

Research Paper

# The Evolving Concept of (Language) Assessment Literacy. Implications for Teacher Education

Ildikó Csépes<sup>1</sup>

Recommended citation:

Csépes I. (2021). The Evolving Concept of (Language) Assessment Literacy. Implications for Teacher Education. *Central European Journal of Educational Research*, 3(1), 120–130. <https://doi.org/10.37441/CEJER/2021/3/1/9360>

## Abstract

Language teachers' assessment knowledge and skills have received considerable attention from language assessment researchers over the past few decades (Davison & Leung, 2009; Hill & McNamara, 2012; Rea-Dickins, 2001; Taylor, 2013). This seems to be linked to the increased professionalism expected of them in classroom-based assessments. However, teachers seem to face a number of challenges, including how large-scale standardized language exams influence their classroom assessment practices. Teachers' assessment literacy, therefore, needs to be examined in order to explain their assessment decisions. In this paper, we review the concept of (language) assessment literacy, how it has evolved and how it is conceptualized currently. Recent interpretations seem to reflect a multidimensional, dynamic and situated view of (language) assessment literacy. Implications for teacher education are also highlighted by presenting research findings from studies that explored teachers' and teacher candidates' assessment literacy in various educational contexts. As a result, we can identify some common patterns in classroom assessment practices as well as context-specific training needs. Finally, we make a recommendation for tackling some of the challenges language teachers are facing in relation to classroom-based assessment in the Hungarian context.

*Keywords:* assessment literacy; language assessment literacy; pedagogical assessment; teacher education

---

## Introduction

In the 21<sup>st</sup> century multilingual Europe, the knowledge of foreign languages, especially English, has become indispensable for the individual both personally and for professional reasons. In Hungary, a successful language learner has at least a B2-level state-recognized foreign language exam certificate or the equivalent advanced level matriculation examination. The certificate provides benefits for admission to higher education, and it is a requirement for earning a degree. Because of the benefits attached to a B2-level certificate, language assessment affects many, but concerns primarily language teachers and students. The question arises: are they equipped with the necessary knowledge and skills to critically evaluate language tests? Stiggins (1991, 1995) introduced the term 'assessment literacy' (AL), which he defines by characterising those individuals who lack this literacy: they are the ones who are unaware of what it takes to design and implement high quality measurements, and thus they accept the results as well as the decisions based on them without any reservation. The situation is further complicated by the fact that the complex procedures used by measurement specialists are not easily accessible and interpretable by lay people. Stiggins (1991) regards assessment literacy as a complex concept that is highly dependent on the user's needs.

As early as the 1990s, Stiggins already pointed out that there was plenty to do about assessment literacy in the classroom context, as no one prepared students to be critical of their teachers' feedback or classroom testing practice. Presumably not many teachers would have been willing to submit themselves to scrutiny because they did not have adequate training in assessment either. According to Stiggins (1995), teachers were afraid of tests and assessments in general, mainly due to their past bad experiences. Teachers reported that the

---

<sup>1</sup>University of Debrecen, Hungary, [csepes.ildiko@artsunideb.hu](mailto:csepes.ildiko@artsunideb.hu)

tests they had taken often increased their sense of failure because of the qualifying and ranking function attached to them. The situation may have changed since then, but some of Stiggins' claims made nearly three decades ago still seem timely today. The results of several studies (e.g. DeLuca & Klinger, 2010; Lam, 2015; Vogt & Tsagari, 2014; Volante & Fazio, 2007) support the view that there is lack of emphasis in this area in teacher education, thus teachers' assessment literacy appears insufficient.

Over the past two decades, language assessment researchers have paid increased attention to classroom-based assessments (Davison & Leung, 2009; Hill & McNamara, 2012; Poehner, 2009; Rea-Dickins, 2001), trying to define what theoretical and practical knowledge language teachers need in this field. The specialized nature of language teaching and learning justifies use of a more specific term: "Language Assessment Literacy" (LAL), which was introduced by Inbar-Lourie (2008: 385). According to her definition, an individual can be described as literate in language assessment if s/he is able to "ask and answer critical questions about the purpose for assessment, about the fitness of the tool being used, about testing conditions, and about what is going to happen on the basis of the results" (Inbar-Lourie, 2008: 389). However, this was only the first attempt to define the concept of language assessment literacy, and the interpretation of the term has been changing ever since. There is still no widely accepted definition, no agreed upon set of skills or principles shared by the entire educational community (Harding & Kremmel, 2016). The lack of a unified concept can be explained by the constantly changing elements that form the multi-componential view of the concept: there were definitions that included elements of knowledge and skills, but sometimes also principles (e.g. Brindley, 2001; Davies, 2008; Fulcher, 2012). Brindley (2001) suggested a definition that distinguished between core and added components, while others proposed connected knowledge elements (Inbar-Lourie, 2008; Xu & Brown, 2016). Inbar-Lourie (2017) claims that in our endeavour to refine the concept, we are now moving towards an ever-expanding conceptual and practical repertoire, and the creation of situated, differential LAL seems to be more feasible instead of a unified concept.

