



---

LETS

Journal of Linguistics and English Teaching Studies

ISSN (print) : 2715-4408

ISSN (online) : 2715-4416

Homepage : [stainmajene.id/index.php/lets](http://stainmajene.id/index.php/lets)

---

## Test Construction Practices: A Narrative Inquiry into English Teachers of Tertiary Education in Bangladesh

Ibrahim Hossain<sup>1</sup>, Farzana Yesmen Chowdhury<sup>2</sup>

Institute of Modern Languages, University of Chittagong<sup>1,2</sup>

[ibrahim@cu.ac.bd](mailto:ibrahim@cu.ac.bd), [farzana.chowdhury@cu.ac.bd](mailto:farzana.chowdhury@cu.ac.bd)

### Abstract

*Assessment is a crucial component of the teaching and learning process within any educational setting. In many educational contexts, teacher-constructed achievement tests serve as the primary means of assessment. Such tests can measure the attainment of educational objectives, facilitate effective instruction, and provide positive washback. This research employs a narrative inquiry approach to explore the test construction practices and to identify the influencing factors that shape these practices. It explores the lived experiences of a group of novice and experienced English language teachers at the tertiary level in Bangladesh. Individual stories are analysed using Brown and Abeywickrama's (2010) five language assessment principles. The findings of this study reveal a substantial gap between the principles of test construction and the actual practices and reveal that practices are predominantly shaped by institutional regulations rather than teachers' grasp of assessment theory and techniques. Findings of the study have implications for teachers in constructing practical tests, teacher trainers in designing training modules for novice teachers, and policymakers in developing appropriate assessment policies.*

**Keywords:** Test construction practice, Influencing factors in test construction, Assessment literacy, Narrative inquiry, Tertiary level English

### 1. Introduction

Assessment plays a critical role in any educational domain. It involves systematically collecting information on a learner's language ability or achievement (Bachman 2004). It is a cornerstone in the educational system that impacts various processes like teaching, learning, and decision-making (Black & Wiliam, 1998; Coombs et al., 2018). Test, a subset of assessment, is widely used in many educational contexts as the primary means of student assessment. Though high-stake standardised tests are usually constructed and administered by testing professionals, achievement tests in educational programs are mostly constructed by the course teachers. While a well-constructed test can provide a reliable measure of learners' ability, a poorly constructed test might negatively affect learning and teaching. Brown and Abeywickrama (2010) highlighted the complexity of good test construction, for it requires the use of both science and art.

Assessment literacy is necessary for all teachers as they might not have the same educational background and experience to navigate the complexities related to test construction. It has primarily been defined as

“teachers’ understanding of assessment processes as well as their capacities to design assessment tasks, develop adequate criteria for making valid judgments on

the quality of students' performances, and understand and act upon the information that is collected through assessment" (Hay & Penney, 2013, pp. 69–70).

Hence, it is essential to know how teachers in different contexts and levels of their teaching service navigate the complexities of test construction. Identifying the dominant factors that shape and regulate their test construction practices is also important.

The education system in Bangladesh is distinctly characterised by its heavy reliance on achievement tests. At the secondary level, teaching and learning mostly revolve around achieving good grades in the exams. (Amin & Greenwood, 2018). The assessment procedures within secondary and higher secondary education are notably shaped by the significant impact of high-stakes examinations administered by various education boards. (Rahman & Khan 2021; Sultana, 2018). Conversely, English language assessment practice at the tertiary level holds distinct characteristics as English teachers adopt a more engaged role in assessment-related processes. They assume wide responsibilities that span the creation of tests, moderation of questions, the administration of tests, and the compilation and dissemination of results. Despite English teachers' heavy involvement in assessment-related activities at the tertiary level, almost no study has explored the assessment practice of English teachers at the tertiary level in Bangladesh to date.

This study explores the lived experiences of English teachers at the tertiary level in Bangladesh with the aim of investigating their regular assessment practices. It mainly focuses on their question-setting strategies for the achievement tests. Individual perceptions of the participants towards such practices are also highlighted to identify the factors that influence or shape their question-setting strategies.

## 2. Literature Review

Among the different assessment processes, question setting plays a significant role in measuring students' learning outcomes, problem-solving activities, and critical thinking. The use of questions in the assessment practice is multifaceted: eliciting existing knowledge, fostering the development of understanding, and cultivating the capacities for critical thinking of the students. While properly constructed questions promote learning, poorly framed questions can impede learning (Tofade et al., 2013). Formatting questions are mainly of two types: convergent and divergent questions. Convergent questions are intended to elicit a single best response from the learner while divergent questions are more open-ended and require critical thinking to answer (McComas & Abraham, 2004). Questions can also be classified according to their ability to measure learners' cognitive level. A hierarchical framework for measuring cognitive ability was initially proposed by Bloom (1956) and later modified by Anderson and Krathwohl (2001). According to this framework, questions should be designed to measure different levels of cognition, from mere memory recall to a higher level of critical thinking and creation.

Teachers' assessment strategies and preferred assessment methods are subject to multiple factors, encompassing assessment education, individual needs, and personal preferences (Coombs et al., 2018). The approaches teachers adopt in their assessment practices are intrinsically tied to their conceptual understanding and practical knowledge, which they have developed within their unique educational environments (DeLuca et al., 2018). Recognising the pivotal role of teachers' assessment practices in shaping students' learning outcomes and achievements (Black & Wiliam, 1998; DeLuca et al., 2018; Hattie, 2008), there exists a pertinent need to understand the disparities and commonalities in teachers' assessment approaches across diverse learning and instructional settings.

