



<https://doi.org/10.36078/987655576>

**Received:** May 11, 2026  
**Accepted:** June 23, 2026  
**Available online:** June 25, 2026

Review article / Sharh maqola

## ASSESSMENT COMPETENCE: TEACHERS' PERSPECTIVES AND AN ANALYSIS OF EXISTING CHALLENGES IN UZBEKISTAN

**Kamola MURADKASIMOVA**

Associate Professor, Doctor of Sciences (DSc) in Pedagogy  
Uzbekistan State World Languages University  
Tashkent, Uzbekistan  
Email: [kmuradkasimova@yahoo.com](mailto:kmuradkasimova@yahoo.com)  
ORCID iD: 0000-0003-4263-9470

### ABSTRACT

Learner assessment is an integral component of the foreign-language teaching and learning process and a powerful decision-making instrument; yet teachers in higher education institutions are frequently not trained sufficiently to make assessment decisions consistent with the core principles of sound assessment.

The aim of this study is to examine English as a Foreign Language (EFL) teachers' perceptions of student assessment, to identify the existing gaps in their assessment training, and to propose recommendations for closing these gaps. To achieve this aim, the study set the following tasks: to determine the methods of assessment teachers actually use; to establish the extent to which their assessment is aligned with the Common European Framework of Reference for Languages (CEFR); and to elicit teachers' views on final and exit examinations.

A mixed-methods design was employed. The qualitative strand explored teachers' perceptions and beliefs about assessment, while the quantitative strand used a test built around authentic, contextualized scenarios to profile teachers' competence. Data were collected from 120 EFL teachers at three language universities in Uzbekistan through structured interviews.

The results reveal, first, the teachers' language assessment literacy (LAL) profile and, second, a markedly stronger need for training in assessment practice than in assessment theory a pattern explained by the gap between subject-matter knowledge and the pedagogy of assessment. Teachers rely heavily on informal, participation-based formative assessment and demonstrate limited experience in aligning assessment with the CEFR.

The study concludes that consistency in teachers' assessment literacy can be achieved through purpose-built courses centred on classroom assessment practice, organized hierarchically so that pre-service programmes build subject knowledge and in-service programmes develop assessment pedagogy.

**Key words:** language assessment literacy, teachers' perceptions, EFL teachers, classroom assessment, formative assessment, CEFR alignment, teacher training, pre-service education, in-service education, higher education.

### For citation:

Muradkasimova, K. (2026). Assessment competence: Teachers' perspectives and an analysis of existing challenges in Uzbekistan. *Philology Matters*, 56(2), 154–169. <https://doi.org/10.36078/987655576>.



## **BAHOLASH KOMPETENSIYASI: O'QITUVCHILAR QARASHLARI VA O'ZBEKISTONDAGI MAVJUD MUAMMOLAR TAHLILI**

**Kamola MURADKASIMOVA**

Pedagogika fanlari doktori (DSc), dotsent  
O'zbekiston davlat jahon tillari universiteti  
Toshkent, O'zbekiston

### **ANNOTATSIYA**

Ta'lim oluvchilarni baholash chet tilini o'qitish va o'rganish jarayonining ajralmas qismi hamda ta'limga oid qarorlar qabul qilishning muhim vositasidir. Biroq oliy ta'lim muassasalarida faoliyat yuritayotgan o'qituvchilar ko'pincha sifatli baholashning asosiy tamoyillariga mos keladigan baholash qarorlarini qabul qilish uchun yetarli tayyorgarlikka ega emas.

Mazkur tadqiqotning maqsadi ingliz tilini chet tili sifatida (EFL) o'qituvchi pedagoglarning talabalarni baholashga oid qarashlarini o'rganish, ularning baholash bo'yicha tayyorgarligidagi mavjud kamchiliklarni aniqlash hamda ushbu kamchiliklarni bartaraf etish yuzasidan tavsiyalar ishlab chiqishdan iborat. Ushbu maqsadga erishish uchun quyidagi vazifalar belgilandi: o'qituvchilar amalda qo'llayotgan baholash usullarini aniqlash; baholashning Chet tillarni bilish bo'yicha umumyeвроpa kompetensiyalari (CEFR) bilan qay darajada uyg'unlashtirilganini belgilash hamda o'qituvchilarning yakuniy va bitiruv imtihonlariga oid fikrlarini aniqlash.

Tadqiqotda aralash metodlar (mixed-methods) yondashuvi qo'llanildi. Sifat tahliliga asoslangan qismda o'qituvchilarning baholashga oid qarashlari va yondashuvlari o'rganildi, miqdoriy qismda esa haqiqiy va kontekstlashtirilgan vaziyatlarga asoslangan test orqali ularning baholash kompetensiyasi aniqlandi. Ma'lumotlar O'zbekistondagi uchta til universitetida faoliyat yuritayotgan 120 nafar EFL o'qituvchisidan tuzilmashtirilgan intervyular yordamida to'plandi.

