ENGLISH TEACHERS' LEARNING NEEDS FOR LANGUAGE TESTING AND ASSESSMENT: AN EXPLORATORY STUDY

Dao Thi Bich Nguyen[*1], Nguyen Thi Hoai[1] and Nguyen Vinh Quang[2]

[1]*Faculty of English, Hanoi National University of Education, Hanoi city, Vietnam*

[2]*Quality Assurance Center, Hanoi National University of Education, Hanoi city, Vietnam*

Corresponding author Dao Thi Bich Nguyen, e-mail: nguyendb@hnue.edu.vn

Abstract. Language assessment has been a critical aspect in the field of English language education, and one's skills and knowledge of language assessment can be regarded as assessment literacy. This exploratory study employed surveys to measure and understand the learning needs for language testing and assessment of a diverse sample of English teachers from various educational contexts in northern Vietnam, as well as the relationship between their learning needs and personal factors including age, years of experience, and educational level. The results showed that English teachers wanted greater training on designing language tests, particularly those tailored to particular educational environments, and there is a correlation between their training needs and their educational background, experience teaching, and present teaching levels, and self-perceived language assessment literacy. The findings of this research contribute to the ongoing discourse on teacher assessment literacy, and it provides suggestions on the design of new materials for language testing and assessment education programs.

Keywords: language testing and assessment, assessment literacy, learning needs.

1. Introduction

Language assessment literacy (LAL) is a major component of language teachers' professional competence (Popham, 2009 [1]; Kremmel and Harding, 2020 [2]). In Vietnam, developing teachers' LAL has been the focus of the Ministry of Education and Training (MOET) and the priority of the National Foreign Languages Project (NFLP) (Duong, 2019) [3]. Specifically, the NFLP has set the goal of enhancing English teachers' assessment capabilities as one of the important tasks for the period 2020-2025 [4]. In 2012, the English Teacher Competency Framework (ETCF) for in-service EFL teacher education programs were proposed and developed by MOET, and attached to the NFLP [5]. The framework consists of five domains, and teachers' knowledge and skills in assessing students' progress and proficiency or LAL is one major descriptor of the domain of knowledge of language teaching. Moreover, developing teachers' competence in testing and assessment is one of the training modules organized by the NFLP every year. However, the content of the modules mostly focuses on designing test items for standardized tests while inadequate attention has been paid to EFL teachers' alternative assessment practices and literacy in Vietnam where traditional assessment has been upheld (Ngo, 2020 [6]; Tran, 2015 [7]).

While teachers who are literate in language assessment are able to create and carry out efficient testing procedures, effectively interpret test results and create appropriate lesson plans for students and make logical decisions about their students' learning, insufficient LAL on the side of educators could lead to poor language assessments, incorrect test score interpretations, and illogical instructional guidance, all of which could be harmful to students (Weng & Shen, 2022) [8].

Despite its critical role in language instruction and evaluation, teachers' LAL continued to be insufficient (Berry et al., 2017 [9]; Xu & Brown, 2017 [10]). Inadequate language evaluation materials and inadequate assessment literacy training for teachers are included in TESOL programs for aspiring teachers (Jeong, 2013 [11]; Popham, 2006 [12]), and in-service teachers receive little opportunity for language assessment training (Crusan et al., 2016) [13]. Furthermore, it was the fault of education officials and administrators at universities or schools to fail to guarantee that instructors receive adequate training before beginning their teaching careers (Coombe et al., 2012) [14]. Together, these elements impede the LAL development of teachers.

In Vietnam, there has been an increase in studies on testing and assessment. However, the majority of research has focused on standardized tests (Chi, 2022) [15]; or difficulties faced by language assessors when administering standardized tests (Chi, 2022) [15]. Other research has concentrated on enhancing teacher-trainees assessment competency before they entered into the teaching profession or enhancing the competency of in-service school teachers (Chi, 2022) [15].

This article describes the findings of a study designed to collect empirical data from language teachers to discover what learning needs they have in language testing and assessment. The intention was to use the outcome of the needs analysis to give empirically derived content to the concept of assessment literacy and inform the design and construction of new materials that can be used in language testing education programs. The study uses surveys to measure and understand the assessment literacy levels of a diverse sample of English teachers from various educational settings in northern Vietnam. Specifically, the study seeks answers to the following questions:

(1) What are English teachers' training needs for language testing and assessment?

(2) What is the relationship between English teachers' training needs for language and testing and assessment and their teaching experience, teaching levels, educational levels, and perceived LAL levels?