Brown and Abeywickrama (2010) proposed five principles of language assessment: practicality, reliability, validity, authenticity, and washback. According to them, such characteristics should be applied to any form of assessment either in question setting, script marking, or test administration for the effectiveness of any formal test. In setting a question paper for any formative or summative test, Brown and Abeywickrama's (2010) five principles of language assessment should be addressed in the following way-

**Practicality:** Instructions in the questions should be clear and manageable for both test takers and administrators. It refers to formulating test questions so that test takers can complete the test within the stipulated time.

**Reliability:** Reliability in question setting refers to formulating questions for a test with clear instructions for both test takers and the scorer. If the test contains multiple-choice items, the distractors must be evenly tricky, and the answers should not be easily singled out. A reliable test item should not have more than one correct answer. Questions should be designed to yield consistent results, ensuring that different test versions produce similar outcomes.

**Validity:** A valid test measures precisely what it is supposed to measure and nothing else. The items of question in the test must be familiar to the students and their learning objectives. As a result, a valid test should be aligned with the course objectives.

**Authenticity:** An authentic test contains test items that replicate real-world tasks. The items in the test should be contextualised, not isolated. The language used in the question paper of tests has to be as natural as possible. When a test is designed in such a way that influences learners to be involved in learning, it achieves positive washback.

**Washback:** It considers the impact of the test on teaching and learning. Questions should be aligned with the curriculum and encourage effective teaching practices. Additionally, it should provide clear and constructive feedback to test takers to help them improve. Positive washback occurs when test questions align with course objectives and help students achieve those objectives through tests. Conversely, negative washback occurs when students merely memorise test items and the test does not effectively measure the course's objectives.

The above five principles can provide valuable guidelines for evaluating the practice of test constructions by English teachers at the tertiary level in Bangladesh.

Assessment practice has been investigated widely in different EFL contexts across the world, namely China, Indonesia, New Zealand, Vietnam, UK (Alderson & Buck, 1993; Alderson, 2010; Fan & Jin, 2013; Fjørtoft, et al. 2024; Phuong et al. 2023; Xu & Liu, 2009; Zulaiha et al., 2020;). Coniam (2009) found that English tests produced by EFL teachers in Hong Kong lack reliability, while Glusac & Milic (2021) found that instructions in teacher-made written tests lack clarity and recommended training teachers in writing instructions. Weng (2023) found that writing assessment literacy (WAL) among Chinese tertiary level EFL teachers was positively influenced by their academic background and training. Islam, et al. (2021) raised concerns about the reliability and validity of English tests in Bangladesh, identifying ineffective assessment as a root cause of English language learning failure and recommending teacher training to improve knowledge, skills, and professionalism. There are some studies that have focused on the effect of testing on the motivation, self-resilience and foreign language anxiety of EFL students (Ahmed et al. et al., 2023; Chakmak et al., 2023; Ismail et al., 2023; Wicaksono, 2023). In addition, a few researchers have investigated the perception of teachers towards English language assessment (Isik, 2021; Sohrabi et al., 2022). Findings of their research emphasised the need for pre-service and in-service

training in integrating teaching and assessment, considering assessment as a complementary component of the teaching process. However, research on the perceptions of teachers towards question-setting strategies in English language education at the tertiary level in Bangladesh is scarce to date.

The current study, therefore, aims to explore the assessment practices, especially the methods of constructing achievement tests in various public and private universities in Bangladesh. It also aims to identify the dominant factors that shape and regulate the testing concepts of the participant teachers. Given the impact of teachers' assessment practices on student learning, this study aimed to investigate the following research questions:

- i. What strategies are employed for question setting by the English teachers at the tertiary level in Bangladesh?
- ii. What influential factors do shape and regulate their question-setting practices?

### **3. Method**

This study adopted a narrative inquiry into the lived experiences of a group of novice and expert English teachers at different public and private universities in Bangladesh to explore their perceptions towards their assessment practice, particularly question-setting practices. This article is part of a larger research on the overall assessment practice of English language teachers at the tertiary level in Bangladesh. Though the original study conceptualised assessment practices into three main areas: question setting, script evaluation and test administration, this article reports on question setting strategies employed by tertiary level English teachers in Bangladesh.

This study collected narratives from eight English teachers using criterion sampling (Patton, 2002). Participants were selected based on three main criteria: teaching experience, education or training in language testing and assessment, their teaching institution, and the type of institute they teach. Data was collected from two main cities: Dhaka, the capital city and Chittagong, the second-largest city in Bangladesh. Participants were selected in terms of their length of teaching experience from both public and private universities in Bangladesh. Participants with more than ten years of teaching were regarded as experienced, while participants with less than three years of teaching experience were regarded as novices. Two experienced and two novice teachers were selected from two public and two private universities for this study. Four of the informants have previous educational backgrounds in testing and assessment. At the initial part of each interview session, the demographic information of the participants, namely age, gender, and educational background, was identified, though the age and gender of informants were not considered essential variables for the study.