Based on the idea of Pill and Harding (2013) as well as Taylor (2013), language assessment literacy can be represented on a scale, it is a relative concept. They propose a multidimensional model of language assessment literacy that takes into account the dynamic nature and complexity of the concept. The most important feature of such a LAL model is that it distinguishes between different stakeholder groups that require different levels of expertise in language assessment. In order to fully account for our current conceptualisation of LAL, however, we need to go back in time and examine more closely how the concept of assessment literacy evolved historically, starting our time travel back in the second half of the previous century.

### **Language assessment literacy: a historical perspective**

As early as the 1970s and 1980s, the literature on pedagogical assessment outlined what skills and knowledge elements teachers should have in the field of assessment (Schafer, 1991; Stiggins, 1991). However, a comprehensive description of the requirements appeared only in the 1990s under the title "Standards for Teacher Competence in Educational Assessment of Students" ['the Standards' for short] (American Federation of Teachers, National Council on Measurement in Education and National Education Association, 1990). According to a study 25 years later (DeLuca et al., 2016), the Standards have been really powerful as even today language teachers' assessment literacy is measured with the help of tools that are based on the Standards. The emphasis on what is regarded as important has changed somewhat over the years: the 1990s were characterized by 'teacher-centeredness' (trying to define the competencies needed for teacher assessment), while after the 2000s the student perspective gradually emerged – thanks to concepts like 'assessment for learning' (Black & Wiliam, 1998) and 'assessment as learning' (Earl & Katz, 2005) – in order to involve students in the assessment process.

The Standards (1990) were intended to guide pre- and in-service teacher education programs in the area of assessment. Altogether seven standards were recommended that teachers should be trained in:

1. choosing assessment methods that are appropriate for instructional decisions;
2. developing these methods; administering, scoring, and interpreting externally produced and teacher-produced assessment results;

3. using results of assessments in making decisions for individual students, planning teaching, developing curriculum, and making school improvements;
4. developing valid, assessment-based grading procedures;
5. communicating results of assessments to student, parent, and other lay audiences, and educators;
6. recognizing methods and uses of assessment that are unethical, illegal, or otherwise inappropriate.

For training purposes the Standards had to be broken down, as appropriate, into topics. These were analysed by Schafer (1993) on the basis of teacher manuals published during the given time period. The result of the comparison between the Standards and teacher manuals showed that the fourth requirement – using results of assessments in making decisions for individual students – was missing. However, it was even more disappointing to find that in many teacher education programs pedagogical assessment itself was not given sufficient attention (Schafer, 1993). Because of the increased interest in the learners' perspective in assessment in the 2000s, the Standards formulated in 1990 were extended in 2015 by Klinger et al. (Joint Committee on Standards for Education). The revision and extension of the Standards were warranted by the research findings that had accumulated over the preceding 25 years.

The importance of formative assessment was put in the limelight in the 1990s by Black and Wiliam (1998), who did a meta-analysis of classroom research only to find that students' achievements could also be enhanced through specific forms of classroom assessment. The emergence of the concept of 'assessment for learning' can be attributed mainly to Black and Wiliam, but Stiggins (2002) also provided detailed guidance on the classroom implementation of assessment for learning. According to this, teachers should

- inform students about the learning goals in terms that students understand;
- become assessment literate and thus able to design assessment exercises and scoring procedures that accurately reflect student achievement;
- use classroom assessments to build students' confidence in themselves as learners and help them take responsibility for their own learning;
- translate classroom assessment results of students into descriptive feedback instead of judgmental feedback, providing them with specific insights as to how to improve;
- engage students in regular self-assessment, with standards held constant so that students can watch themselves grow over time and thus feel in charge of their own success.

Based on the above, we can conclude that by the early 2000s pedagogical assessment gave priority to formative assessment, but at the same time, the role of external standardized examinations increased due to globalization and the need for accountability (Taylor, 2009). This meant that teachers had to face challenges both externally and within their own classroom. Performing daily classroom tasks made new demands on teachers because they required greater professionalism, a higher level of assessment literacy on their part. However, training opportunities were very limited. Popham (2004) labelled the lack of assessment literacy in this situation as "professional suicide". Another problem was that most of the language testing courses available at the time provided only an introduction to basic professional knowledge and were concerned with external, standardized examinations rather than classroom assessment, and more specifically formative assessment (Alderson, 1999). The latter claim is also supported by the results of a European survey (Hasselgreen et al., 2004), which was conducted on behalf of the EALTA (European Association for Language Testing and Assessment). The online survey showed that respondents (language teachers from all kinds of educational backgrounds) were in dire need of in-service training, especially in the area of alternative forms of assessment (e.g. the use of portfolio, self- and peer assessment, diagnostic feedback). The limited access to or complete lack of training opportunities was also made evident as respondents expressed their need for training in more conventional areas too, including testing principles (e.g. validity, reliability) and test development skills (e.g. item writing, item statistics).

