32%	Private: 16	32%	No: 30	60%
Option 3			41–50: 12	24%	11–15 years: 12	24%				
Option 4			51+: 6	12%	16+ years: 8	16%				

Interpretation:

In the same section, a detailed report is provided on what demographic composition of the 50 who participated in the quantitative aspect of the study is like. These are some of the contextual factors that should be known the answer thereof in as far as English language evaluation practices are concerned.

Gender Distribution

A total of 50 English teachers were the participants of the study (28 males (56%) and 22 females (44%)). Even though there is a slight male dominance in the sample, there is relatively fair distribution which enables representation of both sexes. Such a balance assists in having a just look at the attitude of assessment, based without gender-specific disaggregation, which is out of the reach of this research.

Age Group

The participants were of different ages. The highest percentage was found in the age group of 31-40 years (44%), 41-50 (24%), and 25-30 years (20%). Only 12% were aged above 50. This distribution shows that there is a general representation of middle-career professionals, with the assistance of both the beginning-career professionals and the senior professionals providing a mixture of innovation and experience in the assessment practices.

Years of Teaching Experience

The experience in teaching reflected the age trends. The majority of the participants were having 6-10 years of experience (32%), 1-5 years (28%), and 11-15 years (24%). The minority group (16%) had more than 16 years' experience which indicates that emerging assessment practices are heavily influenced by the early and mid-career teachers.

Type of Institution

Instructors at the public universities were the majority (68%), whereas there were 32 percent at the private institutions. This is due to the fact that South Punjab uses more of the public-sector higher education and this can affect the flexibility of assessment and innovativeness.

Formal Training in Language Assessment

Formal assessment training was only received by 40 percent of respondents, which demonstrates the high level of assessment literacy gap, which ultimately supports the necessity to implement specific professional development programs.

4.1 Quantitative Data Analysis

Section: Current Assessment Practices

Table 4.1.1

Statement1: I regularly use written tests to assess my students' English proficiency.

<i>Likert Scale Option</i>	<i>Frequency (n)</i>	<i>Percentage (%)</i>
<i>Strongly Disagree (1)</i>	4	8%
<i>Disagree (2)</i>	8	16%
<i>Neutral (3)</i>	10	20%
<i>Agree (4)</i>	18	36%
<i>Strongly Agree (5)</i>	10	20%
Total	50	100%
Mean	3.42	
Standard Deviation	1.15	

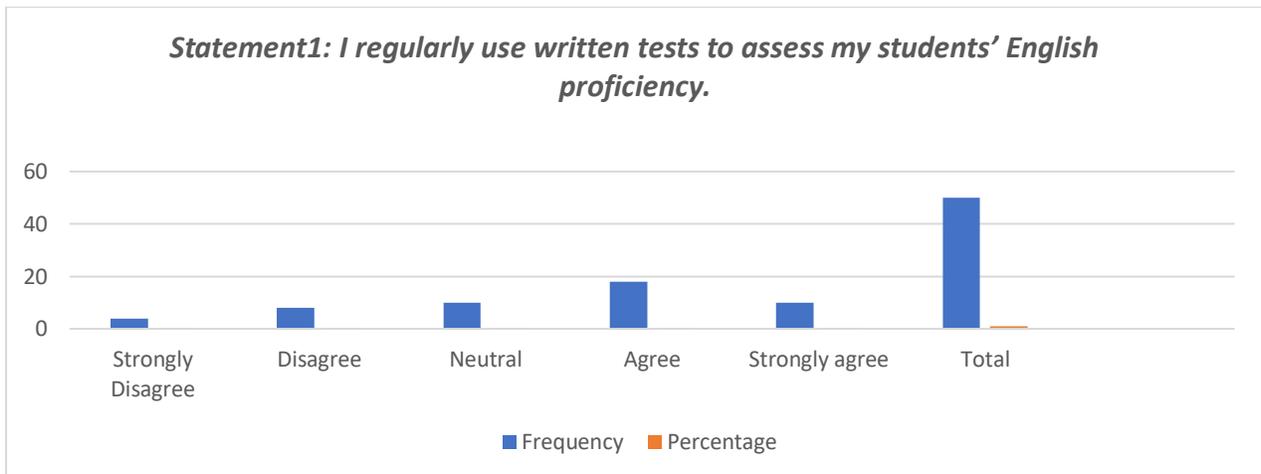


Chart 1

Interpretation:

Findings depict an agreement of 56 percent on the use of written tests, which implies the use of traditional evaluation. Nevertheless, the disagreement and mean of 24% and 3.42 (SD = 1.15) indicate inconsistent practices, and there should be more specific assessment guidelines and institutional variance.

Table 4.1.2

Statement 2: I use oral assessments (presentations, discussions) in my courses.