Natijalar, birinchidan, o'qituvchilarning tilni baholash savodxonligi (Language Assessment Literacy – LAL) darajasini, ikkinchidan esa baholash nazariyasiga nisbatan baholash amaliyoti bo'yicha tayyorgarlikka ehtiyoj ancha yuqori ekanligini ko'rsatdi. Bu holat fan mazmuni bo'yicha bilimlar bilan baholash pedagogikasi o'rtasidagi tafovut bilan izohlanadi. O'qituvchilar asosan norasmiy, ishtirokka asoslangan formativ baholashdan foydalanadilar hamda baholashni CEFR talablari bilan uyg'unlashtirish bo'yicha cheklangan tajribaga ega.

Tadqiqot xulosasiga ko'ra, o'qituvchilarning baholash savodxonligini izchil rivojlantirish sinfdagi baholash amaliyotiga yo'naltirilgan maxsus o'quv kurslari orqali ta'minlanishi mumkin. Bunday kurslar iyerarxik tarzda tashkil etilishi, ya'ni xizmatgacha bo'lgan (pre-service) ta'lim dasturlari fan bo'yicha bilimlarni shakllantirishi, xizmat davridagi (in-service) malaka oshirish dasturlari esa baholash pedagogikasini rivojlantirishga qaratilishi lozim.

**Kalit so'zlar:** tilni baholash savodxonligi (LAL), o'qituvchilarning qarashlari, EFL o'qituvchilari, sinfdagi baholash, formativ baholash, CEFR bilan uyg'unlashtirish, o'qituvchilarni tayyorlash, xizmatgacha bo'lgan ta'lim, xizmat davridagi malaka oshirish, oliy ta'lim.

## **ОЦЕНОЧНАЯ КОМПЕТЕНТНОСТЬ: ВЗГЛЯДЫ УЧИТЕЛЕЙ И АНАЛИЗ СУЩЕСТВУЮЩИХ ПРОБЛЕМ В УЗБЕКИСТАНЕ**

**Камола МУРАДКАСИМОВА**

Доктор педагогических наук (DSc), доцент  
Узбекский государственный университет мировых языков  
Ташкент, Узбекистан

### **АННОТАЦИЯ**

Оценивание обучающихся является неотъемлемым компонентом процесса преподавания и изучения иностранного языка, а также важным инструментом принятия педагогических

решений. Вместе с тем преподаватели высших учебных заведений нередко недостаточно подготовлены к принятию оценочных решений, соответствующих основополагающим принципам качественного оценивания.

Цель настоящего исследования заключается в изучении представлений преподавателей английского языка как иностранного (EFL) об оценивании обучающихся, выявлении существующих пробелов в их подготовке в области оценивания и разработке рекомендаций по их устранению. Для достижения поставленной цели были определены следующие задачи: выявить методы оценивания, фактически используемые преподавателями; определить степень соответствия применяемого ими оценивания требованиям Общеввропейских компетенций владения иностранным языком (CEFR); а также изучить мнения преподавателей относительно итоговых и выпускных экзаменов.

В исследовании использовался смешанный дизайн (mixed-methods). Качественный этап был направлен на изучение представлений и педагогических установок преподавателей в отношении оценивания, тогда как количественный этап предусматривал использование теста, основанного на аутентичных контекстуализированных ситуациях, для определения уровня их оценочной компетентности. Данные были собраны посредством структурированных интервью с участием 120 преподавателей английского языка как иностранного из трёх языковых университетов Узбекистана.

Результаты исследования позволили, во-первых, определить профиль языковой оценочной грамотности преподавателей (Language Assessment Literacy, LAL), а во-вторых, выявить значительно более высокую потребность в подготовке по практике оценивания по сравнению с его теоретическими аспектами, что обусловлено существующим разрывом между предметными знаниями и педагогикой оценивания. Установлено, что преподаватели преимущественно используют неформальное формирующее оценивание, основанное на активности обучающихся, и обладают ограниченным опытом согласования процедур оценивания с требованиями CEFR.

В заключение сделан вывод о том, что последовательное развитие оценочной грамотности преподавателей может быть обеспечено посредством специально разработанных курсов, ориентированных на практику оценивания в учебной аудитории и организованных по иерархическому принципу: программы додипломной подготовки должны формировать предметные знания, тогда как программы повышения квалификации — развивать педагогику оценивания.

**Ключевые слова:** языковая оценочная грамотность (LAL), представления преподавателей, преподаватели английского языка как иностранного (EFL), оценивание в учебной аудитории, формирующее оценивание, согласование с CEFR, подготовка преподавателей, додипломное образование, повышение квалификации, высшее образование.