The most comprehensive description of the components of language assessment literacy to date is based on an analysis of the literature published over several decades by a widely recognized measurement specialist,

Davies (2008). In his view, the components of language assessment literacy include: *skills* (item writing, statistics, test analysis and software programmes for test delivery, analysis and reportage), *knowledge* (relevant background in measurement and language description, as well as in context setting, and may involve an examination of different models of language learning, of language teaching and of language testing), and *principles* (the proper use of language tests, their fairness and impact, including questions of ethics) (Davies, 2008: 328). However, Davies's definition is predominantly for measurement and evaluation professionals, it does not take into account the needs of other stakeholders, thus its validity cannot be generalized. The components of the three dimensions proposed by Davies can also change dynamically, as was evidenced by the content analysis of testing courses across the globe by Brown and Bailey (2008). The authors conducted a repeated questionnaire survey in 2007 in order to update their results obtained 10 years earlier. Data from 2007 showed that the topics covered by testing courses around the world had been expanded to include 20 new topics, although the emphasis continued to be placed upon the same topics: critiquing and analysing tests, measuring the four skills, validity, item analysis (facility value, discrimination index, and content analysis), and basic test statistics (descriptive statistics, reliability, error). There have also been several attempts at measuring language assessment literacy with specially designed tests (Mertler & Campbell, 2005; Newfields, 2006; Plake et al., 1993; Tao, 2014). However, as DeLuca et al. (2016) pointed out, the majority of these tests measured the competencies required for teacher assessment in line with the Standards, which reflected the views on classroom assessment of the 1990s.

Language teachers' assessment literacy can be shaped not only by professional training but also by their social context. Brindley (2001) suggested that teachers need not only learn what and how to evaluate, but they also have to develop their critical sense to properly assess the role and implications of language testing in their own social setting. It is essential to enhance teachers' critical awareness because, for many, external language exams seem automatically valid and authoritative (Alderson, 1999). Thus, critical awareness has become a requirement in language assessment literacy. According to Inbar-Lourie (2008), the critical approach can be developed best if teacher education is characterized by a constructivist learning approach, whereby teachers have the opportunity to discuss, critically analyse their own teaching as well as their social context. O'Loughlin's definition of LAL also emphasises this need for critical awareness by suggesting that it refers to "the acquisition of a range of skills related to test production, test score interpretation and use, and test evaluation in conjunction with the development of a critical understanding about the roles and functions of assessment within education and society" (2013: 363). Fulcher goes even further by proposing that language assessment literacy should also incorporate historical, political and even philosophical aspects. He claims that LAL includes "the ability to place knowledge, skills, processes, principles and concepts within wider historical, social, political and philosophical frameworks in order [to] understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals" (Fulcher, 2012: 125).

The need for a unified concept of LAL was challenged by Taylor (2009, 2013), who asserted that different stakeholders can be characterised by different needs or levels of involvement in language assessment, thus she raised concerns about the depth and not only the width of assessment literacy for specific target groups (see Figure 1.). In an attempt to define language assessment literacy for different stakeholders groups, Taylor (2013: 410) distinguished eight dimensions: (1) knowledge of theory, (2) technical skills, (3) principles and concepts, (4) language pedagogy, (5) sociocultural values, (6) local practices, (7) personal beliefs and attitudes, (8) scores and decision making. Taylor proposes that stakeholders are linked to each dimension to varying degrees, and the depth of their knowledge varies depending on their needs. For example, it is a basic requirement for a testing professional to have a balanced and comprehensive knowledge across all dimensions. However, it is less critical for language teachers to be familiar with measurement and evaluation theories, while it is essential for them to have adequate knowledge of language pedagogy.