Likert Scale Option	Frequency (n)	Percentage (%)
<i>Strongly Disagree (1)</i>	3	6%
<i>Disagree (2)</i>	7	14%
<i>Neutral (3)</i>	12	24%
<i>Agree (4)</i>	16	32%
<i>Strongly Agree (5)</i>	12	24%
Total	50	100%
Mean	3.54	
Standard Deviation	1.12	

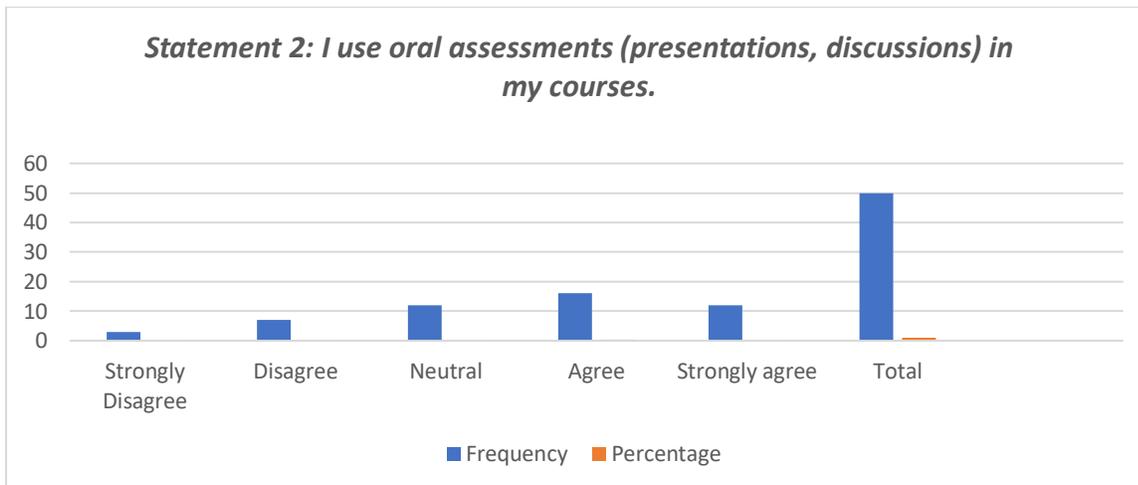


Chart 2

Interpretation:

Moderate integration ($M = 3.54$, $SD = 1.12$, 56% agreement) is observed in oral assessments, but 40% of the oral assessments are neutral or disagree, and this result implies that training and support gaps need to be guided by the institution.

Table 4.1.3

Statement 3: I frequently use rubrics to evaluate student work.

Likert Scale Option	Frequency (n)	Percentage (%)
Strongly Disagree (1)	5	10%
Disagree (2)	9	18%
Neutral (3)	13	26%
Agree (4)	15	30%
Strongly Agree (5)	8	16%
Total	50	100%
Mean	3.24	
Standard Deviation	1.17	

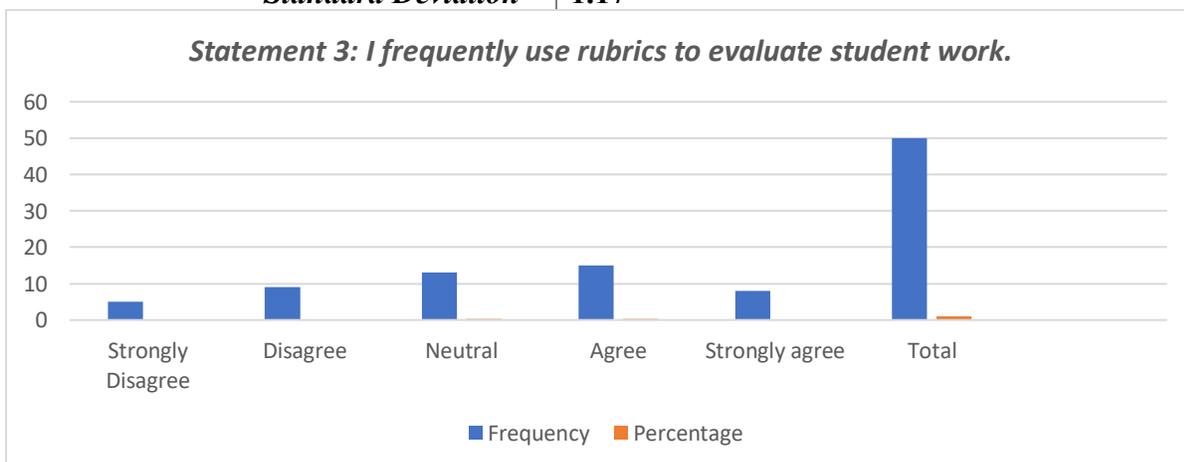


Chart 3

Interpretation:

The rate of agreeing and strongly agreeing to the frequent use of rubrics was only 46% with a significant number of 28% admitting to disagree and 26% showing a no opinion. This signifies the lack of coherent use of transparent evaluation resources such as rubrics that are necessary in fairness and transparency in grading. The average value of 3.24 indicates low or rather a poor use of the rubrics. This may be because they were not exposed, trained, or focused on on rubric design in institutions. By working on this with specific training related to the development of rubrics, it may become possible to standardize assessment practices and enhance the understanding of students as well as teacher accountability.

Section 3: Challenges in Assessment

Table 4.1.6

Statement 6: I struggle to create assessments suitable for large classes.

