

REVIEW

Open Access

Language assessment literacy: what do we need to learn, unlearn, and relearn?

Christine Coombe¹, Hossein Vafadar² and Hassan Mohebbi^{3*}

* Correspondence: hassan.mohebbi973@gmail.com
³European Knowledge Development Institute, Ankara, Turkey
Full list of author information is available at the end of the article

Abstract

Recently, we have witnessed a growing interest in developing teachers' language assessment literacy. The ever increasing demand for and use of assessment products and data by a more varied group of stakeholders than ever before, such as newcomers with limited assessment knowledge in the field, and the knowledge assessors need to possess (Stiggins, *Phi Delta Kappa* 72:534-539, 1991) directs an ongoing discussion on assessment literacy. The 1990 Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, & NEA, *Educational Measurement: Issues and Practice* 9:30-32, 1990) made a considerable contribution to this field of study. Following these Standards, a substantial number of for and against studies have been published on the knowledge base and skills for assessment literacy, assessment goals, the stakeholders, formative assessment and accountability contexts, and measures examining teacher assessment literacy levels. This paper elaborates on the nature of the language assessment literacy, its conceptual framework, the related studies on assessment literacy, and various components of teacher assessment literacy and their interrelationships. The discussions, which focus on what language teachers and testers need to learn, unlearn, and relearn, should develop a deep understanding of the work of teachers, teacher trainers, professional developers, stakeholders, teacher educators, and educational policymakers. Further, the outcome of the present paper can provide more venues for further research.

Keywords: Assessment literacy, Language testing, Assessment knowledge, Assessment education

Introduction

The traditional thought of literacy or illiteracy as the ability or inability respectively to read and write has now begun to take on a new functional aspect. This aspect is conceptualized within different domains as possessing knowledge, skills, and competence for specific purposes and in particular fields. An individual is expected to be able to understand the content related to a given area and be able to engage with it appropriately. As with this growing number of domains and rapid advances in this era, it is imperative to acquire multiple literacies to keep up with this contemporary trend, such as computer literacy, media literacy, academic literacy, and many others. Given this evident growth of new literacies, it should not come as no surprise that assessment literacy began to appear as an early contribution in the general education literature

(Inbar-Lourie, 2008; Popham, 2008; Stiggins, 1999, 2001; Taylor, 2009) and in language testing (Brindley, 2001; Davies, 2008) focusing on identifying the characteristics of testing knowledge and skills of teachers. The 1990 Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, & NEA, 1990) made a considerable contribution to this field. Following these Standards, a substantial amount of for and against research has been published on the knowledge base and skills for assessment literacy, assessment goals, the stakeholders, formative assessment and accountability contexts, and measures examining teacher assessment literacy levels. This paper inquires into the philosophy behind language assessment literacy, its theoretical and conceptual framework, the related studies on assessment literacy, and various components of teacher assessment literacy and their interrelationships.

Language assessment literacy: the road we have gone

Language assessment literacy is generally viewed as a repertoire of competences, knowledge of using assessment methods, and applying suitable tools in an appropriate time that enables an individual to understand, assess, construct language tests, and analyze test data (Inbar-Lourie, 2008; Pill & Harding, 2013; Stiggins, 1999). Davies (2008) suggested a “skills + knowledge” approach to assessment literacy. “Skills” describe the practical know-how in assessment and construction, and “knowledge” to the “relevant background in measurement and language description” (p. 328). As it is evident in the literature, there has been a shift in developing language assessment literacy from a more componential view (e.g., Brindley, 2001; Davies, 2008; Inbar-Lourie, 2008) to a developmental one. For example, Fulcher (2012) believed that language assessment literacy should fall into a classification of (a) practical knowledge, (b) theoretical and procedural knowledge, and (c) socio-historical understanding. Fulcher argued that practical knowledge is the base and more important than all other aspects of language assessment literacy. Focusing on mathematics and science literacy, Pill and Harding (2013) classified language assessment literacy from “illiteracy,” through “nominal literacy,” “functional literacy” and “procedural and conceptual literacy,” to an expert level of knowledge: “multidimensional language assessment literacy” (p.383).

In her review paper, Taylor (2013), having considered these notions, suggested that language assessment literacy requires specific levels of knowledge and thus proposed eight levels (1) knowledge of theory, (2) technical skills, (3) principles and concepts, (4) language pedagogy, (5) sociocultural values, (6) local practices, (7) personal beliefs/attitudes, and (8) scores and decision making. However, Taylor was cautious about calling this a model, but her suggestion offered a useful starting point and paved the way for further research on more conceptualization of language assessment literacy.

For example, Baker and Riches (2018), in response to Taylor’s call for more research, investigated language assessment literacy characterization of 120 Haitian language teachers. They suggested an alternative language assessment literacy aspect required for language teachers and assessors. They elaborated on how language assessment literacy is different for these two groups, but their knowledge could be considered complementary in accomplishing collaborative task. Yan, Zhang, and Fan (2018) investigated the factors, namely experiential and contextual, mediate language assessment literacy development for three secondary-level Chinese teachers. The semi-structured retrospective interviews revealed that teachers had a distinct language assessment literacy profile and

more robust training needs in assessment practice than in theories of assessment. However, the need to study language assessment literacy with larger groups shows a gap in this field of research.

More recently, Kremmel and Harding (2020), through empirical research, investigated language assessment literacy needs of various groups, 1086 persons. They analyzed the responses of language teachers, language testing developers, and language testing researchers to provide research support to their survey's application and findings.

Language assessment and language learning

Language learning is viewed as a transdisciplinary process within multilingual multicultural realities in this current globalization, and this process is largely affected by new genres as a result of technological innovations and affordances of the current era (Leung & Scarino, 2016; Shohamy & Or, 2017). As Kern and Liddicoat (2010) asserted, language learners are perceived as a "social speaker/actor," for he "acts and speaks in multiple communities (scholarly, social, virtual, etc.), and he experiences the intercultural through his affiliations with various communities that often straddle different languages and cultures" (p.22). In these multi-contextually bound environments and discipline-specific assessment literacies, research on assessment literacy in different subject domains seems to be more focused on the combination of disciplinary knowledge and assessment. In language teaching, language teachers need to combine and use their disciplinary pedagogical knowledge or teaching with assessment knowledge in current language-learning constructs (Inbar-Lourie, 2008; Zolfaghari & Ahmadi, 2016). As Farhady (2018) has stated, the current understanding of language learning and use should be matched with the related assessment theory and practice to meet the challenge raised by the realities of different languages and cultures.