Table 1: Participant details

Teacher	Gender	Academic qualification	Length of teaching	University type	Education on testing and assessment
T1	Male	B.A (Hons) M.A in English	21years	Public	Studied a course on language testing and assessment in MA
T2	Female	PhD, B.A (Hons) M.A in English	22years	Public	Studied a course on language testing and assessment in MA
T3	Male	B.A (Hons) M.A in English	15 years	Private	Studied a course on language testing and assessment in MA
T4	Female	B.A (Hons) M.A in English	2 years	Public	Didn't study any course
T5	Female	B.A (Hons) M.A in TESOL	10 years	Private	Studied a course on language testing and assessment in MA
T6	Male	B.A (Hons) M.A in English	22 years	Public	Didn't study any course
T7	Female	B.A (Honos) M.A in Linguistics	7 months	Private	Didn't study any course
T8	Male	B.A (Honos) M.A in English	1 year	Private	Didn't study any course

This method primarily involves individual semi-structured interviews as a research tool to elicit the lived experiences of the participants of this study (Chowdhury & Hamid, 2016). The interview was conducted in two phases. All interviews were audio recorded and transcribed. Based on the transcription, the individual story is written separately for each participant by the researcher. This was followed by a second round of interviews designed to corroborate and further illuminate the information collected during the initial interviews. Later, the story was sent to the participants by email for member checking and data validation. Following Murray (2003), this study adopted a two-phase method for analysing narrative data. The first stage involves describing the narrative content and structure, while the second stage involves interpreting the narrative using theoretical literature.

## 4. Findings

This section analyses the interview data on the participants' experiences of Question-setting strategies and their perceptions towards such practices. The analysis will determine the factors that shape and regulate their question-setting practices.

### 4.1 Question-Setting Strategies

All the participants in the study are involved in teaching Foundation English courses to undergraduate students and preparing achievement tests. The teachers in the public university mostly follow a summative assessment through a four-hour pen-and-paper achievement test. In contrast, the participants in the private universities follow continuous assessment and a summative pen and paper written test at the end of the semester. The very nature of the course and the academic context largely influence teachers' assessment practice.

The English course in public universities focuses mainly on teaching reading, grammar and writing. There is no test of listening and speaking. The year-end summative test is a 100-mark test designed by the course teacher and by a second examiner. The two sets of questions are then moderated by a three-member moderation committee. The items in the test include unseen reading comprehension, grammar and writing. Grammar is tested through discrete point items like sentence corrections, fill-in-the-blanks, changing sentences and so on. The writing section consists of writing letters, applications, paragraphs and essays.

#### 4.1.1 Reliance on Ready-Made Sources

Findings of the interview show that all participants show a similar trend in question-setting strategies. Both experienced and novice English teachers from both public and private universities depend on ready-made sources like English grammar books, English Language Graded books, Reference books, and online resource materials for setting questions. Participants of the novice category mostly copy ready-made questions from English language books published by different publishers.

They reported,

*<sup>1</sup>It is easier to collect reading comprehension questions from guidebooks. Most of the English books contain reading comprehension. If you want to set a reading comprehension question, it will take a lot of time. [T4]*

*I do not test grammar through isolated sentences. I test grammar in context. However, my other colleagues test grammar in isolation because they believe it is easy to set questions and mark. I use the internet for resources to set questions. I take reading articles from the internet. I believe I mainly focus on the time I spend answering the questions. Sometimes, I receive requests to clarify the instructions while I am on invigilation duty in the exam hall. I believe I understand the technicalities related to the importance of writing clear instructions while setting questions. I always teach students how to follow instructions while answering questions. [T8]*

However, experienced teachers claim to change the contents they take from ready-made sources. T1 reported,

*I use many internet resources for setting questions. I adapt the ready-made questions from different websites according to my context. In addition to internet resources, I also use course books for setting questions.*

The above examples show the reliance of the participants on ready-made teaching materials instead of preparing authentic materials based on the level and context of learners, although the experienced teachers make some changes in adapting the commercial materials. Regarding the use of textbooks in question setting, Taylor (2009) states that most accessible textbooks are “highly technical or too specialised for language educators seeking to understand basic principles and practice in assessment” (p. 23). In the same way, Bailey and Brown (1996) and Brown and Bailey (2008) found that teachers usually develop their familiarity with particular textbooks by using them for a longer period, and textbooks are hardly changed. In the same way, the informants of this study are dependent on using textbooks in preparing their question papers.

Reliance on ready-made sources for setting questions indicates that they focus more on personal convenience than the validity and authenticity of their questions. Brown and Abeywickrama (2010) suggest that validity and authenticity should be reflected in the question-setting practices of English language teachers. According to the principle of validity, the question items in the test must be familiar to the students and should be aligned with the course objectives. He further argues that question items should be authentic based on the context of the learners. However, ready-made commercial materials, i.e., textbooks, graded English language books, or online-based reading materials, are published in different contexts. They might not be associated with the needs, context or objectives of

---

<sup>1</sup> The excerpt collected from participants' interviews is presented in italics, and no grammatical changes were applied here to preserve data validity.

English language programs in Bangladesh. Hence, the validity and authenticity of the assessment are not appropriately maintained.

#### 4.1.2 Following Previous Years Test Format

Individual narratives of the informants of this study demonstrate that they follow question patterns of the previously held examination of English language courses at their respective universities. When they were asked to describe their first experiences of question setting for the course they first taught, they all shared a similar experience of following the previous year's question paper. They reported,

*At the beginning of my teaching, when I was asked to set a question, I felt nervous..... when I sought help from one of the senior teachers, she gave me a question from previous years. [T7]*

*I found question setting very difficult and time-consuming in the beginning. I did not receive any instruction for setting questions from my department. I did not receive any support from other colleagues either. I did not ask for any support. In my early teaching days, I simply followed the patterns of previous years' questions. [T1]*