## INTRODUCTION

Teaching and learning a foreign language have become a necessity of the modern world. In the educational sphere, great attention is paid to the development of the teaching and learning process of foreign languages. A variety of seminars, workshops and training courses are organized at different institutions and levels to develop areas such as professional development, teacher training, materials design and others. Considerable work has been done in implementing new methodologies in foreign-language teaching. New methods, techniques and approaches help to develop learners' knowledge, habits and skills in the foreign language. Today's methodology requires the implementation of authentic tasks and the use of authentic language so that learners can transfer what they have learned to real-life communication. This process requires teachers to plan and deliver their lessons effectively, ensuring that learning objectives, instructional activities and assessment are aligned. This triangle

of objectives, instruction and assessment plays a crucial role in the teaching, learning and assessment of foreign-language learners.

Assessment is an integral part of the language-learning process. It is therefore very important for teachers to be assessment literate and to design successful assessment tasks, since sound assessment is the basis for classroom administering and scoring, for aiding instructional decisions, for communicating results to stakeholders, and for being aware of inappropriate and unethical uses of tests [Fulcher, 2012; 115]. Assessment may thus be regarded as a universal attribute of social life. Throughout history, people have been put to the test to prove their capabilities or to establish their credentials [McNamara, 2000; 3]. The Russian scholar V.S. Avanesov observes that any attempt to trace the origin of tests is comparable to seeking the source of a river [Avanesov, 2002]. This clearly demonstrates that the use of tests has a long history, and scholars agree there is no concrete date for its origin. From the earliest periods of history, people have tried to display their capabilities and skills and compete with one another, and each such attempt was evaluated by others. Tests that establish how a person performs in relation to a threshold of performance have become important social institutions, fulfilling a gatekeeping function by controlling entry to many important social roles. We may therefore confidently state that testing and assessment have existed since the appearance of humankind.

The most common and obvious distinction in how language testing is used is drawn between objective (right/wrong) and subjective (judgmental) assessment techniques. It is more useful, however, to view this distinction in its historical context, because language testing and language teaching have developed in parallel throughout history; any change in language teaching, therefore, affects language testing. Assessment should provide teachers, students, parents and stakeholders with evidence of how well students have learned what they were intended to learn. This article reports the results of a study that examines teachers' perceptions of assessment, identifies gaps in teacher training, and offers recommendations to narrow them.

In Uzbekistan, the role of foreign-language assessment has become particularly significant in recent years. National reforms aimed at raising the quality and reach of foreign-language learning have introduced proficiency requirements and incentives for teachers and have placed assessment at the centre of educational quality assurance. In this context, the question of whether teachers themselves possess the knowledge, skills and principles needed to assess learners validly, reliably and fairly becomes especially pressing. Yet relatively little empirical research has examined the language assessment literacy of university teachers in the country. The present study addresses this gap by investigating how EFL teachers at language universities understand and practise assessment, and by identifying the areas where their preparation falls short.

## **Literature review**

### **Conceptualizing assessment**

There are diverse views on the operationalization of assessment terminology in the literature, and these often cause confusion. In the teaching and learning process, terms

such as “assessment”, “evaluation” and “test” are frequently used interchangeably. The confusion arises because the three terms are closely related to one another. R.Berry notes that a test often results in numbers, whereas assessment encompasses different strategies – including testing – as a means of gathering information about students’ learning progress. Evaluation, by contrast, is the value judgment and interpretation of the outcomes of the data-collection process within a decision-making context [Berry, 2008; 6]. According to C.Coombe, K.Folse and N.Hubley, the term “evaluation” is all-inclusive and constitutes the broadest base for collecting information in education; evaluation considers all the factors that influence the process of learning and teaching, such as syllabus objectives, course design and materials [Harris & McCann, 1994]. Assessment is part of evaluation, since it is concerned with the student and with what the student does [Brindley, 1989]. Assessment includes various ways of collecting data about students’ achievement and progress in the learning process, while a test is a subcategory of assessment – usually a systematic and formal (paper-and-pencil) procedure used to obtain information about students’ behaviour [Coombe et al., 2007]. These three terms can therefore be understood as nested levels: evaluation is the widest concept, assessment is contained within it, and the test is one of the instruments through which assessment is carried out. This nested relationship is now well established in the language-testing literature [Fulcher & Davidson, 2007; Brown & Abeywickrama, 2010].

### **The evolution of language testing**

Years of experience show that the methodology of teaching foreign languages progresses and develops from generation to generation. It is important to note that, as methods of teaching change, the assessment of language ability changes accordingly: traditional methods require traditional assessment, and alternative tasks require alternative assessment. It should also be noted that the evolution of language testing lags behind the evolution of language-teaching methodology. The British specialist R.West states that the evolution of language teaching and language testing began around the Second World War and falls into three phases or generations, which broadly parallel three stages in the evolution of language teaching:

- Stage 1 – the grammar/translation approach;
- Stage 2 – the audio-lingual approach;
- Stage 3 – the communicative approach.