Since language learning possesses multifaceted modes and constructs, this requires corresponding assessment practices. Through an examination of the language testing literature, six assessment themes which reflect current language-learning constructs have been identified:

- 1) Assessment to promote language learning: This type of assessment has led to approaches that improve learning in the language learning context such as Learning Oriented Assessment (LOA, Turner & Purpura, 2016) and Dynamic Assessment (Poehner, 2008).
- 2) Classroom assessment: It helps process-oriented learning via appropriate testing methods and by focusing on assessing task performance, which requires competency in the related language construct, the educational settings in which language learning and teaching takes place, and task design (Wigglesworth & Frost, 2017).
- 3) Integrated language assessment: It has led to the integration rather than separation of language skills. According to Lee (2015), multiple competencies are required to complete integrated tasks in multi-mediated contexts. For example, integrated reading-writing tasks require different competencies such as extracting the source texts for ideas, selecting ideas, and organizing ideas.
- 4) Content assessment: In this type of assessment, content and language are supplementary to one another. Any assessment of content requires language and

any assessment of an individual's ability to use language will involve content or topical knowledge. For example, in the Content Language Integrated Learning (CLIL) approach, the language of teaching used in conveying meaning is important in meaning-oriented language learning (Lopriore, 2018).

- 5) Multilingual assessment: It is reflective of translanguaging pedagogies where learners can use their whole language-learning repertoires and multilingual competence. This allows for the assessment of dynamic language use as a result of interactions occurred amongst speakers (Lopez, Turkan, & Guzman-Orth, 2017).
- 6) Multimodal assessment: This classification of assessment is proposed since texts in different languages are conceptualized in multifaceted modes, presented with several meanings, delivered on-screen, live or on paper, presented with various sensory modes, and presented through various channels and media (Chapelle & Voss, 2017).

Different areas of the related assessment research

Reviewing the related literature provides insights into assessment literacy and helps with our understanding of what has worked for developing assessment literacy by examining the links between previous studies. The 1990 Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, & NEA, 1990) has made a considerable contribution to the field. The Standards prescribed that teachers need to attain competence in selecting suitable instructional assessment methods, developing suitable instructional assessment methods, administering, scoring, and interpreting the results assessment methods, applying assessment results in decision-making, developing valid student grading procedures, sharing assessment results to various stakeholders, and identifying unethical and in some cases illegal assessment methods and uses of the related information obtained from assessment tasks and tests.

Following the Standards, a considerable amount of literature has been published on the knowledge base and skills for assessment literacy, assessment goals, and the stakeholders (Abell & Siegel, 2011; Inbar-Lourie, 2008; Taylor, 2013), formative assessment and accountability contexts (Brookhart, 2011; JCSEE, 2015; Stiggins, 2010), and assessment education (DeLuca, Klinger, Pyper, & Woods, 2015). Likewise, instruments were developed to examine teacher assessment literacy levels (Campbell & Collins, 2007; DeLuca, 2012; DeLuca, Klinger, et al., 2015; Fan, Wang, & Wang, 2011; Graham, 2005; Greenberg & Walsh, 2012; Hill, Ell, Grudnoff, & Limbrick, 2014; Koh, 2011; Lam, 2015; Leahy & Wiliam, 2012; Lukin, Bandalos, Eckhout, & Mickelson, 2004; Mertler, 2009; Sato, Wei, & Darling-Hammond, 2008; Schafer & Lizzitz, 1987; Schneider & Randel, 2010; Smith, 2011; Wise, Lukin, & Roos, 1991).

Assessment literacy measures: a complex world of measures

Developing assessment literacy measures is a major area of interest within the field of assessment literacy. Most of the related studies have involved quantitative measures. Eight instruments regarding assessment literacy or teacher competency in assessment for pre-service and in-service teachers published between 1993 and 2012 were identified: Assessment Literacy Inventory (Campbell, Murphy, & Holt, 2002); Assessment Practices Inventory (Zhang & Burry-stock, 1997); Assessment Self-Confidence Survey (Jarr, 2012); Assessment in Vocational Classroom Questionnaire (Kershaw IV, 1993),

Part II; Classroom Assessment Literacy Inventory (Mertler, 2003); Measurement Literacy Questionnaire (Daniel & King, 1998); the revised Assessment Literacy Inventory (Mertler & Campbell, 2005); and the Teacher Assessment Literacy Questionnaire (Plake, Impara, & Fager, 1993) (see Table 1). DeLuca, Klinger, et al. (2015) represented these assessment instruments based on their item characteristics, the instrument's guiding framework, and the instrument's psychometric properties.

Plake et al. (1993) studied the assessment proficiency of 555 teachers and 268 administrators across American states. The results underscored the significant gaps in teachers' pedagogical and technical knowledge of language assessment. The participants had basic problems in assessment results' interpretation and communication. O'Sullivan and Johnson (1993) employed Plake et al.'s (1993) questionnaire with 51 teachers during a measurement course offering performance-based tasks that were related to the standards (AFT, NCME, & NEA, 1990). The results indicated that Classroom Assessment Task responses supported a strong match between performance tasks and the Standards, which further validated the questionnaire. Similarly, Campbell et al. (2002) investigated a revised version of Plake et al.'s (1993) questionnaire with 220 undergraduate students enrolled in a pre-service measurement course. They concluded that teacher participants' competency differed across the seven standards, and respondents were found to have lacked critical aspects of competency upon entering the teaching profession. Mertler (2003), in his study with 67 pre-service and 197 practicing teachers, found results that paralleled Plake et al.'s (1993) and Campbell et al.'s studies. Similarly, Mertler and Campbell (2004), with the aim of restructuring items into scenario-based questions, found low critical assessment competencies across teachers.