<i>Likert Scale Option</i>	<i>Frequency (n)</i>	<i>Percentage (%)</i>
<i>Strongly Disagree (1)</i>	3	6%
<i>Disagree (2)</i>	5	10%
<i>Neutral (3)</i>	7	14%
<i>Agree (4)</i>	19	38%
<i>Strongly Agree (5)</i>	16	32%
Total	50	100%
Mean	3.86	
SD	1.10	

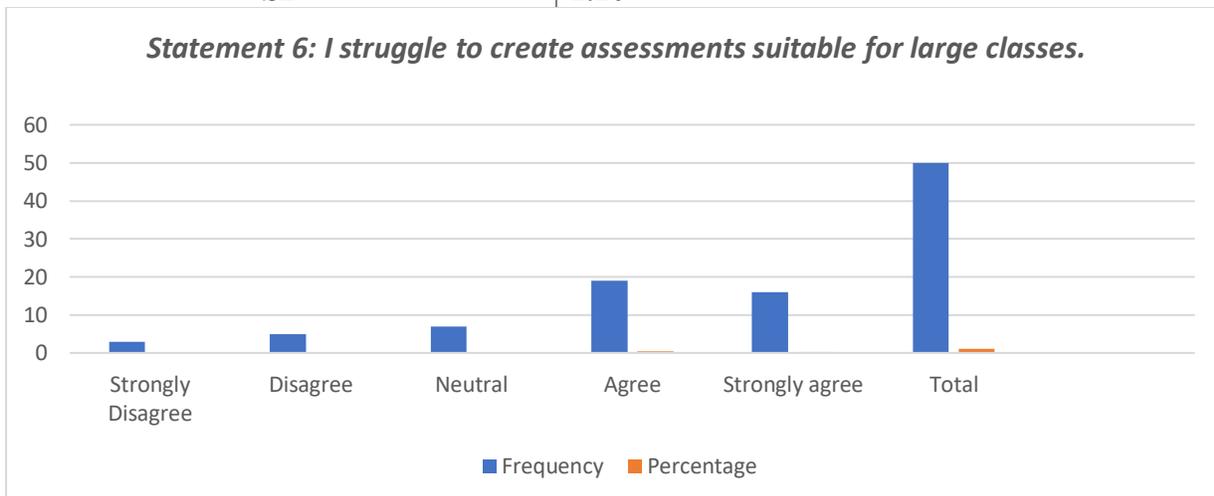


Chart 6

Interpretation:

This question exposes one of the common and unified issues within the teaching fraternity: 38% of the answers were in agreement and 32% were strongly so of the fact that they are unable to create effective assessments when it comes to large classroom environments. It is a high majority and it is indicative of the long-established problem of crowded classroom in higher institutions of learning in places such as South Punjab where the student to teacher ratio is normally high.

Interestingly, 14 percent of teachers indicated the neutral choice, which might imply that they resign or they simply were not aware of other assessment methods when teaching large classes. The reduced percentage 10% not agreeing and 6% strongly disagreeing is most probably instructors in smaller institutions or those institutions with easier means of assessment and teaching support.

This finding is supported by the mean score of 3.86 and small standard deviation of 1.10. It verifies that despite the differences in demographics, the skill to evaluate well in large groups is an urgent issue. This implies that there is an immediate requirement of the training modules that provide the scaled methods of assessments like peer reviews, digital assessments, or group-based evaluations.

Table 4.1.7

Statement 7: I lack access to professional development opportunities in assessment.

<i>Likert Scale Option</i>	<i>Frequency (n)</i>	<i>Percentage (%)</i>
<i>Strongly Disagree (1)</i>	4	8%
<i>Disagree (2)</i>	6	12%
<i>Neutral (3)</i>	5	10%
<i>Agree (4)</i>	17	34%
<i>Strongly Agree (5)</i>	18	36%
Total	50	100%
Mean	3.82	
SD	1.18	

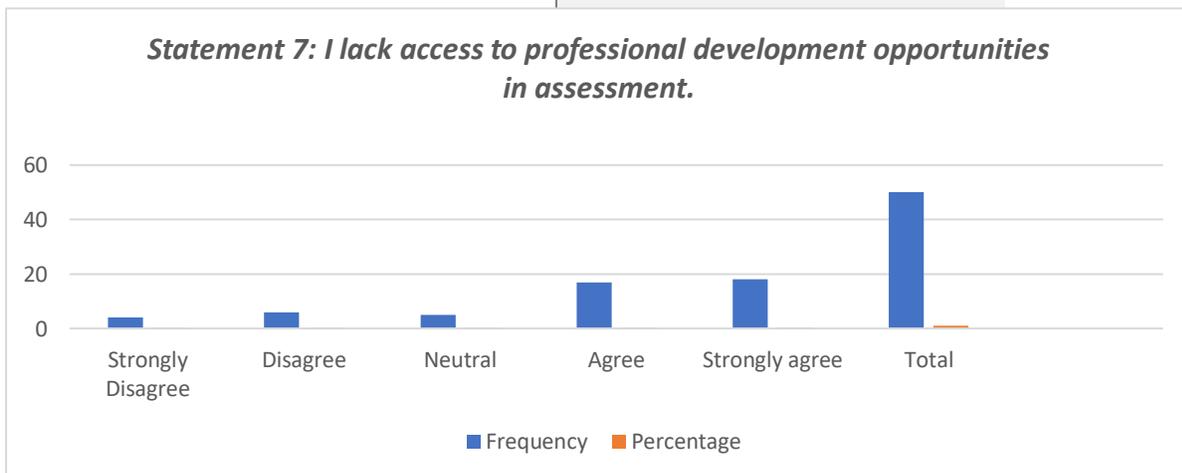


Chart 7

Interpretation:

Two-thirds of instructors (34 percent and 36 percent had the same view, respectively) said they did not have access to professional development. This statistic indicates a structural failure in the infrastructural teacher education, especially in assessment-focused upskilling. Lack of such opportunities may result in the outdated practice, lack of teacher confidence, and insufficient correspondence to the modern pedagogy.

In the meantime, one in ten of the respondents disagreed, and one in eight strongly disagreed, which means that there was a minor proportion that might have gained with the institutional efforts or external training platforms. Nevertheless, it is an exception rather than a rule. The neutral

responses (10%) might indicate some sort of doubt regarding what is considered to be a valid training or even not be exposed to an organized workshop.

The mean and SD of 3.82 and 1.18 are quite agreeable and the variance is small probably due to urban and rural campuses or the differing institutions (public and private). This completes the argument of making language assessment training modules universally and context-specific.

Section 4: Perceptions of Assessment Training

Table 4.1.11

Statement 11: I believe that training in assessment practices is essential for teachers.

Likert Scale Option	Frequency (n)	Percentage (%)
Strongly Disagree (1)	1	2%
Disagree (2)	2	4%
Neutral (3)	7	14%
Agree (4)	20	40%
Strongly Agree (5)	20	40%
Total	50	100%
Mean	4.18	
SD	0.91	

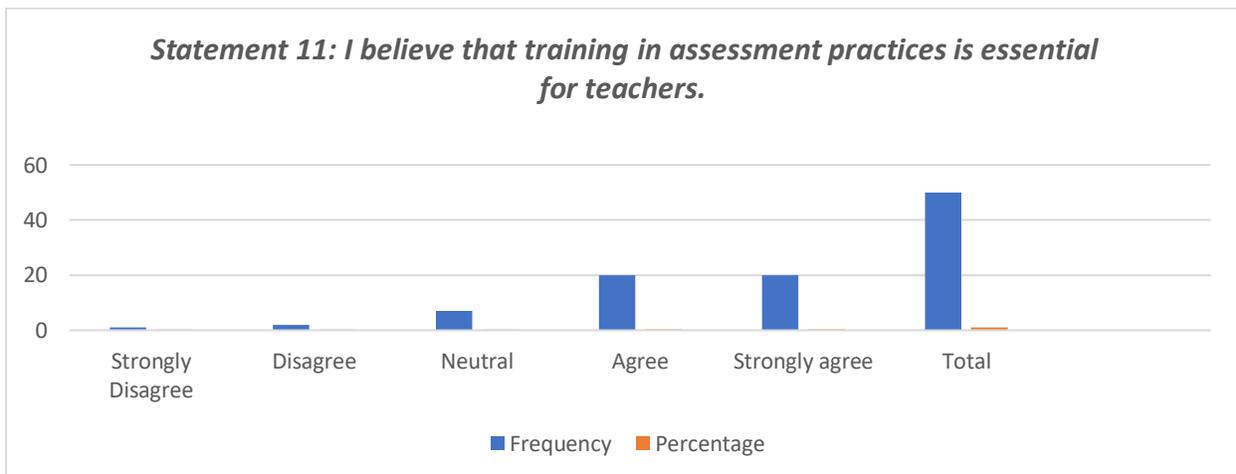


Chart 11

Interpretation:

There was a very impressive 80 percent agreement (40 percent) and strong agreement (40 percent) with the Strong consensus supports assessment training in which the mean (4.18) was high and the variability (SD = 0.91) was low. There is little doubt to support the general acceptance of the value of assessment literacy and prove the significance of the training module suggested.

4.2 Qualitative Data Analysis

The section gives the qualitative results of semi-structured interviews conducted on the selected English language teachers in South Punjab. The participants were chosen purposely out of various institutional backgrounds (10) according to their availability and willingness. Thematic analysis is a manual procedure that was used to identify patterns and categories in the interview transcripts using the six-phase method suggested by Braun and Clarke (2006). The four broad themes were

obtained as a result of the analysis, and each of them was corroborated with the quotes of the participants to make their work authentic and highly saturated in terms of interpretation.

4.2.1 Theme 1: Gaps in Assessment Literacy

Among the most evident themes that appeared in the course of the qualitative data were the general lack of formal training and conceptual clarity of English language assessment. The most common pattern among the teachers irrespective of their experiences was that they had not been given systematic instructions on how to create, analyze, or introspect about assessment tools especially in the language acquisition context where evaluation of speaking, listening and writing involved unique difficulties beyond standardized testing.