**Figure 1.** Levels of AL/LAL differentiated according to stakeholder constituency

(Taylor, 2013: 409)

Unfortunately, the stakeholder groups identified by Taylor exclude a very important group of test users: language learners. Engelsen and Smith (2014) propose that students' language assessment literacy also plays an important role in the outcomes of educational reform or when the goal is to bring about a change in attitudes in pedagogical assessment. If learners are not willing to accept and participate in alternative formative assessment procedures, or they are unable to use the information from such assessments to enhance their learning, all the efforts that were put in the pedagogical reform are unlikely to lead to the desired outcome. Lee and Butler (2020) state that students can play a key role in classroom-based assessment, as they can shape the assessment processes through active participation, yet the language assessment literacy of this stakeholder group has hardly been studied yet. The differing needs of other stakeholders, on the other hand, have been studied extensively (Kremmel & Harding, 2020; Malone, 2013; O'Loughlin, 2013; Pill & Harding, 2013).

The multidimensional view of language assessment literacy can be applied not only across different stakeholder groups, but it can be linked to environmental or contextual factors that are relevant to only one specific stakeholder group. As teachers play a central role among all the stakeholders, researchers explored the impact of contextual factors in relation to this group. Borg (2003) suggested that due to the social and institutional context of classrooms, language teachers may be constrained by a number of factors in their ability to teach (and assess) in line with their beliefs. Borg lists several environmental or contextual factors that can impact on what teachers do: parents, colleagues, school management, society, available infrastructure and resources. Willis et al. (2013) also highlighted the dynamic, socially and contextually determined nature of language assessment literacy. However, the first theoretical framework that aimed to incorporate the impact of contextual variables was proposed by Xu and Brown (2016), who identified the contextual effects at three levels: (i) national, (ii) institutional, (iii) personal level. They proposed a new term for assessment literacy: TALiP (Teacher Assessment Literacy in Practice). TALiP is defined as "a dynamic, complex entity combining teachers' assessment knowledge, their conceptions of assessment, and their responses to the external contexts embedded with actual constraints and affordances in the environment" (2016: 23). Yan et al. (2018) also examined the impact of contextual factors, and as a result they also distinguished three levels naming them as "educational landscape and policies", "institutional mandate", and "local institutional context". Building on the work by Xu and Brown (2016) as well as Yan et al. (2018), Vogt et al. (2020) have retained the three dimensions, but use simpler terms to refer to the levels: (i) macro-level (national /regional level), (ii) meso-level (institutional level) and (iii) micro-level (classroom level). In line with earlier recommendations, Vogt et al. propose that there are complex interactions between the levels: for example, assessment processes in the classroom (at micro-level) can be largely determined by the country- or region-specific assessment culture (macro-level effects) as well as the local educational setting, i.e. the traditions and expectations of the school (meso-level effects). It is important to emphasize, therefore, that although the teacher is the 'agent' of classroom assessment, assessment is shaped by the complex interplay of the three contextual dimensions described above.

Vogt et al. (2020) claim that specific individual characteristics of language teachers, representing the micro-level, may play a crucial role in shaping their language assessment literacy. Scarino (2013: 309) defined teachers' 'individual characteristics' to include their prejudices toward, beliefs about, and perceptions of the world that can impact on their judgments and decisions in assessment. Tsagari and Vogt (2017), for instance, found that those teachers who believe that assessment for learning can effectively support student progress are more likely to use learner-centred forms of assessment. In addition to teacher beliefs, teachers' choice of assessment procedures may be influenced by their past experience as well. When making assessment decisions, teachers may rely on their own successes and failures, and the assessment techniques they experienced as language learners can also play a decisive role. This was reported to be the case by Xu and Liu (2009) in their study that explored one Chinese College EFL teacher's LAL. The effect of teachers' past experiences as learner on their LAL was also discussed by Tsagari and Vogt (2017: 54), who phrased it as a principle of "test as you were tested". Vogt and Tsagari (2014) also relate teachers' past experiences as learners to their present assessment practices based on the type of feedback they give: if it focuses on only language errors and shortcomings, it can be identified as traditional, outcome-oriented assessment of learning. Vogt and Tsagari claim that if teachers do not have the appropriate conceptual framework for learner-centred assessments, or for some reason their language assessment literacy falls short, they may not be able to identify their actual needs

for their own professional development. Sometimes teachers' assessment training needs do not coincide with what they want.

According to Vogt et al. (2020), teachers' experiences in classroom assessment have the potential to gradually overwrite the assessment knowledge and recommendations for good practice learnt on teacher education programmes. The emerging new beliefs over time may have a crucial role in determining teachers' assessment decisions on a daily basis. However, it is also possible that the negative effects of lack of training can be mitigated through teaching experience and practical knowledge gained from colleagues (Sheehan & Munro, 2017).