2. Content

2.1. Theoretical framework

2.1.1. Language Assessment Literacy

The term "language assessment literacy" (Stiggins, 1991 [16], 1997 [17]) has been accepted to refer to the range of skills and knowledge that stakeholders require to deal with the new world of assessment. However, there is little agreement on what assessment literacy comprises, despite the increasing diversity of approaches recommended to encourage its development (Fulcher, 2012) [18]. While Stiggins (2002) [19] and McMillan (2000) [20] define teachers' LAL as an understanding of the principles of sound assessment to appropriately integrate assessment with instruction and to utilize appropriate forms of teaching, Mertler views it as recognizing good assessment practices, understanding assessment methods, reporting assessment results, and integrating assessment and learning (2004) [21]. Similarly, Davies (2008) [22] and Taylor (2009) [23] posit that teachers' LAL includes knowledge, skills, and principles of conducting assessments appropriately. LAL is also used to refer to familiarity with testing practices, the use of assessment methods, the explanation and analysis of collected results, the decision-making, and use of assessment results for teaching (Boyle, 2005) [24]. Pill & Harding compare LAL with a bank of competencies that permit its owners, or teachers, to judge, create, and analyze tests (2013) [25].

The present study adopts Fulcher's (2012) [18] perspective on teachers' LAL, which posits that the components of teachers' LAL include knowledge, skills, and abilities necessary for the creation, development, upkeep, and evaluation of large-scale standardized and/or classroom-based assessments. Teachers' LAL also entails being conversant with test procedures and conscious of the ideas and precepts that support and direct practice, such as codes of ethics. Additionally, in order to assess the function and effects of testing on institutions, society, and individuals, as well as to understand why certain practices have developed the way they have, assessment literacy necessitates the ability to situate knowledge, skills, processes, principles, and concepts within larger historical, social, political, and philosophical frameworks.

2.1.2. Assessment Literacy Survey Instrument

Fulcher (2012) [18] developed and piloted a survey instrument to investigate the assessment literacy of language teachers. The survey was delivered over the Internet using Lime Survey Software and was widely advertised through professional organizations and discussion lists. The intended population was described as "language teachers," but the sample was essentially self-selecting. The survey comprised both closed- and constructed-response items and used a number of innovative design features that encouraged teachers to express needs independently of the suggested response options. The survey was piloted with a group of language teachers before being launched more widely. The survey results were analyzed using both quantitative and qualitative methods, and the outcomes were used to inform the design of new teaching materials and the further development of online resources that could be used to support program delivery.

2.2. Methodology

2.2.1. Research design

This study was an exploratory study which is an investigation into a study topic to find something novel and fascinating (Elman et al., 2020) [26]. Using this research design, it was expected to figure out English language teachers' training needs for language testing and assessment, thus providing interesting data for the material design of testing and assessment programs in Vietnam.

2.2.2. Participants of the study

Convenience sampling was employed in this study. 368 English teachers from three provinces in northern Vietnam, consisting of 48.4% secondary school teachers and 51.6% high school teachers, participated in the study. The proportion of female teachers was approximately 80%, four times as high as that of male teachers. There was also a diversity in the age range, years of experience, and education level among the participants, as shown in Table 1 below.

Table 1. The participants' age range, years of experience, and educational levels

Age range (years old)	Number of participants	Years of experience	Number of participants	Highest educational level	Number of participants
21-25	4 (1.1%)	Less than 5 years	16 (4.3%)	High school graduate	2 (0.5%)
26-30	14 (3.8%)	5-10 years	19 (5.2%)	Bachelor degree	291 (79.1%)
31-35	19 (5.2%)	10-20 years	184 (50%)	Master degree	66 (17.9%)
36-40	63 (17.1%)	More than 20 years	149 (40.5%)	Doctorate degree	0 (0%)
41-45	175 (47.6%)			Others	9 (2.4%)
46-50	79 (21.5%)				
51-55	14 (3.8%)				

2.2.2. Data collection and analysis

A survey instrument adapted from Fulcher (2012) [18] was employed to collect data for the study. There were two main parts to the questionnaire. In the first part, the participants answered 5 personal questions regarding their gender, age range, current educational level, working institution, and years of experience. The second part of the questionnaire consisted of 8 items, both multiple-choice and open ones, about their understanding and needs for language testing and assessment. The survey was delivered to the participants via Google Forms, and it took them about 15 minutes to complete the questionnaire.