The above practice suggests that teachers are not fully aware of the purpose of assessment. Their classroom teaching is also primarily influenced by their assessment practice as they mainly teach the items that students must answer in their final exam. Research shows that ESL teachers prefer the traditional examination system, a practice of traditional teaching and assessment systems in which they had schooling experiences (Singh & Arshad, 2013). It is also evident that ESL teachers encourage students to develop a memorisation culture by engaging them in final exam-specific activities throughout the term or semesters (Chan & Sidhu, 2011). Individual narratives of this study show that traditional test formats are convenient for teachers when they are setting questions. They often prepare templates and resources for creating such tests, which can save time and effort compared to designing more innovative assessment tasks. It usually happens due to several reasons. Firstly, there is a huge workload of teachers in both private and public universities in Bangladesh. Teachers at private universities are engaged with teaching, continuous assessment and results publication throughout the year. This is because they have to teach many courses at a time for each semester. In the same way, English teachers at public universities are highly occupied with question setting and script evaluation as English is a compulsory course for all disciplines, and the exam is usually held in the same periods. In addition, each discipline requires two examiners for the English test due to the double examination system. However, the number of English teachers is minimal compared to the number of departments, institutes, and students. For this reason, teachers have to prepare question papers for many English courses at a time, and they depend on following the previous year's question patterns at their earliest convenience. However, it is essential to balance convenience with the effectiveness of assessment in promoting meaningful learning outcomes for students.

Such practice of 'teaching to the test' raises concern about the washback effect of the test. Once the teachers start following the existing question format, they can hardly come out of it. As a result, the same question pattern keeps repeating year after year. This causes a negative washback effect. Washback is known as the influence of testing on teaching and learning (Alderson & Wall, 1993). According to Brown and Hudson (2002), test items should be directly related to the language teaching/learning process as they represent the

components of the language curriculum at the pre-defined level. As long as the test items are parallel with the objectives of the syllabus/curriculum, they will have potential positive backwash effects on the learners; otherwise, they will negatively influence their learning. Research conducted in the education system of Bangladesh (Sultana, 2019; Hamid, 2011; Islam, 2015; Karim, 2004; Rahman, 2015; Selim & Mahboob, 2001) found similar patterns of teaching to the tests in the secondary and higher secondary levels.

#### 4.1.3 Focus on Time than the Test Constructs

The purpose of the assessment is to assist test takers in knowing what they can perform well and what they should improve. For this reason, the teachers should focus on the assessment task based on the course objectives. However, two novice teachers in this study have reported focusing more on the test duration than the test constructs while setting questions. They ensure they put enough items for the students to write for four hours.

*While conducting class tests I allot time based on marks for the class test. I think test time is based on the total marks. For example, 2 hours to answer a written test of 40 marks, for 30 marks 1 hour and a half, for 15 marks 45 minutes. I must conduct the class tests within the class hour. [T8]*

The above excerpt shows that the participant is unaware of the course objectives. She is not quite sure what she will measure through the test. It represents novice English teachers' current trends and challenges in their regular assessment practices. It has happened to them as they were not provided with any training or instruction on question setting since they were first involved in such assessment activities. It raises questions about the validity of the questions they set and the lack of their assessment literacy.

#### 4.2 Concerns about Assessment Literacy

Language assessment literacy means the knowledge, skills and principles for language testing (Davies, 2008; Fulcher, 2012). In several studies (Giraldo, 2018; Inbar-Lourie, 2008, 2013; Malone, 2008), teachers' assessment literacy is directly connected with good assessment practices. The findings of this study contradict the common belief that teachers' assessment literacy is directly reflected in their assessment practices.

Despite their qualifications, education, and involvement in assessment-related activities, two informants of the experienced category expressed reservations about their assessment literacy and practical implementation. They commented on their assessment practice,

*Though I studied a language testing and assessment course, and I am familiar with the concepts of validity, reliability and practicality, I cannot transform my knowledge into practice. [T2]*

*I don't consider myself an assessment literate though I studied, conducted workshops with teachers and published articles, I cannot practice what I know. [T1]*

This gap between theoretical knowledge and actual practice is also acknowledged by O'Loughlin (2006). He maintains that Language testing is a "notoriously difficult domain of knowledge for students in second language teacher education programs" due to the intricate balance between its highly abstract theoretical concepts and their practical applications (O'Loughlin, 2006, p. 71). Although assessment knowledge is indeed essential for teachers' successful assessment practices, some research has found that professional training



programs often fail to train teachers to develop assessment tasks aligned with course objectives or evaluate the exam script perfectly (Kao, 2023). This implies that the relationship between assessment knowledge and effective assessment practices might be more complex and influenced by various factors, such as teaching context, institutional guidelines, and personal challenges. The study highlights the need to consider these nuances when assessing teachers' assessment literacy and its translation into practice.

### 4.3 Institutional Influence on Assessment Practice

Brindley (2001) states that the social context of testing should be considered for the discussion of effective assessment practices. Individual stories of the novice English teachers of this study show the institutional influence on their assessment practices. T8 reported,

*My senior colleague told me “make easy questions for the students so that they can pass. Don’t make complicated questions”. I often receive unfair requests from the chairman of other departments where I teach the English course. They direct me to set easy questions for the students of their departments. I think a private university is a kind of business. As they have a tendency to attract students by giving more marks.*

The experience of T8's receiving unfair requests from other departments' chairmen to set easier questions raises concerns about external interference in assessment practices in the private universities in Bangladesh. The impact of institutional regulations and culture on T8's assessment practices indicates the influence of external factors on teachers' approaches to assessment.