### **Implications for teacher education**

The most recent, dynamic and situated view of assessment literacy clearly has implications for teacher education. Teachers' assessment literacy is likely to change over time, and their identity as assessors develops gradually as a result of complex and deep learning about assessment, which may take a long time to lead to changes in their conceptions and practices. Xu and Brown (2016) suggest that the basic knowledge of assessment teachers acquire during initial or further training goes through different filters: first the individual's cognitive and emotional filter, then the micro and macro environment filter (e.g. class, school community, education system), and finally the teaching experience filter. Ideally, teachers get a solid assessment literacy knowledge base in a separate course during their initial training, where the emphasis is put on enhancing teacher candidates' self-awareness and self-reflection. Xu and Brown consider the latter critical as teachers who are willing to follow only ready-made patterns, traditions, and never question their own classroom practice are not likely to embrace the current state-of-the-art approach, including assessment for learning.

For the above reasons, teachers and teacher candidates are likely to have different needs for training in assessment literacy. This can be explained by their beliefs, experience and the contextual factors that shape their perceptions and practice. Teacher education (pre- and in-service training) therefore should be contextually adaptive to take all these into account (Vogt et al., 2020). Research carried out in a variety of educational contexts has already explored teachers' and teacher candidates' perceptions of their (language) assessment literacy as well as their training needs. In the following, we will present some important findings to identify common patterns as well as highlight some context-specific needs. Our review is presented in two parts: teachers' and teacher candidates' (language) assessment literacy is discussed separately.

#### *Teachers' (language) assessment literacy*

Teachers' assessment literacy has been examined in a number of studies, some of which have focused specifically on language teachers. A detailed review of the relevant literature is offered by DeLuca et al. (2016), Xu and Brown (2016), Lee and Butler (2020) Coombe et al. (2020), and Lewkowicz and Leung (2021). In Europe, the main line of investigations has been largely determined by a non-representative survey conducted by Hasselgreen et al. (2004), which involved language teachers, teacher trainers and language testing professionals from several countries (mainly Finland, Sweden and the United Kingdom). The main finding of this study was that respondents needed in-service training in all aspects of language assessment, especially with regard to alternative assessment methods. The questionnaire Hasselgreen et al. used inspired several others who tried to explore the training needs of language teachers in specific European countries. Among these studies, the study by Vogt and Tsagari (2014) stands out as it involved 853 secondary school teachers from seven countries (Cyprus, Germany, Greece, Italy, Poland, Macedonia, Turkey). Overall, the results showed that respondents were dissatisfied with their language assessment skills, and the lowest level of confidence was shown by German teachers. The authors' conclusion was similar to that of Hasselgreen et al. (2004) that training needs in alternative, learner-centred forms of assessment were high, especially as regards self- and peer assessment and the use of student portfolios. Similar findings were reported by Kvasova and Kavytska (2014), based on their survey among university foreign language teachers (35 respondents) in Ukraine.

The results of the survey conducted in Poland by Lewkowicz and Zawadowska-Kittel (2010) reflect the trends presented so far. Their findings revealed that English teachers in primary and secondary schools (91 respondents) were poorly skilled in assessment, and they used predominantly summative assessment.

Translation was a highly favoured assessment task, and the higher the school level, the more teachers used this task. Alternative assessment procedures were not popular with teachers, and teachers mostly viewed assessment as ‘something that is done to students’. Lewkowicz and Zawadowska-Kittel concluded that for many teachers assessment seemed to be the black sheep of education.

A similar pattern of teachers’ beliefs and classroom practices to the Polish one was found in the Czech Republic by Seden and Svaricek (2018). Their case study included ten English teachers, whose assessment practice could also be characterised by a predominant use of summative assessment and little use of self- and peer assessment, or student portfolios. According to the teachers interviewed, these alternative forms of assessment do not work well in practice, they do not provide reliable assessment. Teachers believed that students were unable to notice their own mistakes or judge how well or poorly a task was performed.

The findings from surveys conducted in non-European countries are in line with the European trends. Classroom-based assessment is dominated by traditional testing techniques, and in the absence of sufficient professional insight, teachers are unable to take a critical approach to their own testing procedures. Hidri’s (2015) study involved Tunisian high school teachers (336 people) and university lecturers (206 people), and the results showed that the classroom evaluation of the respondents was predominantly norm-oriented and summative. Muhammad and Bardakçi (2019) explored the assessment literacy of Iraqi English teachers (101 people), but his finding reveal a similar pattern. Furthermore, respondents lacked awareness of test misuse, for instance, when a test was used illegally or unethically, or when it measured language proficiency in a professionally unacceptable way. Sultana’s (2019) qualitative research conducted in Bangladesh has shown that, due to the dominance of external tests, teachers (10 respondents) do not have time to use alternative assessment methods, as teaching is heavily focussed on test coaching.