After being collected, data were then analyzed with cross-tabulation by Statistical Product and Services Solutions (SPSS). This helped the researchers to investigate the relationship between the variables in the survey's items. Moreover, data from open items were analyzed using content analysis to figure out different themes in the participants' responses.

2.3. Research findings and discussion

The main purpose of this study is to support the development of language testing and assessment training programs for English teachers in secondary and high schools in Vietnam. It can also provide valuable insights for researchers and educational experts on how to improve LAL among foreign language teachers in secondary education. Additionally, it promotes attention to the important issue of providing effective language testing training and integrating it into foreign language teacher training programs. Therefore, with the collected data, we will analyze the content of the open-ended questions regarding the courses and books on testing and assessment they have undergone, as well as their needs for content-related training in language testing and assessment. For the closed-ended questions, exploratory analysis will be conducted to examine respondents' perspectives on 23 items in question 8 regarding the knowledge and skills needed in language assessment, as identified in Fulcher's (2012) [18] study, as well as their relationships with other items such as teaching experience, educational level, current teaching level, and self-perceived language testing and assessment competence of the participants.

2.3.1. The open questions

Questions 6 and 7 were designed to encourage teachers to compare their training experiences in language testing and assessment with their awareness of what they still need to be trained in order to use assessment effectively in their current teaching. When asked about the training contents (question 6), the majority of the participants said they had received training on constructing test specifications and designing test items. However, for question 7 which asked about their further training needs to effectively carry out language assessment tasks, approximately 40% of the participants expressed a desire for training in methods of designing tests that ensure value and reliability, scoring speaking and writing skills, and analyzing and using test data in teaching.

Question 9 asked about the books/ materials on language testing and assessment teachers have recently read and what they liked or disliked about the materials. The majority of the answers mainly mentioned reading materials of the training programs which provided them with theories and practical examples of language assessment suitable for the current secondary education curriculum and closely related to teachers' practical jobs. However, as commented by the participants, the materials' content was not deep enough, and the examples were not abundant. A few teachers mentioned reading such books as "Testing for Language Teachers" and "Practical Language Testing," which have in-depth content, various examples, and explicit instructions, but some aspects of the books did not align with the Vietnamese secondary education curriculum.

Questions 10, 11, and 12 addressed the topics in books/ materials on language testing and assessment more specifically, which were their perceived crucial topics and their desired ones. The opinions expressed generally emphasized the need for systematic content, covering concepts

such as assessment forms, assessment organizing techniques, matrix design techniques, test item design, language skills testing, and test result analysis and use. They particularly desired specific guidance and examples that are closely related to the current secondary education curriculum.

2.3.2. The close-ended questions

2.3.2.1. Teachers' language assessment training needs

Question 8 consisted of 23 items representing the knowledge and skills necessary for teachers to carry out language assessment practices. The survey data was subjected to exploratory factor analysis (EFA) using SPSS software to identify the main factors in the additional training needs for language testing and assessment among secondary and high school English language teachers in Vietnam. The EFA results, using the Maximum Likelihood method and Varimax rotation, with a cutoff point of 0.5 (considered practically significant according to Hair et al., 1998, p111) [27], revealed the emergence of two factors with eigenvalues greater than 1, explaining 75.6% of the observed variable variance as follows:

Table 2. Total variance explained

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	16.334	71.019	71.019	16.069	69.867	69.867	9.268	40.295	40.295
2	1.524	6.627	77.646	1.318	5.731	75.598	8.120	35.303	75.598
3	.848	3.685	81.331						
4	.498	2.165	83.497						
5	.412	1.791	85.288						
6	.389	1.693	86.980						
7	.351	1.527	88.508						
8	.301	1.308	89.816						
9	.253	1.101	90.917						
10	.253	1.099	92.015						
11	.222	.963	92.978						
12	.207	.900	93.878						
13	.184	.800	94.678						
14	.168	.731	95.409						
15	.156	.678	96.087						
16	.149	.650	96.736						
17	.138	.602	97.338						
18	.125	.544	97.882						
19	.116	.503	98.385						
20	.113	.489	98.874						
21	.092	.401	99.275						
22	.087	.378	99.653						
23	.080	.347	100.000						

Extraction Method: Maximum Likelihood.