Brown (2004) and his later co-authors (e.g., Brown & Harris, 2009; Brown & Hirschfeld, 2008; Brown, Hui, Flora, & Kennedy, 2011) employed the Teachers' Conceptions of Assessment (COA) questionnaire to specify New Zealand primary school teachers' and managers' priorities based on four purposes of assessment: (a) improvement of teaching and learning, (b) school accountability, (c) student accountability, and (d) treating assessment as irrelevant. On this instrument, teachers were asked if they agreed or disagreed with various assessment purposes related to these four conceptions. The results from these studies indicated that teachers' conceptions of assessment were different based on context and career stage and participants agreed with the improvement conceptions and the school accountability conception while rejecting the view that assessment was irrelevant. Subsequently, Brown and Remesal (2012) used COA. They examined teachers' conceptions based on three purposes of assessment: (a) assessment improves, (b) assessment is negative, and (c) assessment shows the quality of schools and students. They reported similar results.

Overall, most studies on teacher's assessment competency use instruments that aim to identify teachers' conceptions toward different assessment aims. Findings from these studies revealed that teachers' assessment competency was inconsistent with the recommended 1990 Standards (Galluzzo, 2005; Mertler, 2003, 2009; Zhang & Burry-Stock, 1997). As Brookhart (2011) argued, the 1990 Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, & NEA, 1990) was no longer useful in supporting assessment practices, or the assessment knowledge teachers require within the current classroom context.

Table 1 Assessment literacy instrument descriptions (Adopted from DeLuca, Klinger, et al.'s (2015) work

Instrument	Item characteristics	Guiding framework	Respondents	Psychometric properties
Assessment Literacy Inventory (ALI) (Campbell et al., 2002)	35 content-based items (5 items per standard)	1990 Standards	220 pre-service teachers	$\alpha = 0.74$; mean total score = 21
Assessment Practices Inventory (API) (Zhang & Burny-stock, 1997)	67 Likert-type items; 5-point scale (1 = not at all skilled to 5 = highly skilled)	1990 Standards	311 in-service teachers	$\alpha = 0.97$ for total instrument; α ranged from 0.79 to 0.93 for 7 subscales
Assessment Self-Confidence Survey (Jarr, 2012)	15 Likert-type items; 7-point scale (1 = not confident at all to 7 = very confident)	(Bandura's 2006) guidelines for constructing self-efficacy scales	201 in-service teachers	$\alpha = 0.90$; mean total score = 64.9 (out of 105 max. points), SD14.2
Assessment in Vocational Classroom Questionnaire, Part II (Kershaw IV, 1993)	26 Likert-type items; 5-point scale (1 = not competent to 5 = extremely competent)	1990 Standards	393 in-service teachers	$\alpha = 0.91$; mean total score = 97.0 (out of 130 max. points), SD = 12.9
Classroom Assessment Literacy Inventory (CALI) (Mertler, 2003)	35 content-based items (5 items per standard)	1990 Standards	197 in-service teachers	$\alpha = 0.57$; mean total score = 22.0, SD = 3.4; $\alpha = 0.74$; mean total score = 19.0, SD = 4.7
Measurement Literacy Questionnaire (Daniel & King, 1998)	30 true/false items	Assessment literature (e.g., Gullickson 1984; Kubiszyn and Borich 1996; Popham 1995)	67 pre-service teachers, 95 in-service teachers	$\alpha = 0.60$; mean total score = 18.2, SD=3.3
Revised Assessment Literacy Inventory (ALI) (Mertler & Campbell, 2005)	35 scenario-based items (5 scenarios; 5 items per standard)	1990 Standards	250 pre-service teachers	$\alpha = 0.74$; mean total score = 23.9, SD=4.6
Teacher Assessment Literacy Questionnaire (TALQ) (Plake et al., 1993)	35 content-based items (5 items per standard)	1990 Standards	555 in-service teachers	$\alpha = 0.54$; mean total score = 23.2, SD = 3.3

The assessment-based teaching practices over the past 20 years (Volante & Fazio, 2007), and the recently revised Classroom Assessment Standards (JCSEE, 2015) set grounds for developing a new instrument for measuring teacher assessment literacy that directs the current demands on teachers. Gotch and French (2014) examined a recent systematic review of 36 literacy measures. They found that these measures do not support psychometric aspects and that existing instruments lack “representativeness and relevance of content in light of transformations in the assessment landscape (e.g., accountability systems, conceptions of formative assessment)” (p. 17). Gotch and French (2014) called for further research and developing an efficient and reliable instrument to measure teacher’s assessment literacy reflecting contemporary demands. In response, DeLuca, LaPointe-McEwan, and Luhanga (2016) studied professional learning communities to support teachers’ assessment practices and data literacy to reflect contemporary practices for classroom assessment. They developed the Approaches to Classroom Assessment Inventory containing 15 assessment standards (1990–2016) from six countries, namely the USA, Canada, UK, Europe, Australia, and New Zealand. They identified eight themes indicating contemporary aspects of teacher assessment literacy (see Table 1).

The themes of assessment standards from 1990 to 1999 included Assessment Purposes, Assessment Processes, Communication of Assessment Results, and Assessment Fairness; from 2000 to 2009 involved Assessment Purposes, Assessment Processes, Communication of Assessment Results, and Assessment Fairness, and Assessment for Learning; 2010–present focused on Assessment for Learning. Assessment Purposes, Assessment Processes, Communication of Assessment Results, and Assessment for Learning have become a more dominant theme in modern assessment standards.

Assessment Purposes refers to selecting the appropriate form of assessment according to instructional purposes. Assessment Processes involves constructing, administering, and scoring assessment and interpreting assessment results to facilitate instructional decision-making. Communication of Assessment Results includes communicating assessment purposes, processes, and results to stakeholders. Assessment Fairness entails providing fair assessment conditions for all learners by considering student diversity and exceptional learners. Assessment for Learning explains the use of formative assessment during instruction to guide teacher practice and student learning.

Assessment training and other affective factors

Harding and Kremmel (2016) have claimed that language teachers, as the primary users of language assessment, need to be “conversant and competent in the principles and practice of language assessment” (p.415). Teacher assessment literacy involves teachers’ mastery of knowledge and skills in designing and developing assessment tasks, analyzing the relevant assessment data, and utilizing them (Fulcher, 2012). Scarino (2013), focusing on assessment literacy for language teachers, argued that two aspects should be considered in developing language teacher assessment literacy, including the identification of relevant domains that contain the knowledge base and the relationship among these domains. He explained that the knowledge base composes some intersecting domains such as knowledge of language assessment, which entails not only various assessment paradigms, theories, purposes, and practices related to elicitation, judgment, and validation in diverse contexts, but also learning theories and practices and evolving

theories of language and culture. Additionally, Scarino stated that assessment could not be separated from its relationship with the curriculum and processes of teaching and learning in schooling. Scarino (2013) further claimed that "... it is necessary to consider not only the knowledge base in its most contemporary representation but also the processes through which this literacy is developed" (p. 316).