Most writers agree that the generations of language teaching parallel the generations of language testing, although they do not agree on what these generations should be called. Some of the labels suggested by different writers are presented in Table 1 [West, 1997; 29-30] (See Table 1):

**Table 1.** Labels suggested for the three generations of language testing

| <b>Generation</b> | <b>B.Spolsky<br/>(1984)</b> | <b>J.B. Carroll &amp; P.Hall<br/>(1985)</b> | <b>K.Morrow (1979)</b> | <b>H.S. Madsen<br/>(1983)</b> |
|-------------------|-----------------------------|---------------------------------------------|------------------------|-------------------------------|
| 1                 | traditional                 | grammatico-literacy                         | Garden of Eden         | intuitive                     |
| 2                 | modern                      | psycho-linguistic                           | Vale of Tears          | scientific                    |
| 3                 | post-modern                 | socio-communicative                         | Promised Land          | communicative                 |

Even though the writers suggest different names for the generations, the content remains the same.

### **First-generation tests**

R. West agrees with Spolsky in referring to the first generation as “traditional”, since both teaching and testing were based on translation from the first language (L1) to the target language (TL) [West, 1997]. The common techniques used to assess learners were the translation of passages mostly literary texts from and into the target language; lengthy compositions written without context and based on product-oriented writing (for example, “Write 150 words about your hobby”); and oral interviews that often-posed questions without meaning or sequence. On this basis, several conclusions may be drawn about first-generation tests:

- The tasks were not authentic. While translation is used in real life, we do not translate literary extracts; moreover, when we write and speak there is normally a purpose, yet no such purpose can be seen in first-generation tasks.
- These tasks and techniques lacked context. When students write a few words about a topic, they do not know why they are writing, who they are writing for, or how they are supposed to write, and they therefore lack information about the language appropriate to the context.
- The scoring was subjective, requiring assessors to make their own judgments based on personal opinion and experience. On this basis, first-generation language testing and assessment cannot be considered reliable.

### **Second-generation tests**

Second-generation assessment techniques set out to eliminate the shortcomings of the first generation, especially the problems associated with subjectivity. Ways of assessing language on a right/wrong basis were therefore explored. These techniques were considered much more reliable and fairer, because experimental results showed that the assessment could be replicated and yield the same results: factors such as who assessed the item, when it was assessed, and the mood of the assessor no longer mattered. This means that second-generation test items could be marked by people who were not teachers of English, or even by people who did not understand English, and could also be marked by computers. The most common and famous test item was the multiple-choice question (MCQ). The following characteristics of second-generation tests may be identified:

- The techniques were entirely objective, designed to overcome the shortcomings of the subjective first-generation items.
- The multiple-choice format is even more disembodied than first-generation tasks. There is no context in multiple-choice questions, which is a serious problem because distractors may confuse or penalize the better students, who have broader knowledge and the imagination to see other possibilities.
- The same applies to authenticity: second-generation tests are less authentic than first-generation tests. We do not make discrete language choices in real-world communication, so this type of test does not prepare students for real-life language use.



### **Third-generation tests**

The main aim of communicative teaching and communicative testing is the use of real-world situations. The quotation presented by B.J. Carroll clearly sets out the aim of the communicative approach: the communicative approach stands or falls by the degree of real-life, or at least life-like, communication that is achieved in the foreign language in the classroom situation [Carroll, 1982]. The third-generation approach was a genuine reaction against the previous approaches. Some characteristics of third-generation tests are as follows:

- Third-generation techniques are as authentic as possible. Authenticity refers both to texts (selected from real-life sources such as newspapers, radio and lectures) and to tasks (which attempt to simulate the language tasks needed in real life). Total authenticity may be impossible, but learners will at least be able to use language that is relevant to situations beyond the classroom.

- Because third-generation tests attempt to be authentic and to elicit language as it should be used in real life, all third-generation techniques are contextualized. For example, students of English for Specific Purposes (ESP) are required to produce the skills demanded by their future professional contexts.

- Since third-generation tasks are concerned with all four skills, they should assess all four skills in an integrated way. Listening and reading can be assessed objectively, but speaking and writing cannot; the main task of testers has therefore been to improve the reliability of subjective assessment [Bachman & Palmer, 1996; Hughes, 2003], through measures such as using more than one assessor, retraining assessors frequently, and using banding descriptors.

### **Assessment literacy**

In today's world we encounter many new terms connected with literacy: computer literacy, media literacy, health literacy, scientific literacy, statistical literacy and others. What, then, does "literacy" mean? The word is conventionally defined as the ability to read and write, but its meaning varies depending on the context of use. Language assessment literacy (LAL) represents a broadening of the traditional association of literacy with reading and writing skills. LAL may be understood as a competence that enables an individual to understand, evaluate and create language tests and to analyse test data. The majority of teachers today know relatively little about educational assessment. For some teachers, the word "test" is, both figuratively and literally, a four-letter word [Popham, 2009; 5]. The most obvious explanation for this situation was partly noted above: teachers complete teacher-education programmes that specialize in teaching and in exploring new methods and techniques of teaching methodology, yet there is no requirement to study educational assessment. Only a few sessions may be conducted, which is not enough to produce an assessment-literate teacher [La Marca, 2006]. R.J. Stiggins had already warned that teachers entering the profession in the twenty-first century would need to be assessment literate [Stiggins, 1995], and Inbar-Lourie subsequently argued for a dedicated language-assessment knowledge base built around the questions of why, what and how to assess [Inbar-Lourie, 2008].