#### *Teacher candidates’ (language) assessment literacy*

The studies targeted at teacher candidates’ assessment literacy have brought to light that the training they get mostly focuses on classroom procedures linked to assessment of learning rather than assessment for learning (DeLuca & Klinger, 2010; Jin, 2010; Lam, 2015; Kovácsné Dúró, 2004; Volante & Fazio, 2007). In Kovácsné Dúró’s study involving 163 Hungarian teacher candidates and in DeLuca and Klinger’s analysis of Canadian teacher trainees, a similarly unfavourable phenomenon can be detected: teacher trainees’ assessment literacy seems to be shaped more by the knowledge and experiences gained during their teaching practice than on assessment-related courses provided throughout their university studies. Teacher candidates’ perceptions of classroom assessment is likely to be affected by their mentor teachers, whose assessment literacy – if it is limited – can potentially curb the assessment-related behaviour of teacher candidates. Kovácsné Dúró even points out that teacher trainees do not really have the opportunity to try out or to observe ‘in operation’ various assessment methods and procedures during their university training, or in actual classroom settings. Thus, the theory–practice gap remains a problem. In this situation, the past experiences of teacher trainees can play a decisive role in the development of their assessment literacy. For example, this can account for the fact that they favour well-known, mostly traditional assessment methods. According to Kovácsné Dúró, the effectiveness of such assessment training is questionable. On the other hand, providing no training on classroom-based assessment for future teachers is also a cause for concern. Unfortunately, courses on pedagogical assessment are not mandatory in many teacher education institutions (cf. DeLuca & Klinger, 2010; Lam, 2015). In sum, the overall picture of teacher education in terms of enhancing assessment literacy does not appear to be favourable in the contexts shown above. Assessment courses seem to have little effect on the development of the assessment literacy of teacher candidates, whose beliefs, experiences and classroom practice seem to reflect a summative assessment culture, and alternative, formative assessment methods play only a minor role.

According to the Hungarian National Core Curriculum (A Nemzeti alaptanterv 110/2012 (VI. 4.) Korm. rendelet, 2012), formative assessment should include the use of self- and peer assessment by language learners in public education. However, Hubai and Lázár (2019) claim that language teachers are not given guidance on how to integrate these forms of assessment into their everyday practice. The study by Kovácsné Dúró (2004), for example, brought to light that for Hungarian teacher candidates formative assessment is limited to continuous error correction during practice activities, and summative evaluation is interpreted only as administering and marking achievement tests. There seems to have been no investigation to find out whether this situation has changed since 2004. A recent study by Vigh and Kiss-Kovács (2019) explored the impact of a university training course on pedagogical assessment. The authors compared 59 teacher trainees' initial perceptions of classroom-based assessment, which were mostly summative in nature, to their views at the end of the course. The training was intended to help teacher candidates appreciate formative assessment that is differentiated and promotes learning. The results of the questionnaire survey showed that there was some change in students' perceptions by the end of the course, but the authors concluded that a single course on pedagogical assessment was unlikely to guarantee a long-term change in attitudes.

### Conclusion

In this paper, we have reviewed how the conceptualisation of (language) assessment literacy has evolved over the past few decades. According to the most recent views of language assessment literacy, the concept is interpreted in terms of a multi-level and multi-componential framework, in which there is a unique, dynamic interaction among the proposed variables for each individual. In the context of language teaching and learning, Davies (2008) identified three essential areas that form the core of assessment literacy: skills, knowledge and principles. The depth and width of coverage of these areas, however, seem to depend on the needs of test users. This perception has motivated Taylor (2009; 2013) to differentiate the concept of AL/LAL according to various stakeholder groups.

Research focusing on teachers' (language) assessment literacy has revealed that in many educational contexts in and beyond Europe, classroom-based assessment appears to be dominated by traditional testing techniques, the emphasis seems to be placed on summative assessment with little or no attention to alternative forms of assessment that require learners to get engaged actively in the assessment process. This lack of evidence for assessment for learning (Black & Wiliam, 1998) calls for action from teacher educators. The need for change is reinforced by the OECD project called "Future of Education and Skills 2030"<sup>2</sup>, which promotes the concept of student agency. According to this, learners exercise their sense of purpose and responsibility while learning, which includes classroom-based assessment as well. Training Hungarian teacher candidates how to face such a challenge is a tall order: the development of their assessment literacy has already been found problematic owing to the theory-practice gap, and the supposedly low levels of assessment literacy of mentor teachers may further weaken the outcomes of pre-service teacher education (cf. Kovácsné Dúró, 2004). Another area for concern is related to the limitations of university teacher training courses, their questionable potential to generate long-term impact on teacher candidates' beliefs, attitudes and skills in pedagogical assessment. In our view, teacher candidates of foreign languages appear to be challenged even further: the usefulness and effectiveness of their assessment training can be hampered by the fact that they also need subject-specific (language-specific) knowledge. This is due to the fact that assessment outcomes are often communicated in the target language to learners who have a developing level of second language proficiency. In order to tackle the current challenges in classroom-based assessment, we suggest that the first steps should be taken in language teacher education. In Hungary, this could involve re-examining pre- and in-service courses on language testing and assessment so as to match them with current views on language assessment literacy for teachers as well as recommendations for good practice.