Some researchers have emphasized assessment in training (Boyles, 2005), establishing a framework of core competencies of language assessment (Inbar-Lourie, 2008), developing language testing textbooks (Davies, 2008; Fulcher, 2012; Taylor, 2009), and developing online tutorial materials (Malone, 2013). In a study with 66 Hong Kong secondary school teachers, Lam (2019) examined knowledge, conceptions, and practices of classroom-based writing assessment. He found that most teachers had related assessment knowledge and positive notions about alternative writing assessments; some teachers had a partial understanding of the assessment of learning and assessment for learning, but not assessment as learning as they could only follow the procedures without internalizing them. In a study of EFL teachers in Colombia, Mendoza (2009) found that teachers frequently and inappropriately use summative rather than formative assessments; they used test scores not to facilitate the learning process; they lacked knowledge of different types of language assessments and what information each type provides; how to give more effective feedback to students; how to empower students to take charge of their learning; ethical issues related to test and assessment use and how results are used; the role of the language tester; and concepts such as validity, reliability, and fairness. The authors concluded that teachers lack adequate language assessment training.

In a skill-based study with 103 Iranian teachers, Nemati, Alavi, Mohebbi, and Masjedlou (2017) pointed to the inadequacy of teachers' assessment knowledge and training in writing skill. Crusan, Plakans, and Gebriel (2016) also surveyed 702 second language writing instructors from tertiary institutions and studied teachers' writing assessment literacy (knowledge, beliefs, practices). Teachers reported training in writing assessment through graduate courses, workshops, and conference presentations; however, nearly 26% of teachers in this survey had little or no training. The results also showed the relative effects of linguistic background and teaching experience on teachers' writing assessment knowledge, beliefs, and practices.

The majority of studies are related to assessment education for both pre- and in-service teachers, teachers' training needs, conceptions of assessment, and efficacy (Brown, 2008; DeLuca & Lam, 2014; Gunn & Gilmore, 2014; Hill et al., 2014; Levy-Vered & Alhija, 2015; Quilter & Gallini, 2000; Smith & Galvin, 2014; Smith, Hill, Cowie, & Gilmore, 2014).

With this training-supportive perspective, the first focus was on the quality of assessment courses (Greenberg & Walsh, 2012), course content (Brookhart, 1999; Popham, 2011; Schafer, 1991), assessment of course characteristic factors (e.g., instructors, content, students, and alignment with professional standards) (Brown & Bailey, 2008; DeLuca & Bellara, 2013; Jeong, 2013; Jin, 2010), and pedagogies that reflect knowledge about assessment (DeLuca, Chavez, Bellara, & Cao, 2013). (Bailey & Brown's, 1995, and replicated in 2008) study was a starting point in the investigation of language assessment courses. They examined the teachers' backgrounds, the topics they taught, and their students' perspectives toward those courses. Although their study offered useful

findings, the issue of non-language teachers who teach language assessment courses was a missing link in their research. Kleinsasser (2005) also explored language assessment courses from the teachers' attitude. Kleinsasser argued that the major problem with teaching a language assessment course the failure of bridging between theory and practice. For example, the connection between the class discussions and the final assessment product is not well constructed. Qian (2014) found that English teachers did not have marking skills when assessing learners speaking in a school-based assessment in Hong Kong. DeLuca and Klinger (2010) found that Canadian teachers (288 candidates) knew how to conduct a summative assessment, but they were not familiar with formative assessment. They stressed the importance of direct instruction in developing teacher assessment literacy.

The usefulness of assessment education in both pre- and in-service programs is the other focus. The related studies stressed that assessment education should take different forms and integrate different stakeholders' views (DeLuca, 2012; Hill et al., 2014; Mertler, 2009), assessment literacy should be part of teacher certification and qualification (Sato et al., 2008; Schafer & Lizzitz, 1987), mentors should attend to student teachers' prior beliefs about assessment (Graham, 2005), and the instruction of content should be localized, subject-area specific that allow for teachers' free choice (Lam, 2015; Leahy & Wiliam, 2012). In a European study, Vogt and Tsagari (2014) reported that most teacher respondents lacked adequate assessment training, and they had only on-the-job experiences. For those who are unable to attend formal instruction may learn from on-line learning resources (Fan et al., 2011), workplace (Lukin et al., 2004), and daily classroom practices (Smith, 2011), instructional rounds (DeLuca, Klinger, et al., 2015), and design of assessment tasks and rubrics (Koh, 2011). For example, in a recent study, Koh, Burke, Luke, Gong, and Tan (2018) investigated the development of the task design aspect of assessment literacy in 12 Chinese language teachers. They found that although teachers quickly perceived many aspects of task design, they found it difficult to incorporate specific knowledge manipulation criteria into their assessments.

Although there has been much attention devoted to assessment-related training, most teachers are not well-equipped to perform classroom-based assessment confidently and professionally (DeLuca & Johnson, 2017). Thus, a large and growing body of literature has investigated how to improve teacher assessment knowledge via course work, professional development events, on-the-job training and self-study (Harding & Kremmel, 2016), assessment textbooks (Brown & Bailey, 2008), university-based coursework (DeLuca, Chavez, & Cao, 2013), and curriculum-related assessment (Brindley, 2001). Despite a large body of research on training, the assessment knowledge remains a challenge as teachers believe that it is theoretical and pedagogically irrelevant to everyday classroom assessment practices (Popham, 2009; Yan et al., 2018); the knowledge is not contextualized, and they usually learn about related assessment knowledge with a cookie-cutter approach (Leung, 2014); most training programs only include a generic assessment course which provides insufficient detail for developing an adequate assessment knowledge base.