4.2.1.1 Direct Voices from the Field

A 12-year-old instructor in a government-funded university taught the middle of the career, and told me:

I have been teaching English more than 10 years, and frankly speaking, nobody ever educated me about how to assess it in the proper way. I simply made-up tests as I considered appropriate. No standard manual or anything like that is given to us. Indeed, I think this will significantly aid me in handling such issues going forward. In fact, I believe that this will go a long way to help me deal with such issues in the future.

Equally, a private institution teacher who was younger and had just joined the institution as an English lecturer commented:

Such remarks are an indication of a basic lack of relationship between instructional knowledge and evaluation assurance. Even the instructors who portrayed to be comfortable with pedagogy admitted that they evaluated using methods that were not rigorous, consistent or had sound pedagogy.

4.2.1.2 Absence of Assessment Literacy in Pre-Service Training

Some of the participants indicated that they were not taught any meaningful assessment modules in their teacher education programs. Rather, the emphasis was on the lesson planning, classroom management and content delivery.

In my M.A. and even M.Phil coursework, assessment was not taught much. We always assumed that we would learn it in the job. But how do you know rightly when nobody has instructed you on what a validity or a reliability is? The reason is that the public university has a bigger population of students who can more easily engage in diverse extracurricular activities. This is because the population of students who are able to participate in various extracurricular activities is larger in the public university.

Such inadequate foundation has resulted in the use of more or less outdated methods of assessment, like essay writing or multiple-choice tests. In consequence, new practices such as rubric-based grading, peer assessment or performance-based assessment are not actively used.

4.2.1.3 The Myth of Experience Equals Competence

One of the most impressive revelations that came out during the interviews with both junior and senior educators is the fact that the many years of teaching experience do not necessarily make them assessment literate.

I have 15 years of experience but it is just recently that after discussing with my younger colleagues who have studied abroad that I came to realize how much I do not know about assessment. I have never used a rubric in my life. However, when compared to other participants, the proportion of the first one is lower. Nevertheless, the ratio of the former one is less compared with other participants. (Participant 2, Public University).

This brings out a significant difference: practice without reflective training might not bring about improvement. Although experience comes with familiarity, it does not necessarily increase better evaluation practices without continuous professional learning.

4.2.1.4 Inconsistent Understanding of Assessment Purpose

Teachers also failed to understand the reason as to why they were testing students other than giving out grades. Few of them expressed a logical sense of formative vs summative assessment and how assessments might be utilized to inform instruction and aid student learning.

I administer quizzes on a weekly basis primarily because it keeps the students at their toes and not necessarily as a way to get the student-teacher relationship on track. To be honest, I do not think I have ever put together a change in my lesson plan depending on the student performance. Therefore, while the films portray Chicago as a repository of crime, the reality is that it stands as a city with one of the highest crime rates. Thus, although the movies depict Chicago as a place where crimes thrive, it is a fact that it is a city that has one of the highest crime rates.

4.2.1.5 Lack of Confidence in Designing Fair Assessments

One of the issues raised repeatedly was the lack of confidence with which teachers felt about their capacity to aim valid, fair, and objective tests, in particular, in speaking and writing aspects. Teachers believed that their evaluation was usually subjective and could be easily inconsistent especially when dealing with a large classroom.

I occasionally assign one student more marks than the other who gives an oral presentation because they speak better English, although I am not really certain I should do that because I should consider the effort made by the students. It does not have a rubric, and it is all in my head. As a result, the child's enthusiasm for particular learning tasks will increase (Delta, 2009). Consequently, the child will be more interested in specific learning activities (Delta, 2009).

This quote reflects the uncertainty and insecurity regarding the assessment standards and equity, which result in unequal results and a potential student dissatisfaction.

4.1.2.6 Supportive Data from Quantitative Results

These qualitative observations are direct supplements to the data of the quantitative section of the research. Indicatively, in the survey, 60 percent of the people who were surveyed claimed that they had never been given any form of training on assessment. Also, the phrase I feel competent in the process of creating valid and fair assessments (Q10) had a mean score of 3.42, and only 20% of the participants strongly agreed. The combination of qualitative and quantitative data makes this theme more credible and shows a definite need to develop professionally.

4.1.2.7 Implications for Professional Development

The general tendency is indisputable: teachers have to carry out assessment tasks without being properly prepared to undertake them. Although the institutions might think that trial and error will make teachers work out, this model is not sufficient in the 21st century learning environment, which requires valid, ethical, and student-oriented assessment practices.

To deal with this, assessment training should be included in the teacher development programs of the institutions not only in the pre-service education but also in the continuous professional development (CPD). The future training module must cover such topics as rubric construction, formative feedback, diagnostic testing, and assessment ethics.

4.1.2.8 Conclusion of Theme 1

To conclude, Theme 1 laments the systemic absence of assessment literacy within university level English teachers in South Punjab irrespective of their type of institute and years of experience. This loophole not only impairs the equity and effectiveness of the language teaching, but also

constrains the development of students and the responsibility of the institution. The findings confirm the arguments behind the generation and delivery of a contextuality-dependent, talent-based instruction unit of the English language assessment guidelines an intervention that is ultimately presented.