On the other hand, T2 and T1 from a public university background tell their experiences-

*As we maintain a double examiner system in our university, when I set questions as a second examiner, I usually don't pay much attention to the details of the questions. In my university, the moderation board usually takes 80% of the questions from the first examiner and 20% from the second examiner. I think the double examiner system in public universities largely influences our question-setting practice. We act differently in terms of our roles as first examiners, second examiners, and members of the moderation committee. When I set questions as a first examiner, my question setting is mostly regulated by the contents I taught during the course. In contrast, when we set questions as a second examiner we hardly bother about the contents taught during the course. Again, when I work as a member of the moderation committee I focus on the standards of the questions.*

*The situation is totally different in the private university as they have a single examiner system in the private universities. The course teachers are empowered to set questions as they want. Though some universities have moderation systems, the moderation committee does not usually make any major changes. They just check if the calculation of the total marks is okay or if there are any grammatical mistakes in the instructions of the questions. [T2]*

*I perform a range of assessment activities like setting questions, question moderation, and script evaluation. I also conduct quiz tests, and class tests as part*

*of continuous assessment in my classes. My assessment practice is largely guided by the university system. My university rules force me to follow a summative assessment system, but I am aware of the significance of the formative assessment that I want to employ in my classroom teaching such as quizzes, class tests, assignments, and presentations. [T1]*

Understanding how institutional norms shape assessment practices is essential for creating effective assessment policies. This finding highlights the need to maintain the integrity and validity of assessments and the importance of fostering an environment that values fair assessment practices.

#### **4.4 Novice Teachers' Assessment Challenges**

When novice teachers were asked about their experiences with assessment practices, they all expressed a wide range of challenges that they face regularly. They reported,

*After joining the university as a teacher, I attended a conference where they gave a speech on language assessment and testing, but to be very honest, I didn't have any idea. All of them said our assessment is not valid but what type of assessment is valid? I have not got the actual idea. They were saying IELTS assessment is standard but though they were university teachers they were not saying anything about university assessment. [T7]*

This remark explains how novice teachers feel attending talks on language testing. It is even difficult to follow technical terms like validity. Despite completing degrees in Language and Linguistics, T7 lacked formal language testing and assessment training. Her limited training might have contributed to her uncertainty in assessment practices. T8 explains,

*I don't consider the authenticity of the tasks while setting questions. I go with the traditional items, and I don't dare to break the tradition.... but I believe I will be able to add real-world tasks when I will be a bit experienced.*

The above excerpt shows the fear of novice teachers not to do any innovative practices in question setting at this initial stage of their career. Like him, other novice teachers T4, T5 and T7 believe that the content of the questions should be familiar to the students and reflect real-world language use so that test takers can relate the task to their needs and context. Nevertheless, it is an interesting finding of this study that they all love to stay in the traditional bubbles of assessment practices.

Such perceptions of novices are also reflected in the stories of experienced teachers (T1 and T3) when they reported their experiences as a novice teacher-

*I believe, my assessment practice evolved with the passage of time. With the passage of time, my assessment practice changed. I have moved from the traditional approach to the modern one. I further believe that my training and experience have helped me to prepare more standard questions than before. I currently use a lot of internet resources for setting questions whereas I previously just followed previous years' questions. At present, I adapt the ready-made questions from different websites according to my context. In addition to internet resources, I also use course books for setting questions.....[T1]*

*Novice teachers in my university are not given any training in question setting. There are no written instructions for either question setting or script grading. Teachers are simply given the syllabus and the previous year's question paper to set questions. At that time teachers considered previous years' questions as instructions. [T3]*

The narratives of the experienced teachers of this study shows that they initially followed the traditional question patterns of previous years due to the lack of assessment literacy, instructional support and training. However, years of teaching experiences have provided them with adequate knowledge, skills and confidence to develop their assessment tasks incorporating the principles of assessment: authenticity and validity. The importance of adequate training and support for novice teachers in assessment methods emerges as a significant finding for this group of novice English teachers.

## **5. Discussion**

This study focuses on the assessment practices of tertiary-level English as a Foreign Language (EFL) teachers in Bangladesh. Adopting narrative inquiry as its research approach, the research findings have been presented based on the lived experiences of current English teachers at different public and private universities in Chittagong region. The narrative of each participant is analysed in terms of Brown and Abeywickrama's (2010) five principles of language testing: practicality, reliability, validity, authenticity, and washback. It has also identified their in-depth perception and attitudes towards their assessment knowledge, skills, and experiences as English teachers.

The first research question focuses on the lived experiences of the informants to identify the strategies they employ in their question-setting practices. This mainly found reliance of English teachers on ready-made resources, i.e., textbooks, internet resources and previous years' question papers as part of question-setting strategies. In narrating their experiences, they highlighted the challenge they have to undergo as novice teachers due to their lack of experience and adequate assessment knowledge in formulating effective questions following the principles of assessment practices. Existing assessment theory proposes that the primary goal of the achievement tests is to measure what extent learners have achieved the goals set in the curriculum. As the English courses taught at tertiary level education in Bangladesh follow achievement tests, English teachers need to be well conversant with the objectives of the course.

The test items must be designed to appropriately align with the course objectives (Brown, 2004). While writing the test items, teachers should consider the fundamental principles of assessment like validity, reliability, authenticity, practicality and washback (Brown & Abeywickrama, 2010). Both teachers and students should have a clear understanding of what is going to be taught and how it will be assessed. An analysis of the data shows that all the participants are either uninformed or have a vague idea about the course objectives. In their classroom teaching, they try to cover the items in the syllabus and final tests and set questions on the topics covered during the course. They are hardly concerned about the course objectives. As a result, there is a gap between the course objectives and the measurement of the objectives. Such practice raises questions about the validity of the test.