One potentially fruitful approach is to import conceptualizations of literacy from other fields into LAL. An approach adopted in the fields of scientific and mathematical literacy education is the rejection of a dichotomy between literacy and illiteracy in favour of viewing literacy as a continuum. A seminal work is R.W. Bybee’s *Achieving Scientific Literacy: From Purpose to Practices* [Bybee, 1997], in which five stages of literacy are identified and described: illiteracy, nominal literacy, functional literacy, procedural and conceptual literacy, and multidimensional literacy [Pill & Harding, 2013; 383]. A brief description of each stage is presented in Table 2:

**Table 2.** Stages of literacy as a continuum

| Stage of literacy                  | Description                                                                                                               |
|------------------------------------|---------------------------------------------------------------------------------------------------------------------------|
| Illiteracy                         | ignorance of language assessment concepts and methods                                                                     |
| Nominal literacy                   | understanding that a specific term relates to assessment, but with a possible misconception                               |
| Functional literacy                | sound understanding of basic terms and concepts                                                                           |
| Procedural and conceptual literacy | understanding of central concepts, and the use of that knowledge in practice                                              |
| Multidimensional literacy          | knowledge extending beyond ordinary concepts, including the philosophical, historical and social dimensions of assessment |

Let us consider some further definitions provided by scholars. G.Fulcher notes that the earliest attempt to define assessment literacy for teachers was made by the American Federation of Teachers (1990), which included competencies in “selecting and developing assessments for the classroom, administering and scoring tests, using scores to aid instructional decisions, communicating results to stakeholders, and being aware of inappropriate and unethical uses of tests” [Fulcher, 2012; 115]. Subsequently, R.J. Stiggins defined assessment literacy as “having a basic understanding of the meaning of high- and low-quality assessment and being able to apply that knowledge to various measures of student achievement” [Stiggins, 1991; 535]. Drawing on these definitions, we may regard assessment literacy as an understanding of both the theory and the practice of assessment – the principles and methods related to all aspects of assessing students’ learning in the classroom.

### **Dimensions and components of language assessment literacy**

More recent scholarship has sought to map the components of language assessment literacy in greater detail and to recognize that different stakeholders require different profiles of competence. L.Taylor argues that LAL is not a single, uniform construct but a profile made up of several dimensions – including knowledge of theory, technical skills, principles and concepts, language pedagogy, sociocultural values, local practices, and scores and decision-making – and that the relative weighting of these dimensions differs for test writers, classroom teachers, administrators and other stakeholders [Taylor, 2013]. Earlier component models had similarly distinguished the knowledge, skills and principles that underpin assessment competence [Davies, 2008], and O.Inbar-Lourie framed these as a language-assessment knowledge base organized around the purposes, content and methods of assessment [Inbar-Lourie,



2008]. M.E. Malone drew a useful contrast between the needs of professional test developers and those of test users such as classroom teachers [Malone, 2013], while B.Kremmel and L.Harding developed and validated a survey instrument that profiles assessment literacy across different stakeholder groups [Kremmel & Harding, 2020]. For classroom teachers, the dimensions most heavily weighted are language pedagogy and classroom-based assessment skills, rather than the advanced psychometric knowledge required of professional test developers. This profiling approach is valuable because it reframes the development of teachers' LAL: rather than expecting every teacher to master the full technical apparatus of measurement, it focuses attention on the specific competences teachers need in order to integrate assessment into everyday teaching.

This view is consistent with the continuum model described above and with the position that LAL develops gradually and unevenly. A teacher may possess functional literacy in some areas – for example, an understanding of basic terminology – while remaining at the nominal level in others, such as the design of valid speaking and writing tasks or the interpretation of results in relation to external standards. Recognizing this unevenness is important for the design of teacher-education programmes, because it implies that training must be targeted at the particular gaps in teachers' profiles rather than delivered as a single, generic course. Empirical studies of foreign-language teachers in Europe have confirmed that their assessment-training needs are uneven and often unmet [Vogt & Tsagari, 2014], and A.Scarino has emphasized that developing LAL also involves teachers becoming aware of the interpretive frameworks they bring to assessment [Scarino, 2013]. It also implies that the development of LAL is best understood as a long-term, career-spanning process that begins in pre-service education and continues through in-service professional development. These conceptual considerations directly informed the present study, which sought to establish the LAL profile of EFL teachers in Uzbekistan and to locate the specific points at which their preparation requires strengthening. To establish the level of teachers' assessment literacy and the types of misconception they hold about language testing, the study was conducted with EFL teachers at language universities in Uzbekistan.