The literature also shows that some research has been carried out on teachers' conceptions about assessment. The conception of assessment is believed to diagnose and improve learners' performance and the quality of teaching (Crooks, 1988), account for

quality instruction offered by schools and teachers (Hershberg, 2002), make students individually responsible for their learning through assessment (Guthrie, 2002), and show that teachers do not use assessment as a formal, organized process of evaluating student performance (Airasian, 1997). Cizek, Fitzgerald, and Rachor (1995) examined elementary school teachers and argued that many teachers have their assessment policies based on their conceptions of teaching. Kahn (2000) studied high school English classes and argued that teachers employed different assessment types because they eclectically held and practiced transmission-oriented and constructivist models of teaching and learning. And yet, conceptions may be individualistic because they are socially and culturally shared cognitive phenomena (van den Berg, 2002). In their study, Looney, Cumming, van Der Kleij, and Harris (2018) worked on a conceptualization of Teacher Assessment Identity. They argued that language teacher's professional identity, their beliefs about language assessment, their practice and performance in language assessment related tasks, and their cognition of their perceived role as language assessors play a vital role in evaluation of their effectiveness in the field of language assessment.

Some researchers also suggested that perceptions might be resistant to training (Brown, 2008), while others claimed a positive relationship between assessment training and teacher assessment literacy (Levy-Vered & Alhija, 2015; Quilter & Gallini, 2000). Interestingly, a study by Smith et al. (2014) investigated the assessment beliefs of pre-service teachers and found that the teachers' assessment beliefs were framed by their past experiences rather than by what they had been taught about assessment theories or policy requirements.

Some studies have also suggested that teachers' conceptions and assessment practices are dependent on specific contexts (Forsberg & Wermke, 2012; Frey & Fisher, 2009; Gu, 2014; Lomax, 1996; Wyatt-Smith, Klenowski, & Gunn, 2010; Xu & Liu, 2009). Willis, Adie, and Klenowski (2013) viewed assessment literacy as "a dynamic context-dependent social practice" (p. 242). According to this contextualized view, teacher assessment literacy is considered to be a joint property that needs input and support from many stakeholders, such as students, school administrators, and policymakers (Allal, 2013; Engelsen & Smith, 2014; Fleer, 2015).

Assessment literacy has also been investigated by considering different stakeholders in various educational contexts. For example, Jeong (2013) studied the difference between language assessment courses for language testers and non-language testers. The results revealed significant differences in the content of the courses based on the teachers' background in test specifications, test theory, basic statistics, classroom assessment, rubric development, and test accommodation. Additionally, the results indicated that non-language testers were less confident in teaching technical assessment skills than language testers and were willing to focus more on classroom assessment issues. Malone (2013) examined assessment literacy among language testing experts and language teachers through an online tutorial. The results reported from both language testing experts and language teachers revealed that testing experts stressed the need to develop an increasing knowledge of the theories, and teachers emphasized the need to increase their knowledge of the "how to" components in the tutorial. Pill and Harding (2013) explored policymakers' understanding of language testing. They observed that how a lack of understanding of both language and assessment issues and lack of familiarity with the tools used and with their intentions can result in meaningful

misconceptions which can undermine the quality of education. Evidently, such misconceptions may lead to misinformed and misguided decisions by the policy makers on crucial issues.

Web-Based Testing (WBT), another area of interest within the field of assessment, has been employed in different educational settings (He & Tymms, 2005; Sheader, Gouldsborough, & Grady, 2006; Wang, 2007; Wang, Wang, Wang, Huang, & Chen Sherry, 2004). WBT is used to administer tests, correct test papers, and record scores on-line. WBT can be presented in the form of online presentation, application- or software-product representation, hypermedia, audio and video representation, and so on. For example, Wang, Wang, and Huang (2008) investigated “Practicing, Reflecting and Revising with Web-based Assessment and Test Analysis system (P2R-WATA) Assessment Literacy Development Model” for improving pre-service teacher assessment literacy. They reported improvement in teachers’ assessment knowledge and assessment perspectives.

In a different focal area, O’Loughlin (2013) studied the IELTS (International English Language Testing System) using an online survey. Fifty staff completed the survey. The study examined how well the test goals were met and how they might be best addressed in the future. The results showed that the participants mainly considered the minimum test scores required for entry. They needed information about the ways IELTS scores can be interpreted and used validly, reliably, and responsibly in decision-making in higher education contexts. O’Loughlin suggested information sessions and online tutorials for learning about the IELTS test.

Conclusion

To sum up, the studies reviewed on assessment literacy clarify the fact that language assessment literacy is a multi-faceted concept and that defining it presents a major challenge. Clearly, it is related to educational measurement and influenced by current paradigms in this field. It is not answered what is the relative or general balance between what agreed-upon issues, and themes represent knowledge in this field and make it distinct. This uncertainty reflects the lack of unanimity within the professional assessment community as to what shapes the assessment knowledge that will be passed on to future experts in the field.

The studies reviewed on assessment literacy also indicate that teachers need assessment knowledge. Assessment courses programs should be part of teachers’ qualifications and requirements. Additionally, the content of the assessment knowledge base needs to be kept up with what is most recent, based on research and policy innovations. Teacher assessment training needs to become long and sustainable enough to engage teachers in profound learning about assessment, which will possibly help them improve and expand conceptions and practices about assessment. Further, assessment training needs to take the knowledge base and the context of practice into account and make connections between them. In other words, assessment literacy should be developed by considering various educational contexts and necessities of times and contexts. Assessment literacy also needs support from different stakeholders. Teachers as individuals and professionals need to be considered because teachers’ conceptions, emotions, needs, and prior experiences about assessment may help to improve the efficacy of training, assessment knowledge, and skills of teachers. Teacher assessment literacy

development does not only mean an increased assessment knowledge, but it also needs to expand and broaden contextual-related knowledge and inter-related competencies. In line with teacher professionalization in assessment, it requires a consideration of many inter-related factors such as teacher independence, identity as assessor, and critical perspectives. Teachers need to engage in learning networks where they can understand each other through a common language, communicate, and decide about their assessment practices.

Last but not the least, this review about assessment literacy challenges provides researchers with both general predictions and needs for further related investigations in developing assessment literacy and workable solutions to cope with such challenges. Besides, the present information may help teachers, policymakers, stakeholders, and researchers know where they are, where they need to be, and how best to proceed with their developmental work and research.