The interviews with the participants sought to know whether the teachers consider the issue of practicality and reliability in their assessment practice; especially in designing the achievement test. While setting the questions, participants were unaware of the scoring criterion: concept, skill, and knowledge, and the reliability factors: the length (or total number

of questions), the quality of the questions, and the fit to the group being measured. Furthermore, they hardly focus on using authentic materials in developing tasks for the question paper. It is also identified that the assessment practices of such educators are predominantly shaped by institutional regulations rather than their individual understanding of assessment theory and techniques.

The second research question focuses on factors influencing question-setting practices of English teachers at tertiary level education in Bangladesh. Individual narratives of informants offered a unique perspective on the influential factors shaping assessment practices when their interview asked them to express their perception towards their experiences of question-setting practices. Factors identified here are institutional conventions and regulations, length of experience of question setting, and perceived assessment literacy.

Institutional context is the most critical factor which influences teachers' assessment practice. Both for teaching and testing English teachers have to comply with the syllabus and the assessment system set by the respective department of the university. This is because English is a compulsory course for students of all disciplines. In the context of the public university, it is evident that the only assessment is a pen-and-paper summative test at the end of a year based on the syllabus and assessment practices set by the particular department. English teachers, therefore, hardly feel to bring any changes in their assessment system. Even if they feel like bringing any changes, they do not have the authority due to the cumbersome process of changing the ordinances.

The experienced and novice category teachers, regardless of their previous language testing and assessment education, perform similar patterns in question setting strategies in public universities. In contrast, English teachers in private universities invariably follow the semester system. As part of their institutional requirements, they follow a continuous assessment system and a summative written test at the end of the semester. Consequently, they must set questions for class tests, quizzes, and assignments. However, due to the double examiner system in public universities, teachers assume different roles while setting questions as first and second examiners. The first examiner focuses on the contents taught during the course, while the second examiner usually is not equally concerned about this. Such roles have a direct influence on teachers' question-setting practice.

The study aimed to develop a better understanding of the complex interplay between personal, institutional, and contextual factors that inform English language teachers' assessment methodologies by identifying and categorising the emerging themes. The findings of this research demonstrate a significant gap between the assessment principles and the actual practices employed by English teachers at the tertiary level in Bangladesh. This discrepancy is highlighted by the discriminatory practices observed between teachers at public and private universities. The assessment practices of these teachers are predominantly influenced by institutional regulations rather than the teachers' individual assessment knowledge and expertise. In contrast to the public universities, the single examiner system in private universities provides teachers more flexibility to remain focused on course objectives.

The second most significant factor is teachers' experience in question setting. Several experienced participants believe they have developed adequate proficiency in question

settings due to their long teaching experience despite the challenges they faced at the initial stage of their profession. Some experienced teachers in both public and private universities created their personal collections of question banks, which they use to set new questions. The novice category in both public and private universities believes that they would be able to gain more proficiency and confidence if they progress in their teaching career.

Narratives of the participants of this study reveal that the assessment practices of their schooling are one of the factors. Most novice teachers, with or without language testing and assessment education, follow a common strategy for setting questions for the first time. As students, they developed a familiarity with question patterns by appearing in the exams for an extended period in their undergraduate and postgraduate programs. Participants graduating from private universities hold different beliefs and concepts about assessment than those from public universities. However, both groups apply their previous understanding of test items in their question-setting practices when they become teachers.

Finally, inadequate assessment literacy and associated training programs impact their question-setting experiences. Novice teachers struggle to find a way to do their best in their question setting as they were not provided with any formal training. However, some of the participants reported that their previous education in testing and assessment helped them cope with the challenges of question setting, though they cannot apply their knowledge fully in practice due to excessive workload and institutional norms. Nevertheless, they feel their education in assessment helped them consider issues like practicality and authenticity in question setting. In contrast, the experienced teachers have developed their skills through attending various professional training programs and higher studies from abroad, in addition to their wide range of experience in question settings each year, but they are not also able to use their knowledge of testing and assessment in practices due to the prevailing traditional assessment system and curricular requirements set by the respective university authority. However, there is hope among them as the government of Bangladesh has taken the initiative to implement an outcome-based education policy. According to the new curriculum, teachers of both public and private universities have to set questions following Bloom's taxonomy. which has to be implemented in all tertiary-level education very shortly.

## 6. Conclusion

Although this study has a small sample size, it provides insight into question-setting strategies by exploring the lived experience of a group of novice and experienced English teachers at the tertiary level of education in Bangladesh. It is the first in-depth qualitative study in the field of testing and assessment practices presenting the attitudes and individual perceptions of this group of teachers towards their experiences of question-setting practices. Although the existing research on assessment practices focuses on the effectiveness of question-making in the regular classroom setting, this study is the only one that has investigated the question-setting practices for achievement tests in English language education at the tertiary level in the context of Bangladesh. It has presented the personal stories of English teachers, including both novices and experienced teachers, to visualise their daily challenges and coping strategies with institutional norms and regulations. The individual lived experiences highlighted the urgent need for assessment literacy for teachers, educators and policymakers. This is because an individual teacher cannot make any changes in the question-setting patterns until the institution reforms its traditional approach to follow the previous year's question. Stiggins (1997) states that

assessment literacy refers to a wide range of skills and knowledge that stakeholders, namely test writers, classroom teachers, university administrators and professional language testers, need to deal with the new world of assessment in which we have been thrust.