The aim of the current study is to interview university teachers in order to identify their current assessment practices and the conditions under which they work. The main objectives were to determine whether the course content, tasks and assessment are aligned, what methods of assessment are used, and what obstacles teachers face in assessing their students.

- What is the level of language assessment literacy of in-service EFL educators in Uzbekistan?

- What gaps are observed in the conceptual and practical knowledge of assessment of both pre- and in-service teachers?

## **METHODS**

The study is based on 120 structured interviews with EFL teachers from

three language universities in Uzbekistan. A mixed-methods design was adopted: the qualitative component explored teachers' perceptions and beliefs about student assessment, while the quantitative component used a test built around authentic and contextualized scenarios to profile teachers' assessment competence. The design drew on established principles of classroom-based and teacher assessment [Brown & Abeywickrama, 2010; Hill & McNamara, 2012]. The structured interviews and the accompanying instrument were designed to gather data on three main areas:

- the methods of assessment teachers use;
- the alignment of their assessment with the CEFR;
- teachers' views about the final and exit examinations.

The responses were collated and analysed to identify recurring patterns, to rank the assessment activities reported by teachers, and to describe the grading systems in use. The combination of qualitative and quantitative data made it possible both to describe teachers' beliefs and to profile their practical competence against authentic assessment situations.

# RESULTS

## Methods of assessment

The data gathered from teachers referred mostly to methods of formative assessment, which was largely informal in nature. In their answers, teachers focused on the daily or ongoing assessment of their learners, where the main factors were students' participation and readiness rather than academic work. The data reveal that teachers commonly rely on "participation" to evaluate their students, paying attention to how actively students take part in class discussions as well as to task performance. Respondents also spoke in general terms about the classroom activities used to assess students' preparation and knowledge; however, there is little evidence of how these activities relate to the syllabus. The activities mentioned as being used to assess students are shown in Table 3 (See Table 3):

**Table 3.** Methods of formative assessment (N = 120)

| Type of activity          | Number of respondents (rank) |
|---------------------------|------------------------------|
| Knowledge-based MCQ tests | 19 (1)                       |
| Tasks on the four skills  | 18 (2)                       |
| Oral tasks                | 16 (3=)                      |
| "Can-do" evaluation       | 16 (3=)                      |
| Participation             | 15 (5)                       |
| Grammar tasks             | 10 (6)                       |
| Written tasks             | 9 (7)                        |
| Vocabulary tasks          | 7 (8)                        |
| Translation tasks         | 6 (9)                        |
| Presentation              | 5 (10=)                      |
| Self-study                | 5 (10=)                      |
| Project work              | 4 (12)                       |
| Portfolio                 | 3 (13)                       |
| Quizzes                   | 1 (14)                       |

Some respondents (15) regarded the assessment criteria as a rating system common to all universities across the country and stated that they assess students according to the systems shown in Table 4 (See Table 4):

**Table 4.** Assessment grading systems (N = 120)

| Rank | Current / mid-term works                                                              | Self-study | Final / Summative | Total | Respondents |
|------|---------------------------------------------------------------------------------------|------------|-------------------|-------|-------------|
| 1.   | 3 points per lesson $\times$ 18 weeks = 54                                            | 16         | 30                | 100   | 7           |
| 2.   | 10 points per month $\times$ 4 months = 40;<br>15 points per mid-term $\times$ 2 = 30 | –          | 30                | 100   | 5           |
| 3.   | 10 points per month $\times$ 3 months = 30;<br>15 points per mid-term $\times$ 2 = 30 | 10         | 30                | 100   | 3           |

Every higher education institution develops its own rating system, aligned with the Ministry curriculum for assessment, which is usually organized on a 100-point scale per semester. The rating consists of three types: self-study (daily assessment), midterm (formative assessment) and final (summative assessment). One semester consists of daily assessments, two midterms and one final test. All the tests and assessment tasks are designed on the basis of the course content. The prominence of participation and knowledge-based multiple-choice tests in the ranking suggests that, despite familiarity with the language of formative assessment, teachers' practices remain oriented towards informal monitoring and towards the assessment of receptive knowledge rather than towards systematic, criterion-referenced assessment of productive skills.

#### **Assessment alignment with the CEFR**

The Common European Framework of Reference for Languages provides a common basis for describing language proficiency and for relating assessments to one another (Council of Europe, 2001). Those interviewed about how assessment relates to the CEFR responded in a variety of ways. Some teachers (22%) stated that assessment could be carried out by integrating the four skills (speaking, reading, writing and listening); others (18%) suggested that aligning the assessment process with the CEFR would require a great deal of time and effort; some (20%) stated that the existing rating system, which does not relate to the CEFR, is not efficient; some (22%) recommended that assessment should follow the CEFR, since it is the international standard in most countries; and some (18%) stated that students' active participation and class attendance could provide the expected results. These results indicate that teachers have little experience in aligning assessment with the CEFR and that, in some cases, they have no clear idea of what CEFR-based assessment involves.