**Conflicts of Interest:** The author declares no conflict of interest.

---

<sup>2</sup> <https://www.oecd.org/education/2030-project/teaching-and-learning/learning/>



## References

- A Nemzeti alaptanterv kiadásáról, bevezetéséről és alkalmazásáról 110/2012. (VI. 4.) Korm. rendelet. (2012). *Magyar Közlöny*, 66, 10635-10848.
- Alderson, J. C. (1999). *Testing is Too Important to be Left to Testers*. Plenary Address to the Third Annual Conference on Current Trends in English Language Testing, United Arab Emirates University, Al Ain and Zayed University, Dubai Campus. Retrieved from <http://www.lancaster.ac.uk/fass/projects/examreform/Media/Article02.pdf>
- American Federation of Teachers, National Council on Measurement in Education & National Education Association, (1990). Standards for Teacher Competence in Educational Assessment of Students. *Educational Measurement: Issues and Practice*, 9(4), 30–32.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment, *Phi Delta Kappan International*. Retrieved from [http://edci770.pbworks.com/w/file/etch/48124468/BlackWiliam\\_1998.pdf](http://edci770.pbworks.com/w/file/etch/48124468/BlackWiliam_1998.pdf)
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe and do. *Language Teaching*, 36(2), 81–109.
- Brown, J. B., & Bailey K. M. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25(3), 349–383.
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: what do we need to learn, unlearn, and relearn? *Language Testing in Asia*, 10(3).
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347.
- Davison, C., & Leung, C. (2009). Current Issues in English Language Teacher-Based Assessment. *TESOL Quarterly*, 43(3), 393–415.
- DeLuca, C., & Klinger, A. (2010). Assessment literacy development: identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419–438.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: a review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251–272.
- Earl, L., & Katz, S. (2005). *Rethinking classroom assessment with purpose in mind*. Winnipeg, Manitoba: Western Northern Canadian Protocol. Retrieved from <https://digitalcollection.gov.mb.ca/awweb/pdfopener?smd=1&did=12503&md=1>
- Engelsen, K. S., & Smith, K. (2014). Assessment Literacy. In C. Wyatt-Smith, V. Klenowski & P. Colbert (Eds.), *Designing Assessment for Quality Learning* (pp. 91–107). Dordrecht: Springer Science+Business Media.
- Fulcher, G. (2012). Assessment Literacy for the Language Classroom. *Language Assessment Quarterly*, 9(2), 113–132.
- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 413–428). Amsterdam, The Netherlands: Mouton De Gruyter.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European Survey of Language Testing and Assessment Needs. Part 1: General findings*. Gothenburg, Sweden: European Association for Language Testing and Assessment. Retrieved from <http://www.ealta.eu.org/documents/resources/survey-report-pt1.pdf>
- Hidri, S. (2015). Conceptions of assessment: investigating what assessment means to secondary and university teachers. *Arab Journal of Applied Linguistics*, 1(1), 19–43.
- Hill, K., & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395–420.
- Hubai K., & Lázár I. (2018). Assessment of Learning in the Hungarian Education System with a Special Focus on Language Teachers' Views and Practices. *WoPaLP*, 12, 83–93.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402.
- Inbar-Lourie, O. (2017). Language Assessment Literacy. In E. Shohamy, I. Or & S. May (Eds.), *Language Testing and Assessment. Encyclopedia of Language and Education*, 3rd ed. (pp. 1–19). Springer International Publishing.
- Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555–584.
- Klinger, D. A., McDivitt, P. R., Howard, B. B., Munoz, M. A., Rogers, W. T., & Wylie, E. C. (2015). *The classroom assessment standards for PreK-12 teachers*. Kindle Direct Press.