The findings of this study have raised the issue of the quality and effectiveness of teacher education and the process that should develop the assessment literacies of all stakeholders. It has implications for various stakeholders within the educational domain as follows:

Analysis of the data suggests that there is room for improvement in aligning assessment practices with established principles. Regular professional development programs on assessment knowledge and skills will enable them to create more effective and fair assessment methods that accurately measure student learning outcomes. By acknowledging the gap between principles and practices, they can contribute to developing more valid and reliable language tests to measure students' language proficiency levels more accurately. Furthermore, the findings of this study inform the educators of each discipline and policymakers in tertiary-level education to create changes in their decisions about existing assessment regulations and guidelines. By addressing the identified gap, policymakers can work towards creating a more equitable and standardised assessment framework that ensures consistency and quality in tertiary-level English education. To conclude, it can be said that English teachers, policymakers, university authorities, and higher education ministries can collectively contribute to improving the overall quality of assessment and, subsequently, the teaching and learning processes within the educational system.

In summary, the study highlights the need for closer alignment between assessment principles and practices in English language education at the tertiary level in Bangladesh. Future research is required to explore the overall assessment practices in English language education at the university level for different contexts worldwide.

## References

- Ahmed Abdel-Al Ibrahim, K., Cuba Carbajal, N., Zuta, M. E. C., & Bayat, S. (2023). Collaborative learning, scaffolding-based instruction, and self-assessment: impacts on intermediate EFL learners' reading comprehension, motivation, and anxiety. *Language Testing in Asia*, 13(1), 16.
- Amin, A., & Greenwood, J. (2018). The examination system in Bangladesh and its impact: on curriculum, students, teachers and society. *Language Testing in Asia*, 8(1), 1-18 <https://doi.org/10.1186/s40468-018-0060-9>
- Alderson, J. C. (2010). A survey of aviation English tests. *Language Testing*, 27(1), 51-72.
- Alderson, J. C., & Buck, G. (1993). Standards in testing: a study of the practice of UK examination boards in EFL/ESL testing. *Language Testing*, 10(1), 1-26.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., & Pintrich, P. R. (2001). A revision of Bloom's taxonomy of educational objectives. *A Taxonomy for Learning, Teaching and Assessing*. Longman, New York.
- Bachman, L. F. (2004). *Statistical Analyses for Language Assessment Book*. Cambridge University Press.
- Bachman, L.F. & Palmer, A.S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.

- Bailey, K.M. & Brown, J.D. (1996). Language testing courses: What are they? In A. Cumming & R. Berwick (Eds.), *Validation in language testing* (pp. 236-256). Clevedon, UK: Multilingual Matters.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5 (1), 7–74.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain* (pp. 201-207). New York: McKay.
- Brindley, G. (2001). Language assessment and professional development. *Experimenting with Uncertainty: Essays in Honour of Alan Davies*, 11, 137–143.
- Brown, H.D. (2004). *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education
- Brown, H. D., & Abeywickrama, P. (2010). *Language Assessment Principles and Classroom Practices* (2nd ed.). White Plains, NY: Pearson Education.
- Brown, J. D., & Bailey, K. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25 (3), 349-383. <https://doi.org/10.1177/0265532208090157>.
- Brown, J. D. & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge: Cambridge University Press.
- Çakmak, F., Ismail, S. M., & Karami, S. (2023). Advancing learning-oriented assessment (LOA): mapping the role of self-assessment, academic resilience, academic motivation in students' test-taking skills, and test anxiety management in Telegram-assisted-language learning. *Language Testing in Asia*, 13(1), 20.
- Chan., F. Y., & Sidhu, G. K. (2011). School-based assessment among ESL teachers in Malaysian secondary schools. *Malaysian Education Deans' Council Journal*, 9, 75-87.
- Chowdhury, F. Y., & Hamid, M. O. (2016). Language, migration and social wellbeing: A narrative inquiry into the lives of low English proficiency Bangladeshi migrants in Australia. *Australian Review of Applied Linguistics*, 39(1), 8-30.
- Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education*, 71, 134–144.
- Coniam, D. (2009). Investigating the quality of teacher-produced tests for EFL students and the effects of training in test development principles and practices on improving test quality. *System*, 37(2), 226-242.
- Das, S., Shaheen, R., Shrestha, P., Rahman, A., & Khan, R. (2014). Policy versus ground reality: Secondary English language assessment system in Bangladesh. *Curriculum Journal*, 25(3), 326-343.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327– 347 <https://doi.org/10.1177/0265532208090156>.
- DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). Teachers' approaches to classroom assessment: A large-scale survey. *Assessment in Education: Principles, Policy & Practice*, 25(4), 355–375.
- Fan, J., & Jin, Y. (2013). A survey of English language testing practice in China: The case of six examination boards. *Language testing in Asia*, 3(1), 1-16.