#### **Teachers' views on the final and exit examinations**

Among the teachers interviewed, 26 (22%) did not comment on the current exit-test forms and processes because they are not involved in them. Ninety-four (78%) expressed positive views, stating that the final and exit tests are aligned with the Ministry curriculum and the institutional syllabus; 43 (36%) stated that all language skills are included; while 45 (38%) said that the tests are basically knowledge-based,

addressing translation, sentence construction, reading and retelling the content. Two respondents stated that writing is the main skill assessed. Worryingly, most respondents felt satisfied with examination processes that largely do not use criteria or “can-do” statements and that are usually English for General Purposes (EGP) rather than ESP in content, and they expressed no wish to change these processes or even to attempt to align the tests with the CEFR. The reason for this is a lack of knowledge about assessment.

As suggested by Stiggins (1999), a list of seven content requirements or competences is intended to provide a comprehensive foundation in assessment practices. By following these requirements, teachers will be able to carry out successful assessment procedures. They include:

- connecting assessment to clear purposes;
- clarifying achievement expectations;
- applying appropriate assessment methods;
- developing quality assessment exercises and scoring criteria and sampling appropriately;
- avoiding bias in assessment;
- communicating effectively about student achievement;
- using assessment as an instructional intervention [Herrera & Macías, 2015; 307].

## DISCUSSION

The findings of the present study contribute to the growing body of research on language assessment literacy (LAL) by demonstrating that teachers’ assessment competence is shaped not only by their knowledge of assessment terminology but also by their ability to apply assessment principles in authentic classroom contexts. Although most participants demonstrated familiarity with the concepts of formative assessment and continuous assessment, the results suggest that this knowledge does not consistently translate into sound classroom assessment practices. Assessment literacy should therefore be understood as a practice-based competence rather than merely a body of theoretical knowledge. Likewise, the multidimensional model of language assessment literacy explains why teachers may possess adequate conceptual knowledge while experiencing difficulties in classroom implementation [Xu & Brown, 2016; Taylor, 2013].

A second important finding concerns the limited alignment of classroom assessment with the Common European Framework of Reference for Languages (CEFR). While many teachers acknowledged the importance of the CEFR, relatively few demonstrated practical experience in designing CEFR-aligned assessment tasks. Previous studies have similarly reported that insufficient professional preparation, rather than teachers’ unwillingness, constitutes the principal obstacle to the effective implementation of CEFR-based assessment. The present findings therefore suggest that greater emphasis should be placed on developing teachers’ practical competence in criterion-referenced assessment and performance-based assessment [Vogt &

Tsagari, 2014; Tsagari & Vogt, 2017]. However, unlike several European studies, the present study indicates that teachers expressed generally positive attitudes towards the existing examination system despite their limited familiarity with criterion-referenced assessment. This discrepancy may reflect institutional assessment traditions and assessment policies specific to the higher education context of Uzbekistan.

The predominance of informal participation-based formative assessment observed in this study may also reflect institutional assessment traditions and established classroom practices within higher education rather than deliberate pedagogical choices. Although formative assessment is widely recognized as an effective strategy for supporting learning, its educational value depends on clearly defined assessment criteria, constructive feedback, and systematic monitoring of student progress. Without these components, formative assessment may lose much of its diagnostic and developmental function [Hill & McNamara, 2012; Malone, 2013].

From a practical perspective, the findings indicate that teacher education programmes should place greater emphasis on classroom assessment design, rubric development, feedback strategies, CEFR alignment, and authentic performance assessment. Professional development activities should provide teachers with opportunities to design, implement, evaluate, and revise assessment tasks in realistic educational contexts. Such practice-oriented training is likely to contribute more effectively to the development of language assessment literacy than theoretical instruction alone [Malone, 2013].

Several limitations should be acknowledged. First, the study involved teachers from only three language universities in Uzbekistan; therefore, the findings should not be generalized to all higher education institutions without caution. Second, much of the evidence was obtained through structured interviews, which may have been influenced by self-report bias. In addition, the study did not investigate actual classroom assessment practices through direct observation; therefore, the findings primarily reflect teachers' reported practices rather than their observed classroom behaviour. Finally, although the mixed-methods design strengthened the validity of the findings, future studies could incorporate classroom observations, document analysis, and longitudinal research to provide a more comprehensive understanding of teachers' assessment practices and their development over time.

Overall, the findings suggest that improving teachers' language assessment literacy requires systematic and continuous professional development integrating theoretical knowledge with authentic classroom practice. Future research should examine the effectiveness of targeted assessment-literacy programmes, investigate assessment practices across different educational contexts, and explore the relationship between teachers' assessment literacy and student learning outcomes.