Taken together, there are still many unanswered questions about assessment literacy, and further studies are required give us a more comprehensive insight into language assessment literacy and enrich this ongoing discussion. More research can also be used to verify and validate, as well as the question, the current issues in assessment literacy. Further investigations are needed to guide policymakers in conducting standards that address both the contemporary development of assessment research and the cultural aspects of assessment. Also, more research can identify specific problems in pre- or in-service assessment education in specific contexts and provide supplementary methods to achieve better implementation of professional standards or policies. Further work can enrich teachers' assessment knowledge with insights from the latest assessment research findings since the assessment knowledge base is dynamic. Also, due to the importance of teacher conceptions in shaping teacher assessment literacy, further studies can provide greater insight into their conceptions and practice of assessment. We do need to learn, unlearn, and relearn about language assessment literacy which is of primary importance in enhancing the quality of language education.

Abbreviations

LAL: Language assessment literacy

Acknowledgements

Not applicable

Authors' contributions

Christine Coombe and Hassan Mohebbi, as the Guest Editors of the Special Issue of Language Assessment Literacy for Language Testing in Asia gave the idea of the paper. Hossein Vafadar wrote the first draft, and then the paper was edited and revised by Christine Coombe and Hassan Mohebbi. The authors read and approved the final manuscript.

Authors' information

No interest.

Funding

Not applicable

Availability of data and materials

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

¹Higher Colleges of Technology, Dubai, United Arab Emirates. ²Universiti Sains Malaysia (USM), George Town, Malaysia.

³European Knowledge Development Institute, Ankara, Turkey.

Received: 15 April 2020 Accepted: 23 April 2020

Published online: 26 May 2020

References

- Abell, S. K., & Siegel, M. A. (2011). Assessment literacy: What science teachers need to know and be able to do. In D. Corrigan, J. Dillon, & R. Gunstone (Eds.), *The professional knowledge base of science teaching* (pp. 205–221). Dordrecht: Springer.
- Airasian, P. W. (1997). *Classroom assessment*. New York: McGraw-Hill.
- Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy and Practice*, 20(1), 20–34.
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association (AFT, NCME, & NEA). (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, 9(4), 30–32.
- Bailey, K. M., & Brown, J. D. (1995). Language testing courses: What are they? In A. Cumming & R. Berwick (Eds.), *Validation in language testing* (pp. 236–256). Cleveon, UK: Multilingual Matters.
- Baker, B. A., & Riches, C. (2018). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557–581.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents, adolescence and education* (pp. 307–337). Greenwich: Information Age Publishing.
- Boyles, P. (2005). Assessment literacy. In M. Rosenbusch (Ed.), *National assessment summit papers* (pp. 11–15). Ames, IA: Iowa State University.
- Brindley, G. (2001). Outcomes-based assessment in practice: Some examples and emerging insights. *Language Testing*, 18(4), 393–407.
- Brookhart, S. M. (1999). Teaching about communicating assessment results and grading. *Educational Measurement: Issues and Practice*, 18(1), 5–14.
- S. M. What will teachers know about assessment, and how will that improve instruction. In R. W. Lizzitz & W. D. Schafer (Eds.), *Assessment in educational reform: Both means and ends* (pp. 2–17). Boston, MA: Allyn & Bacon.
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.
- Brown, G. T., & Harris, L. R. (2009). Unintended consequences of using tests to improve learning: How improvement-oriented resources heighten conceptions of assessment as school accountability. *Journal of MultiDisciplinary Evaluation*, 6(12), 68–91.
- Brown, G. T., & Hirschfeld, G. H. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy and Practice*, 15(1), 3–17.
- Brown, G. T., Hui, S. K., Flora, W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5-6), 307–320.
- Brown, G. T., & Remesal, A. (2012). Prospective teachers' conceptions of assessment: A cross-cultural comparison. *The Spanish Journal of Psychology*, 15(1), 75–89.
- Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy and Practice*, 11(3), 301–318.
- Brown, G. T. L. (2008). Assessment literacy training and teachers' conceptions of assessment. In C. Rubie-Davies & C. Rawlinson (Eds.), *Challenging thinking about teaching and learning* (pp. 285–302). New York: Nova Science.
- Brown, J. D., & Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25(3), 349–384.
- Campbell, C., & Collins, V. L. (2007). Identifying essential topics in general and special education introductory assessment textbooks. *Educational Measurement: Issues and Practice*, 26(1), 9–18.
- Campbell, C., Murphy, J. A., & Holt, J. K. (2002). Psychometric analysis of an assessment literacy instrument: Applicability to pre-service teachers. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, OH.
- Chapelle, C. A., & Voss, E. (2017). Utilizing technology in language assessment. In E. Shohamy & I. G. Or (Eds.), *Encyclopedia of language and education* (3rd ed., S. May (Ed.). Language testing and assessment (pp. 149–161). Cham, Switzerland: Springer Science + Business Media LLC/Pal.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159–179.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438–481.
- Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43–56.
- Daniel, L. G., & King, D. A. (1998). Knowledge and use of testing and measurement literacy of elementary and secondary teachers. *The Journal of Educational Research*, 91(6), 331–344.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347.
- DeLuca, C. (2012). Preparing teachers for the age of accountability: Toward a framework for assessment education. *Action in Teacher Education*, 34(5/6), 576–591.
- DeLuca, C., & Bellara, A. (2013). The current state of assessment education: Aligning policy, standards, and teacher education curriculum. *Journal of Teacher Education*, 64(4), 356–372.
- DeLuca, C., Chavez, T., Bellara, A., & Cao, C. (2013). Pedagogies for preservice assessment education: Supporting teacher candidates' assessment literacy development. *The Teacher Educator*, 48(2), 128–142.
- DeLuca, C., Chavez, T., & Cao, C. (2013). Establishing a foundation for valid teacher judgement on student learning: The role of pre-service assessment education. *Assessment in Education: Principles, Policy and Practice*, 20(1), 107–126.
- DeLuca, C., & Johnson, S. (2017). Developing assessment capable teachers in this age of accountability. *Assessment in Education: Principles, Policy and Practice*, 24(2), 121–126.
- DeLuca, C., Klinger, D., Pyper, J., & Woods, J. (2015). Instructional rounds as a professional learning model for systemic implementation of assessment for learning. *Assessment in Education: Principles, Policy and Practice*, 22(1), 122–139.

- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy and Practice*, 17(4), 419–438.
- DeLuca, C., & Lam, C. Y. (2014). Preparing teachers for assessment within diverse classrooms: An analysis of teacher candidates' conceptualizations. *Teacher Education Quarterly*, 41(3), 3–24.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2015). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*. Advance online publication. <https://doi.org/10.1007/s11092-015-9233-6>.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Approaches to classroom assessment inventory: A new instrument to support teacher assessment literacy. *Educational Assessment*, 21(4), 248–266.
- Engelsen, K. S., & Smith, K. (2014). Assessment literacy. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *The enabling power of assessment: Designing assessment for quality learning* (pp. 140–162). New York: Springer.
- Fan, Y. C., Wang, T. H., & Wang, K. H. (2011). A web-based model for developing assessment literacy of secondary in-service teachers. *Computers in Education*, 57(2), 1727–1740.
- Farhady, H. (2018). History of language testing and assessment. In J. I. Liontas, T. International Association, & M. DelliCarpini (Eds.), *The TESOL encyclopedia of English language teaching*. doi:<https://doi.org/10.1002/9781118784235.eelt0343>
- Fleer, M. (2015). Developing an assessment pedagogy: The tensions and struggles in re theorising assessment from a cultural-historical perspective. *Assessment in Education: Principles, Policy and Practice*, 22(2), 224–246.
- Forsberg, E., & Wermke, W. (2012). Knowledge sources and autonomy: German and Swedish teachers' continuing professional development of assessment knowledge. *Professional Development in Education*, 38(5), 741–758.
- Frey, N., & Fisher, D. (2009). Using common formative assessments as a source of professional development in an urban American elementary school. *Teaching and Teacher Education*, 25(5), 674–680.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132.
- Galluzzo, G. R. (2005). Performance assessment and renewing teacher education the possibilities of the NBPTS standards. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 78(4), 142–145.
- Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice*, 33(2), 14–18.
- Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through preservice teacher education. *Teaching and Teacher Education*, 21(6), 607–621.
- Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teach about K-12 assessment: A review*. Washington, DC: National Council on Teacher Quality.
- Gu, P. Y. (2014). The unbearable lightness of the curriculum: What drives the assessment practices of a teacher of English as a Foreign Language in a Chinese secondary school. *Assessment in Education: Principles, Policy and Practice*, 21(3), 286–305.
- Gullickson, A. R. (1984). Teacher perspectives of their instructional use of tests. *The Journal of Educational Research*, 77(4), 244–48.
- Gunn, A. C., & Gilmore, A. (2014). Early childhood initial teacher education students' learning about assessment. *Assessment Matters*, 7, 24–38.
- Guthrie, J. T. (2002). Preparing students for high-stakes test taking in reading. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 370–391). Newark, DE: International Reading Association.
- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 413–427). Germany: De Gruyter.
- He, Q., & Tymms, P. (2005). A computer-assisted test design and diagnosis system for use by classroom teachers. *Journal of Computer Assisted Learning*, 21(6), 419–429.
- Hershberg, T. (2002). Comment. In D. Ravitch (Ed.) *Brookings papers on education policy* (324–333). Washington DC: Brookings Institution Press.
- Hill, M. F., Ell, F., Grudnoff, L., & Limbrick, L. (2014). Practise what you preach: Initial teacher education students learning about assessment. *Assessment Matters*, 7, 90–112.
- Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402.
- Inbar-Lourie, O. (2013). Language assessment literacy. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 2923–2931). Oxford: Blackwell.
- Jarr, K. A. (2012). *Education practitioners' interpretation and use of assessment results*. University of Iowa Retrieved from <http://ir.uiowa.edu/etd/3317>.
- Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing*, 30(3), 345–362.
- 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.
- Joint Committee on Standards for Education Evaluation (JCSEE). (2015). *Classroom assessment standards for PreK-12 teachers*. [Kindle version]. Retrieved from <https://www.amazon.com/Classroom-Assessment-Standards-PreK-12-Teachersebook/dp/B00V6C9RVO>.
- Kahn, E. A. (2000). A case study of assessment in a grade 10 English course. *The Journal of Educational Research*, 93(5), 276–286.
- Kern, R., & Liddicoat, A. J. (2010). From the learner to the speaker/social actor. In G. Zarate, D. Lévy, & C. Kramsch (Eds.), *Handbook of multilingualism and multiculturalism* (pp. 17–23). Paris, France: Éditions des Archives Contemporaines, Ltd.
- Kershaw IV, I. (1993). Ohio vocational education teachers' perceived use of student assessment information in educational decision-making (Doctoral dissertation). Ohio State University. Retrieved from https://etd.ohiolink.edu/letd.send_file?accession=osu1244146232&disposition=inline
- Kleinsasser, R. C. (2005). Transforming a postgraduate level assessment course: A second language teacher educator's narrative. *Prospect*, 20, 77–102.
- Koh, K., Burke, L. E. C. A., Luke, A., Gong, W., & Tan, C. (2018). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, 22(3), 264–288.
- Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, 22(3), 255–276.