5.1 Discussion and Conclusion

The research was a mixed method case study that investigated the practices of English language assessment in university teachers in South Punjab. Results showed that there was low levels of assessment literacy, institutional barriers, and emotional burdens and high motivation towards professional development. The quantitative and qualitative data have met, and professional development is structured and context-specific, and it should be informed to design specific assessment training module.

5.2 Discussion of Key Findings

The current paper provides a dimensional conceptualization of the practice of English language assessment by university teachers in South Punjab, and it shows that there are structural, professional, and emotional strains inherent in the assessment practice. The results of both quantitative and qualitative data show that there are enormous gaps in assessment literacy, institutional support, and policy coherence. Whereas, the survey findings reveal the moderate confidence in assessment design (Q10: Mean = 3.42), the interviews revealed the uncertainty at large scale with teachers basing their assessments on inherited or improvised formats. It means that there are weaknesses in the whole system of teacher education and not personal weaknesses. The absence of consistency in the usage of rubrics and feedback is another example of this gap. Though the quantitative data indicated partial usage of rubrics (Q8: Mean = 3.58), the qualitative data have indicated usage of those rubrics superficially or abandoned because of a lack of time and direction. Big class sizes and high-pass expectations were a strong source of institutional pressures which were reported (Q13: Mean = 4.10), and caused conflict between academic integrity and administrative requirements.

Although these are the restrictions, teacher motivation to grow professionally was high (Q17-Q18: Means > 4.30). Interview information showed the desire not only to receive technical training but also to develop identity as reflective assessors. Nonetheless, high-level assessment cultures and low team work restrained innovation and ownership.

5.3 Integration with Existing Literature

The results are consistent with both Pakistani and international studies, as they validate the existing gap in assessment literacy due to lack of training and institutional limitations (Shafqat and Mahmood, 2022; Saeed and Ahmad, 2020). This paper builds upon the previous research by presenting the emotional and moral costs of assessment such as fear of student retaliation and institutional pressure (Rehmani, 2003; Brookhart, 2011). Regardless of these issues, educators showed a great desire to have context-sensitive, collaborative, professional development to endorse situated learning views (Lave and Wenger, 1991; Shahid and Shabbir, 2022).

5.4 Implications for Policy and Practice

The research results of this study are significant to enhance the practice of English language assessment in both the public and the private universities of South Punjab. The findings underscore the fact that assessment literacy should be repositioned as a central professional skill but not peripheral skill. The design of assessment, educational practices with the use of rubrics, and provision of feedback should be offered through established and continuous training that is mandatory to policymakers and institutional leaders with a quality assurance system in place. The

key elements of sustainable reform are embedding assessment literacy into pre-service and in-service teacher education, encouraging teachers to work together, eliminating structural limitations, and encouraging inclusive and student-centred methods of assessment.

The study is limited even though it makes its contributions. The sample size was small and the data was limited to the universities in South Punjab so that generalisation of the findings to other areas was restricted. Self-selection bias might have been created by purposive and voluntary sampling. Also, all aspects of assessment practices may not be captured by the use of self-reported data and descriptive statistics. However, the research provides a solid empirical basis of the empirical research in the future as well as context-specific professional development programs in assessment of higher education.

5.6 Recommendations for Future Research

Future studies ought to be expanded geographically to cover the whole of Pakistan, to cover bigger and more varied samples, and to involve perspectives of the students. An additional study should be longitudinal to determine the effects of assessment training through time and qualitative study should also be conducted to examine the emotional, ethical and policy aspect that influence the English language assessment practice.

5.7 Conclusion of the Study

This paper has explored the process of English language assessment in South Punjab university instructors which has shown that there is a multifaceted interaction of professional, emotional and institutional forces. Results indicate that assessment is a negotiated activity and not a neutral one and is influenced by lack of training, institutional limitations and ethical pressures. As teachers were highly motivated to develop professionally, policy constraints, support, and time were obstacles to their ability to practice effective assessment. The paper demonstrates a distinct paradox between the desire and the possibility of change, that reform in assessment is possible only through systemic, as opposed to workshop-based, changes. Finally, equitable and learning-focused assessment practices can only be developed through empowerment of teachers with long-term training and collaboration with partners and institutions in the Pakistani higher education.