Table 3. Rotated factor matrix

Rotated Factor Matrix[a]		
	Factor	
	1	*2*
c8.4	.867	
c8.3	.864	
c8.5	.861	
c8.9	.809	
c8.14	.797	
c8.18	.783	
c8.2	.771	

c8.6	.768	
c8.8	.693	.552
c8.13	.675	
c8.7	.670	
c8.15	.664	
c8.21		.778
c8.19		.778
c8.23		.770
c8.20		.747
c8.17	.518	.723
c8.12		.678
c8.10	.524	.655
c8.11	.518	.643
C8.1		
c8.22		
c8.16		
Extraction Method: Maximum Likelihood.		
Rotation Method: Varimax with Kaiser Normalization.		
a. Rotation converged in 3 iterations.		

The results showed that out of the proposed 23 items, it was possible to condense them into 3 groups, namely "History of Language Testing," "Large-scale testing," and "The uses of tests in society." The remaining items were categorized into 2 main factors based on their content, tentatively labeled as factor 1 "Skills in designing and developing language tests" and factor 2 "General knowledge of testing and assessment." Notably, items such as "Standard setting," "Reliability," "Validation," and "Test analysis" were mentioned in both factors, indicating that respondents considered these contents important and in need of both theoretical knowledge and practical implementation techniques.

2.3.2.2. The nexus between the training needs and teaching experience, educational level, current teaching level, and self-perceived LAL of teachers

To assess the nexus between the teachers' training needs and teaching experience, educational level, current teaching level, and self-perceived LAL of teachers, the cross-tabulation technique was employed, and the data was visualized with illustrative charts for a more intuitive understanding explanation.

* The nexus between training needs and teaching experience

From the data and charts, it can be observed that the demand for "Skills in designing and developing language tests" is higher than that for "General knowledge of testing and assessment." In particular, teachers with more teaching experience have a lower need for additional training in language assessment compared to newly qualified teachers. This initial observation may be attributed to the fact that experienced teachers have more practical expertise and have received more training in language testing and assessment. However, upon closer examination of the responses from experienced teachers (especially those with over 20 years of experience) regarding their educational level and self-perceived LAL, it is not the case (details will be discussed later). Teachers with over 20 years of experience often reported having a low educational level such as vocational school or even high school, which does not meet current standards, and self-perceived competence at a poor to average level.
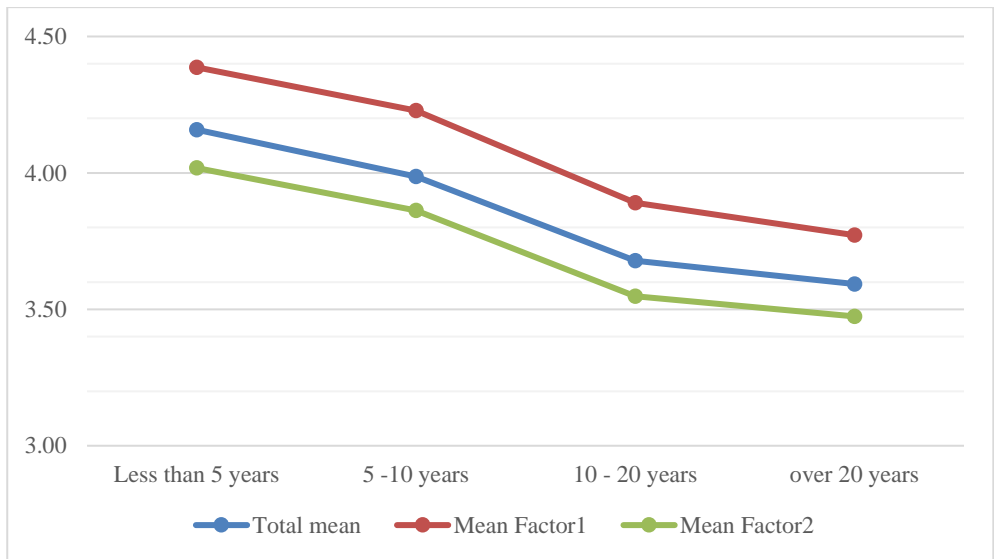
Figure 1. The nexus between training needs and teaching experience

The question arises as to with such a low initial training level and self-perceived LAL, why there is a low demand for training and professional development. When conducting additional interviews with some teachers in this category, most of them shared that it was due to age and nearing retirement, they were reluctant to change and learn challenging new knowledge. This information is quite interesting and useful for managers in organizing training and professional development programs in LAL.