- 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.
- Kubiszyn, T., & Borich, G. (1996). *Educational testing and measurement: classroom application and practice* (5th ed.). New York: HarperCollins.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197.
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, 81, 78–89.
- Leahy, S., & Wiliam, D. (2012). From teachers to schools: Scaling up professional development for formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 49–72). London: SAGE.
- Lee, H. W. (2015). Innovative assessment tasks for academic English proficiency: An integrated listening-speaking task vs. a multimediamediated speaking task (Graduate Theses and Dissertations. 14584). Retrieved from <https://lib.dr.iastate.edu/etd/14584>
- Leung, C. (2014). Classroom-based assessment issues for language teacher education. In A. J. Kunnan (Ed.), *The companion to language assessment* (Vol. 3, pp. 1510–1519). Chichester, UK: Wiley Blackwell.
- Leung, C., & Scarino, A. (2016). Reconceptualizing the nature of goals and outcomes in language/s education. *The Modern Language Journal*, 100(S1), 81–95.
- Levy-Vered, A., & Alhija, F. N.-A. (2015). Modelling beginning teachers' assessment literacy: The contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation*, 21(5-6), 378–406.
- Lomax, R. G. (1996). On becoming assessment literate: An initial look at pre-service teachers' beliefs and practices. *The Teacher Educator*, 31(4), 292–303.
- Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy and Practice*, 25(5), 442–467.
- Lopez, A. A., Turkan, S., & Guzman-Orth, D. (2017). Conceptualizing the use of translanguaging in initial content assessments for newly arrived emergent bilingual students. *ETS Research Report Series*, 2017(1), 1–12.
- Lopriore, L. (2018). Reframing teaching knowledge in Content and Language Integrated Learning (CLIL): A European perspective. *Language Teaching Research*, 136216881877751. doi:<https://doi.org/10.1177/1362168818777518>
- Lukin, L. E., Bandalos, D. L., Eckhout, T. J., & Mickelson, K. (2004). Facilitating the development of assessment literacy. *Educational Measurement: Issues and Practice*, 23(2), 26–32.
- Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344.
- Mendoza, A. A. L., & Arandia, R. B. (2009). Language testing in Colombia: A call for more teacher education and teacher training in language assessment. *Profile Issues in Teachers Professional Development*, 11(2), 55–70.
- Mertler, C. A. (2003). Pre-service versus in-service teachers' assessment literacy: Does classroom experience make a difference? Paper presented at Annual meeting of the Mid-Western Educational Research Association, Columbus, OH.
- Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(2), 101–113.
- Mertler, C. A., & Campbell, C. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, 33, 49–64.
- Mertler, C. A., & Campbell, C. S. (2005). 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, Montreal, Quebec, Canada.
- Nemati, M., Alavi, S. M., Mohebbi, H., & Masjedlou, A. P. (2017). Teachers' writing proficiency and assessment ability: the missing link in teachers' written corrective feedback practice in an Iranian EFL context. *Language Testing in Asia*, 7(1), 21.
- O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, 30(3), 363–380.
- O'Sullivan, R. S., & Johnson, R. L. (1993, April). Using performance assessments to measure teachers' competence in classroom assessment. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.
- 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.
- Popham, W. J. (1995). *Classroom assessment: What teachers need to know*. Boston: Allyn-Bacon.
- Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development* (Vol. 9). Germany: Springer Science & Business Media.
- Popham, W. J. (2008). Formative assessment: Seven stepping-stones to success. *Principal Leadership*, 9(1), 16–20.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48(1), 4–11.
- Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, 46(4), 265–273.
- Qian, D. D. (2014). School-based English language assessment as a high-stakes examination component in Hong Kong: Insights of frontline assessors. *Assessment in Education: Principles, Policy and Practice*, 21(3), 251–270.
- Quilter, S. M., & Gallini, J. K. (2000). Teachers' assessment literacy and attitudes. *The Teacher Educator*, 36, 115–131.
- Sato, M., Wei, C. R., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal*, 45(3), 669–700.
- 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), 2–6.
- Schafer, W. D., & Lizzitz, R. W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education*, 38(3), 57–63.
- Schneider, M. C., & Randel, B. (2010). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 251–276). New York: Routledge.

- Shader, E., Gouldsborough, I., & Grady, R. (2006). Staff and student perceptions of computer-assisted assessment for physiology practical classes. *Advances in Physiology Education*, 30(4), 174–180.
- Shohamy, E., & Or, I. G. (2017). Editors' introduction. In E. Shohamy & I. G. Or (Eds.), *Encyclopedia of language and education* (3rd ed., S. May (Ed.)). *Language testing and assessment* (pp. ix–xviii). Cham, Switzerland: Springer Science + Business Media LLC.
- Smith, K. (2011). Professional development of teachers as a prerequisite for AFL to be successfully implemented in the classroom. *Studies in Educational Evaluation*, 37(1), 55–61.
- Smith, L. F., & Galvin, R. (2014). Toward assessment readiness: An embedded approach in primary initial teacher education. *Assessment Matters*, 7, 39–63.
- Smith, L. F., Hill, M. F., Cowie, B., & Gilmore, A. (2014). Preparing teachers to use the enabling power of assessment. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 418–445). Dordrecht: Springer.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappa*, 72, 534–539.
- Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23–27.
- Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5–15.
- Stiggins, R. J. (2010). Essential formative assessment competencies for teachers and school leaders. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 233–250). New York, NY: Taylor & Francis.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21–26.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412.
- Turner, C. E., & Purpura, J. E. (2016). Learning-oriented assessment in second and foreign language classrooms. In D. Tsagari & J. Baneerjee (Eds.), *Handbook of second language assessment* (pp. 255–272). Boston, MA: De Gruyter, Inc..
- van den Berg, B. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research*, 72, 577–625.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of an European study. *Language Assessment Quarterly*, 11(4), 374–402.
- 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.
- Wang, T. H. (2007). What strategies are effective for formative assessment in an e-learning environment? *Journal of Computer Assisted Learning*, 23, 171–186.
- Wang, T. H., Wang, K. H., & Huang, S. C. (2008). Designing a web-based assessment environment for improving pre-service teacher assessment literacy. *Computers in Education*, 51(1), 448–462.
- Wang, T. H., Wang, K. H., Wang, W. L., Huang, S. C., & Chen Sherry, Y. (2004). Web-based Assessment and Test Analyses (WATA) system: Development and evaluation. *Journal of Computer Assisted Learning*, 20(1), 59–71.
- Wigglesworth, G., & Frost, K. (2017). Task and performance-based assessment. In E. Shohamy & I. G. Or (Eds.), *Encyclopedia of language and education* (3rd ed., S. May (Ed.)). *Language testing and assessment* (pp. 121–134). Cham, Switzerland: Springer Science + Business Media LLC.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualizing teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40(2), 241–256.
- Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education*, 42(1), 37–42.
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy and Practice*, 17(1), 59–75.
- 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.
- Zhang, Z., & Burry-Stock, J. A. (1997). Assessment Practices Inventory: A multivariate analysis of teachers' perceived assessment competency. Paper presented at Annual meeting of the American Educational Research Association, Chicago, IL.
- Zolfaghari, F., & Ahmadi, A. (2016). Assessment literacy components across subject matters. *Cogent Education*, 3(1), 1252561.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