- Kovácsné Duró A. (2004). Tanár szakos egyetemi hallgatók értékelési tudása és kompetenciái. *Magyar Pedagógia*, 104(2), 203–224.
- 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.
- Kvasova, O., & Kavyska, T. (2014). The assessment competence of university foreign language teachers: A Ukrainian perspective. *CercleS*, 4(1), 159–177.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197.
- Lee, J., & Butler, Y. G. (2020). Reconceptualizing language assessment literacy: Where are the learners? *TESOL Quarterly*, 54(4), 1098–1111.
- Lewkowicz, J., & Leung, C. (2021). Classroom-based assessment. *Language Teaching*, 54(1), 47–57.
- Lewkowicz, J., & Zawadowska-Kittel, E. (2011). Classroom-based Assessment: the Polish Experience. In D. Tsagari & I. Csépes (Eds.), *Classroom-Based Language Assessment* (pp. 147–167). Frankfurt am Main: Peter Lang GmbH.
- Malone, M. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344.
- Mertler, C. A., & Campbell, C. S. (2005, April). *Measuring teachers' knowledge and application of classroom assessment concepts: development of the Assessment Literacy Inventory*. Paper presented at the annual meeting of the American Educational Research Association, Montréal, Quebec, Canada. Retrieved from [https://www.academia.edu/2684680/Measuring\\_teachers\\_knowledge\\_and\\_application\\_of\\_classroom\\_assessment\\_concepts\\_Development\\_of\\_the\\_Assessment\\_Literacy\\_Inventory\\_?auto=download](https://www.academia.edu/2684680/Measuring_teachers_knowledge_and_application_of_classroom_assessment_concepts_Development_of_the_Assessment_Literacy_Inventory_?auto=download)
- Muhammad, F. H. N., & Bardakçi, M. (2019). Iraqi EFL Teachers' Assessment Literacy: Perceptions and Practices. *Arab World English Journal*, 10(2), 431–442.
- Newfields, T. (2006). Teacher development and assessment literacy. In T. Newfields, I. Gledall, M. Kawate-Mierzejewska, Y. Ishida, M. Chapman & P. Ross (Eds.), *Authentic communication: Proceedings of the 5th Annual JALT Pan-sig Conference* (pp. 48–73). Shizuoka, Japan: Tokai University College of Marine Science. Retrieved from <https://hosted.jalt.org/pansig/2006/PDF/Newfields.pdf>
- O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, 30(3), 363–380.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402.
- Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment Competencies of Teachers: A National Survey. *Educational Measurement: Issues and Practice*, 12(4), 10–12. Poehner, M. E. (2009). Group Dynamic Assessment: Mediation for the L2 Classroom. *TESOL Quarterly*, 43(3), 471–491.
- Popham, W. J. (2004). All About Accountability/Why Assessment Illiteracy Is Professional Suicide. *Educational Leadership*, 62(1), 82–83.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18(4), 429–462.
- 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.
- Schafer, W. D. (1991). Essential Assessment Skills in Professional Education of Teachers. *Educational Measurement: Issues and Practice*, 10(1), 3–6.
- Schafer, W. D. (1993). Assessment literacy for teachers. *Theory into Practice*, 32(2), 118–126.
- Seden, K., & Svaricek, R. (2018). Teacher Subjectivity Regarding Assessment: Exploring English as a Foreign Language Teachers' Conceptions of Assessment Theories that Influence Student Learning. *CEPS Journal*, 8(3), 119–139.
- Sheehan, S., & Munro, S. (2017). Assessment: Attitudes, practices and needs. *ELT Research Papers*, 17.08. London, UK: British Council.
- Stiggins R. J. (2002). Assessment Crisis: The Absence of Assessment for Learning. *Phi Delta Kappan*, 83(10), 758–765.
- Stiggins, R. J. (1991). Relevant Classroom Assessment Training for Teachers. *Educational Measurement: Issues and Practice*, 10(1), 7–12.

- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238–245.
- Sultanana, N. (2019). Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1–14.
- Tao, N. (2014). *Development and validation of classroom assessment literacy scales: English as a Foreign Language (EFL) instructors in a Cambodian higher education setting*. Unpublished doctoral dissertation, Victoria University. Retrieved from <http://vuir.vu.edu.au/25850/>
- Taylor, L. (2009). Developing Assessment Literacy. *Annual Review of Applied Linguistics*, 29, 21–36.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412.
- 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.
- Vígh T., & Kiss-Kovács R. (2019). Hogyan értelmezik a tanárszakosok a tanulói értékelés fogalmát? *Modern Nyelvoktatás*, 25(3–4), 156–173.
- Vogt, K., & Tsagari, D. (2014). Assessment Literacy of Foreign Language Teachers: Findings of a European Study. *Language Assessment Quarterly*, 11(4), 374–402.
- Vogt, K., Tsagari, D., & Spanoudis, G. (2020). What Do Teachers Think They Want? A Comparative Study of In-Service Language Teachers' Beliefs on LAL Training Needs. *Language Assessment Quarterly*, 17(4), 386–409.
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30(3), 749–770.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40(2), 241–256.
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162.
- 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), 493–513.
- Yan, X., Zhang, C., & Fan, J. J. (2018). "Assessment knowledge is important, but": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158–168.