- Fjørtoft, H., Lai, M. K., & Li, M. (2024). How curriculum and assessment policies affect the role of reading in an assessment culture: a New Zealand case study. *Assessment in Education: Principles, Policy & Practice*, 1–31. <https://doi.org/10.1080/0969594X.2024.2365645>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132 <https://doi.org/10.1080/15434303.2011.642041>.
- Giraldo, F. (2018). Language assessment literacy: implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179-195 <https://doi.org/10.15446/profile.v20n1.62089>
- Glušac, T., & Milić, M. (2022). Quality of written instructions in teacher-made tests of English as a foreign language. *English Teaching & Learning*, 46(1), 39-57.
- Hay, P., & Penney, D. (2013). *Assessment in Physical Education: A Sociocultural Perspective*. London: Routledge.
- Hamid, M. O. (2011). Planning for failure: English and language policy and planning in Bangladesh. In J. A. Fishman & O. Garcia (Eds), *Handbook of language and ethnic identity-the success-failure continuum in language and ethnic identity efforts* (Vol. 2, 2nd ed., pp. 192–203). New York: Oxford University Press.
- Hattie, J. (2008). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Isik, A. (2021). Exploring How ELT Teachers Perceive and Practice English Language Assessment. *Journal of Language and Education*, 7(1), 109-126. <https://doi.org/10.17323/jle.2021.10296>
- Islam, S. M. A. (2015). *Language policy and practice in secondary school contexts in Bangladesh; challenges to the Implementation of Language-in-Education Policy (Unpublished doctoral dissertation)*. Aalborg: Aalborg University.
- Islam, M. S., Hasan, M. K., Sultana, S., Karim, A., & Rahman, M. M. (2021). English language assessment in Bangladesh today: principles, practices, and problems. *Language Testing in Asia*, 11, 1-21.
- Ismail, S. M., Nikpoo, I., & Prasad, K. D. V. (2023). Promoting self-regulated learning, autonomy, and self-efficacy of EFL learners through authentic assessment in EFL classrooms. *Language Testing in Asia*, 13(1), 27.
- Inbar-Lourie, O. (2013). Guest editorial to the special issue on language assessment literacy. *Language Testing*, 30(3), 301–307 <https://doi.org/10.1177/0265532213480126>.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: a focus on language assessment courses.
- Karim, H. B. B. A. (2004). *A study of teachers' perceptions of factors affecting curriculum change (unpublished doctoral dissertation)*. Penang: Universiti Sains Malaysia.
- Kao, Y. T. (2023). Empowering preservice English teachers with language assessment literacy concepts and practices: Application of Vygotskian concept-based language instruction. *The Modern Language Journal*, 107, 68-94.
- Malone, M. (2008). Training in language assessment. In E. Shohamy & N. Hornberger (Eds.), *Encyclopedia of language and education: language testing and assessment* (2nd



ed., Vol. 7, pp. 225–233). New York: Springer. [https://doi.org/10.1007/978-0-387-30424-3\\_178](https://doi.org/10.1007/978-0-387-30424-3_178).

- Murray, M. (2003). Narrative psychology and narrative analysis. In P. Camic, J.E. Rhodes, & L. Yardley (Eds.), *Qualitative research in psychology: expanding perspectives in methodology and design* (pp. 95-112). Washington, DC: American Psychological Association. Doi: 10.1037/10595-006
- McComas, W. F., & Abraham, L. (2004). Asking more effective questions. *Rossier School of Education*, 1-16.
- O'Loughlin, K. (2006). Learning about second language assessment: Insights from a postgraduate student online subject forum. *University of Sydney Papers in Tesol*, 1, 71-85.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Phuong, H. Y., Phan, Q. T., & Le, T. T. (2023). The effects of using analytical rubrics in peer and self-assessment on EFL students' writing proficiency: a Vietnamese contextual study. *Language Testing in Asia*, 13(1), 42.
- Rahman, A., & Khan, R. (2021). English language testing practices at the secondary level: A case study from Bangladesh. *Challenges in language testing around the world: Insights for language test users*, 447-464.
- Rahman, S. (2015). English language policy initiatives and implementation in Bangladesh: Micro political issues. *Asian EFL Journal*, 88, 59-96. *Language Testing*, 25(3), 385–402. <https://doi.org/10.1177/0265532208090158>.
- Selim, A., & Mahboob, T. S. (2001). ELT and English language teachers of Bangladesh: a profile. In F. Alam, N. Zaman, & T. Ahmed (Eds.), *Revisioning English in Bangladesh* (pp. 141–151). Dhaka: The University Press Limited.
- Singh, C. K. S., & Samad, A. A. (2013). Portfolio as an Assessment Tool and Its Implementation in Malaysian ESL Classrooms: A Study in Two Secondary Schools. *Pertanika Journal of Social Sciences & Humanities*, 21(4), 1255–1273.
- Sohrabi, Z., Ghanbari, N., & Abbasi, A. (2022). Iranian EFL teachers' perceptions of writing assessment literacy: a countrywide study. *Language Testing in Asia*, 12(1), 23.
- Stiggins, R. (1997). *Student-centred classroom assessment*. Upper Saddle River, NJ: Prentice Hall.
- Sultana, N. Test review of the English public examination at the secondary level in Bangladesh. *Lang Test Asia* 8, 16 (2018). <https://doi.org/10.1186/s40468-018-0068-1>
- Sultana, N. Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9, 1 (2019). <https://doi.org/10.1186/s40468-019-0077-8>.
- Taylor L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21–36. <https://doi.org/10.1017/S0267190509090035>
- Tofade, T., Elsner, J., & Haines, S. T. (2013). Best practice strategies for effective use of questions as a teaching tool. *American journal of pharmaceutical education*, 77(7).
- Wicaksono, B. H., Ismail, S. M., Sultanova, S. A., & Abeba, D. (2023). I like language assessment: EFL learners' voices about self-assessment, self-efficacy, grit tendencies,

academic resilience, and academic demotivation in online instruction. *Language Testing in Asia*, 13(1), 37.

Weng, F. (2023). EFL teachers' writing assessment literacy: Surveying teachers' knowledge, beliefs, and practises in China. *Porta Linguarum Revista Interuniversitaria de Didáctica de las Lenguas Extranjeras*, (40), 57-74.

Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *Tesol Quarterly*, 43(3), 492-513.

Zulaiha, S., Mulyono, H., & Ambarsari, L. (2020). An Investigation into EFL Teachers' Assessment Literacy: Indonesian Teachers' Perceptions and Classroom Practice. *European Journal of Contemporary Education*, 9(1), 189-201.