### **Recommendations**

On the basis of these findings, several recommendations can be made. The development of teachers' assessment literacy should be treated as a long-term, career-spanning process rather than as a single intervention. To this end, purpose-built courses on teachers' competence in the educational assessment of students should be developed

and embedded in both pre-service and in-service programmes. These courses should be organized hierarchically: pre-service programmes should build a solid foundation of subject-matter knowledge about assessment terminology, principles, validity, reliability and fairness while in-service programmes should focus on the pedagogy of assessment, that is, on the practical design, administration and interpretation of classroom assessment. Framing assessment literacy as a situated, practice-based competence [Xu & Brown, 2016] and grounding it in models of classroom-based assessment [Hill & McNamara, 2012] can help ensure that such courses translate into changes in teachers' everyday practice.

Professional development should be sustained and practice-based rather than confined to occasional sessions. Workshops, seminars and teacher-training courses can give teachers the opportunity to collaborate, to share their experiences, and to develop and trial assessment tasks under guidance. Particular attention should be paid to the assessment of productive skills speaking and writing and to the use of criteria and "can-do" descriptors, since these are the areas in which the study found teachers to be least confident and least experienced. Training in CEFR alignment, in the development of scoring rubrics, and in the moderation of subjective assessment through the use of multiple raters and rater retraining would directly address the gaps identified in the findings. Finally, institutional rating systems should be reviewed so that they reward demonstrated assessment competence and support, rather than constrain, the use of valid and authentic assessment.

## CONCLUSION

Language assessment literacy is an essential professional competence that enables teachers to make valid, reliable, and pedagogically meaningful assessment decisions throughout their careers. The findings of the present study indicate that although EFL teachers demonstrate a general awareness of assessment concepts, important gaps remain in the practical application of assessment principles, particularly in classroom-based assessment, CEFR alignment, and the use of criterion-referenced assessment practices.

The study further suggests that strengthening teachers' language assessment literacy requires systematic and continuous professional development that integrates assessment theory with authentic classroom practice. In particular, teacher education programmes should provide balanced preparation in both the conceptual foundations and the pedagogical implementation of language assessment. A hierarchical model of professional preparation is therefore recommended, in which pre-service programmes establish a solid foundation of assessment knowledge, while in-service professional development focuses on classroom assessment design, feedback practices, rubric development, and CEFR-aligned assessment.

The findings also highlight the importance of designing purpose-built assessment literacy courses that respond to teachers' actual professional needs and support the effective integration of assessment into everyday teaching practice. Such an approach has the potential to improve not only teachers' assessment competence but also the



overall quality, validity, and consistency of language assessment in higher education.

Despite these contributions, the study is limited by its focus on teachers from three language universities in Uzbekistan and by its reliance primarily on self-reported data. Future research should therefore include larger and more diverse samples, combine interview data with classroom observations and document analysis, and examine the impact of targeted assessment-literacy programmes on teachers' classroom practices and student learning outcomes. Such research would provide a more comprehensive understanding of language assessment literacy and contribute to the development of evidence-based teacher education policies and practices.

## REFERENCES

1. Avanesov, V.S. (2002). *Composition of test items* (3rd ed., rev. and enl.). Moscow: Center for Testing.
2. Bachman, L.F., & Palmer, A.S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
3. Berry, R. (2008). *Assessment for learning*. Hong Kong: Hong Kong University Press.
4. Brindley, G. (1989). *Assessing achievement in the learner-centred curriculum*. Sydney: National Centre for English Language Teaching and Research.
5. Brown, H.D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). White Plains, NY: Pearson Education.
6. Bybee, R.W. (1997). *Achieving scientific literacy: From purposes to practices*. Portsmouth, NH: Heinemann.
7. Coombe, C., Folse, K.S., & Hubley, N. (2007). *A practical guide to assessing English language learners*. Ann Arbor, MI: University of Michigan Press.
8. Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
9. Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347.
10. Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132.
11. Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. London: Routledge.
12. Harris, M., & McCann, P. (1994). *Assessment*. Oxford: Heinemann English Language Teaching.
13. Herrera, L., & Macías, D.F. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language. *Colombian Applied Linguistics Journal*, 17(2), 302–312.
14. Hill, K., & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395–420.
15. Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
16. Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402.
17. Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the Language Assessment Literacy Survey. *Language Assessment Quarterly*, 17(1), 100–120.
18. La Marca, P.M. (2006, June). *Assessment literacy: Building capacity for improving student learning*. Paper presented at the National Conference on Large-Scale Assessment, Council of Chief State School Officers, San Francisco, CA.
19. Malone, M.E. (2013). The essentials of assessment literacy: Contrasts between testers and

- users. *Language Testing*, 30(3), 329–344.
20. McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
  21. Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402.
  22. Popham, W.J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48(1), 4–11.
  23. Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309–327.
  24. Stiggins, R.J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534–539.
  25. Stiggins, R.J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238–245.
  26. Stiggins, R.J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23–27.
  27. Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders. *Language Testing*, 30(3), 403–412.
  28. Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41–63.
  29. Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402.
  30. West, R. (1997). *Assessment in language learning: Introduction and principles of language testing* (Unit 0, MD338). Manchester: University of Manchester.
  31. Xu, Y., & Brown, G.T.L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162.