References

- Adie, L. (2012) 'The development of teacher assessment identity through participation in online moderation', *Assessment in Education: Principles, Policy & Practice*, 20(1), pp. 91–106. <https://doi.org/10.1080/0969594X.2011.650150>
- Akram, H. et al. (2020) 'Factors contributing low English language literacy in rural primary schools of Karachi, Pakistan', *International Journal of English Linguistics*, 10(6), p. 335. <https://doi.org/10.5539/ijel.v10n6p335>
- Akram, M. (2017) 'Learning and teaching English in Pakistan: Predicaments and solutions', *International Journal of Educational Sciences*, 19(1). <https://doi.org/10.31901/24566322.2017/19.01.03>
- Al-Mansoor, N.S. (2014) 'The effect of an extensive reading program on writing performance', *International Journal of Linguistics*, 6(2), pp. 247–255.
- Al-Tameemi, R.A.N. et al. (2023) 'Determinants of poor academic performance', *International Journal of Educational Research Open*, 4, 100232. <https://doi.org/10.1016/j.ijedro.2023.100232>
- Andersson, C. and Palm, T. (2018) 'Formative assessment through professional development', *Assessment in Education*, 25(6), pp. 576–597. <https://doi.org/10.1080/0969594X.2018.1430685>

- Awad, M.J. and Adwan, M.A.A. (2024) 'Alternative assessment methods', in *Advances in Educational Technologies*. <https://doi.org/10.4018/979-8-3693-0880-6.ch020>
- Bennett, R.E. (2011) 'Formative assessment: A critical review', *Assessment in Education*, 18(1), pp. 5–25. <https://doi.org/10.1080/0969594X.2010.513678>
- Brookhart, S.M. (2011) 'Educational assessment knowledge and skills for teachers', *Educational Measurement: Issues and Practice*, 30(1), pp. 3–12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3(2), pp. 77–101.
- Cheema, S., Siddiqui, M. and Masood, S. (2022) 'Teachers' perceptions of assessment practices', *Global Educational Studies Review*, VII(II), pp. 263–274.
- Cisterna, D. et al. (2025) 'Score reports for teacher educators', *Educational Assessment*, pp. 1–25.
- Clark, I. (2012) 'Formative assessment and self-regulated learning', *Educational Psychology Review*, 24(2), pp. 205–249.
- D'Elia, P. et al. (2023) 'Socio-economic status and English language learning', *OSF Preprints*. <https://doi.org/10.31234/osf.io/n3xms>
- Finn, J.D., Pannozzo, G.M. and Achilles, C.M. (2003) 'The why's of class size', *Review of Educational Research*, 73(3), pp. 321–368.
- Furtak, E.M. et al. (2016) 'Teachers' formative assessment abilities', *Instructional Science*, 44(3), pp. 267–291.
- Gallardo, K. (2020) 'The importance of assessment literacy', in *Intelligent Systems Reference Library*.
- Giraldo, F. and Quintero, D.M. (2019) 'Language assessment literacy and teacher development', *Colombian Applied Linguistics Journal*, 21(2), pp. 243–259.
- Guskey, T.R. (2002) 'Professional development and teacher change', *Teachers and Teaching*, 8(3), pp. 381–391.
- Habib, M. (2013) 'Education in Punjab: Outcomes and interventions', *Lahore Journal of Economics*, 18(S), pp. 21–48.
- Heikonen, L. et al. (2017) 'Classroom interaction strategies', *Journal of Education for Teaching*, 43(5), pp. 534–549.
- Herrera, L. and Macías, D.F. (2015) 'A call for language assessment literacy', *Colombian Applied Linguistics Journal*, 17(2), p. 302.
- Iqbal, Z., Saleem, K. and Arshad, H.M. (2020) 'Measuring teachers' knowledge of assessment', *Educational Studies*, 49(1), pp. 166–183.
- Khan, K. and Khan, W. (2019) 'Perceptions of English proficiency', *International Journal of English Linguistics*, 10(1), p. 15.
- Khan, T.J. and Khan, N. (2016) 'Obstacles in learning English as a second language', *Open Journal of Social Sciences*, 4(2), pp. 154–162.
- Likuru, L.C. and Mwila, P.M. (2022) 'Overcrowded classrooms', *Asian Journal of Education and Social Studies*, pp. 75–87.
- Looney, A. et al. (2017) 'Teacher assessment identity', *Assessment in Education*, 25(5), pp. 442–467.
- Mahara, K.K. (2024) 'In-service teacher training challenges', *KMC Journal*, 6(1), pp. 153–175.
- Makhambetova, A., Zhiyenbayeva, N. and Ergesheva, E. (2021) 'Personalized learning strategies', *International Journal of Web-Based Learning*, 16(6).

- Mansory, M. (2020) 'Non-traditional assessment in ELT', *International Journal of Linguistics*, 12(5), p. 210.
- Muhammad, I.M.S. et al. (2024) 'Assessment methods and learning outcomes', *LATAM Revista*, 5(5).
- Munir, J. et al. (2023) 'Impact of socio-economic status on achievement', *Journal of Social Sciences Review*, 3(2), pp. 695–705.
- Naz, S., Jabeen, S. and Rashid, A. (2020) 'Evaluation of PEELI project', *Global Regional Review*, V(I), pp. 332–339.
- Popham, W.J. (2011) 'Assessment literacy overlooked', *The Teacher Educator*, 46(4), pp. 265–273.
- Razzaq, N. (2023) 'Technology-based English instruction in Pakistan', *IARS International Research Journal*, 13(2), pp. 12–17.
- Rozilat, A.N. and Matore, M.E. (2023) 'Teacher assessment literacy framework', *IJARPED*, 12(2).
- Sainsbury, E.J. and Walker, R.A. (2008) 'Assessment as a vehicle for learning', *Assessment & Evaluation in Higher Education*, 33(2), pp. 103–117.
- Saleem, A. et al. (2022) 'Teachers' insight on continuous assessment', *IJSS Entrepreneurship*, 2(2), pp. 266–281.
- Saleem, A., Gul, R. and Dogar, A. (2021) 'Professional development effectiveness', *İlköğretim Online*, 20(3).
- Sarwat, S., Panhwar, A. H., Shahzad, W., & Shahzad, S. K. (2024). Critical Discourse Analysis of the Speech of Recep Tayyip Erdogan at the United Nations General Assembly. *Qlantic Journal of Social Sciences and Humanities*, 5(1), 95-106.
- Sarwat, S., Iftikhar, I., Sahito, J. K. M., & Shahzad, W. (2024). Impact of student engagement in language support classes through cooperative learning: A study of Pakistani educational institutions. *Research Journal for Societal Issues*, 6(1), 199-210.
- Segool, N.K. et al. (2013) 'Test anxiety among students', *Psychology in the Schools*, 50(5), pp. 489–499.
- Shafqat, A. and Mahmood, A. (2022) 'Assessment literacy in Pakistani universities', *International Journal of Assessment and Evaluation*, 12(2), pp. 40–53.
- Shahid, M. and Shabbir, A. (2022) 'Teacher agency and curriculum implementation', *South Asian Journal of Education*, 18(1), pp. 66–82.
- Shahzad, W., Shahzad, K., Farooq, A., & Shah, L. (2020). An Introduction to the Language Policies (Lp) of Pakistan as a Multilingual Country: An analysis of the use of different ideologies in the LP from 1947-2009. *Hamdard Islamicus*, 43(1&2), 216-227.
- Stoyhoff, S. (2008) 'Developments in language assessment', *Language Teaching*, 42(1), pp. 1–40.
- Tayyaba, S. (2012) 'Rural-urban gaps in achievement', *International Journal of Educational Management*, 26(1), pp. 6–26.
- Wiliam, D. and Thompson, M. (2017) 'Integrating assessment with learning', in Routledge eBooks.
- Younas, A., Shahzad, W., Naveed, A., Aziz, S., & Younas, A. (2024). An Analysis of Figurative Language in Taufiq Rafat's Poem "The Time to Love". *Qlantic Journal of Social Sciences and Humanities*, 5(1), 399-412.
- Zyngier, D. (2014) 'Class size and academic results', *Evidence Base*, 2014(1).

Appendices

Appendix A: Structured Questionnaire for Quantitative Data Collection

Title: *English Language Assessment Practices among University Teachers in South Punjab*

Instructions: Please indicate the extent to which you agree or disagree with the following statements by selecting the most appropriate option.

Scale:

- 1 – Strongly Disagree
- 2 – Disagree
- 3 – Neutral
- 4 – Agree
- 5 – Strongly Agree

Section 1: Demographic Information

(Please tick or fill where appropriate)

1. Gender: Male Female Other
2. Age: _____
3. Years of teaching experience: 1–5 6–10 11–15 16+
4. Type of institution: Public Private
5. Do you hold any formal training in language assessment? Yes No

Section 2: Current Assessment Practices

I regularly use written tests to assess my students' English proficiency.

I use oral assessments (presentations, discussions) in my courses.

I frequently use rubrics to evaluate student work.

My assessments are aligned with course learning outcomes.

I feel confident designing valid and fair assessments.

Section 3: Challenges in Assessment

I struggle to create assessments suitable for large classes.

I lack access to professional development opportunities in assessment.

Time constraints affect the quality of the assessments I design.

I find it difficult to give timely and constructive feedback to students.

I feel pressure to conform to traditional or outdated testing methods.

Section 4: Perceptions of Assessment Training

I believe that training in assessment practices is essential for teachers.

I am interested in participating in training programs to improve my assessment literacy.

I would benefit from learning how to design rubrics and formative assessments.

I believe professional training can improve fairness and consistency in student evaluation.

I feel more confident assessing students when I receive institutional support.

Appendix B: Semi-Structured Interview/Focus Group Guide (Qualitative Data Collection)

Title: *Exploring the Challenges and Needs of University Teachers in English Language Assessment*

Instructions: These open-ended questions are designed to guide the discussion. The interviewer may probe further for clarification or elaboration.

Section 1: Experience with Assessment

1. Can you describe your usual approach to assessing students' English language proficiency?
2. What types of assessment tools do you use most frequently, and why?

Section 2: Challenges in Assessment

3. What are the main challenges you face when assessing students in English courses?
4. How do classroom size, institutional policies, or time constraints affect your ability to assess fairly?

Section 3: Assessment Literacy and Training

5. Have you ever received any formal training in language assessment? If so, what was your experience?
6. In what areas do you feel you need the most support or training related to assessment?

Section 4: Suggestions and Improvement

7. If you were offered a training module on English language assessment, what topics should it include?
8. What changes would you suggest to improve assessment practices at your institution

Optional Reflection Question

9. How do you think improving assessment practices could affect student learning outcomes in your context?