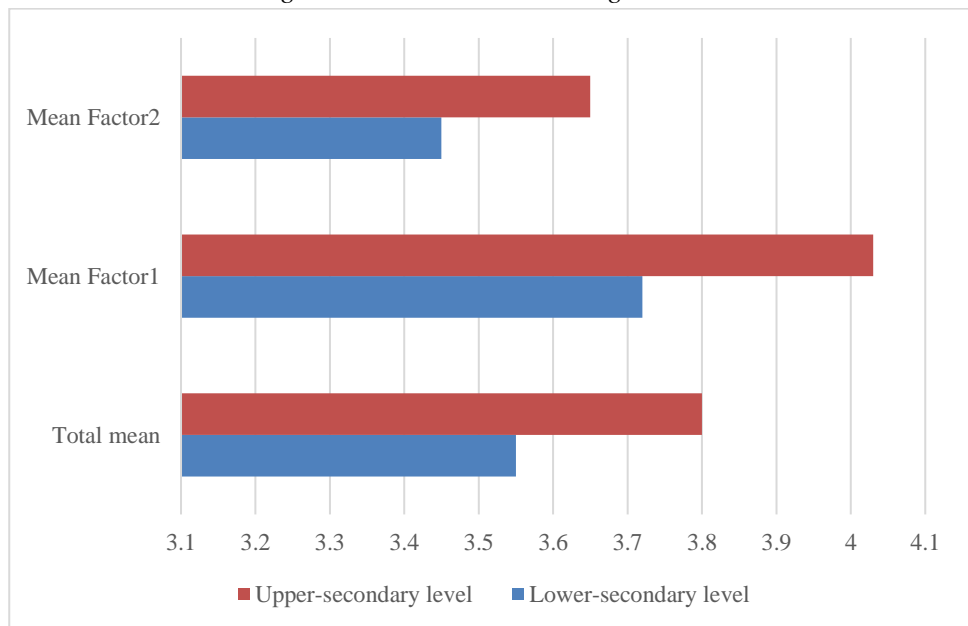** The nexus between training needs and current teaching levels*

Figure 2. The nexus between training needs and current teaching levels

It can be observed that high school teachers have a higher demand for training compared to lower-secondary school teachers. Specifically, the demand for "Skills in designing and developing language tests" is higher than that for "General knowledge of testing and assessment" in both groups.

The nexus between training needs and educational levels
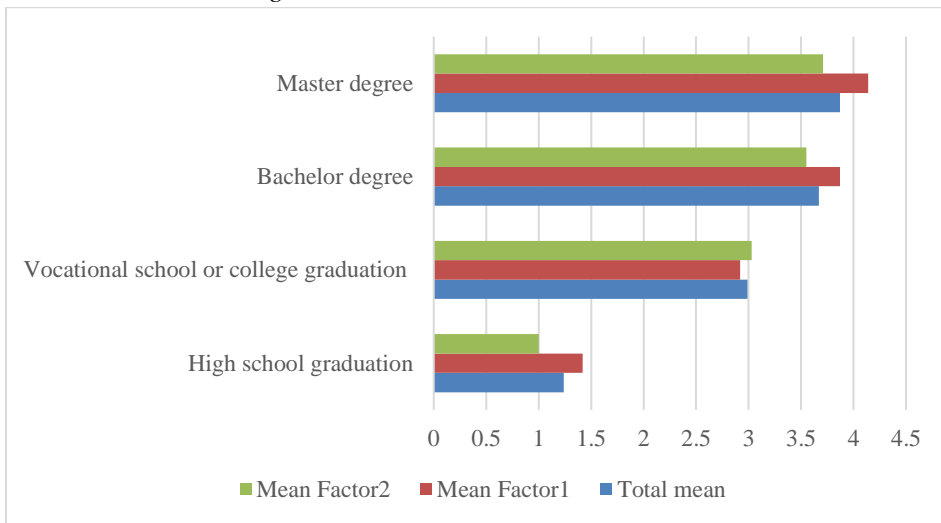

Figure 3. The nexus between training needs and educational levels

The data above indicates that teachers with lower educational levels such as high school or vocational school degrees have less demand for additional training in language assessment. This may seem illogical based on conventional reasoning, but further investigation reveals that this group largely consists of experienced teachers and the reasons for their lower demand have been explained in the previous section. On the other hand, teachers with vocational school or college education degrees have a higher demand for additional training in "General knowledge of testing and assessment." This could be because they may not have received comprehensive and systematic training in these areas compared to those with a Bachelor's or Master's degree.

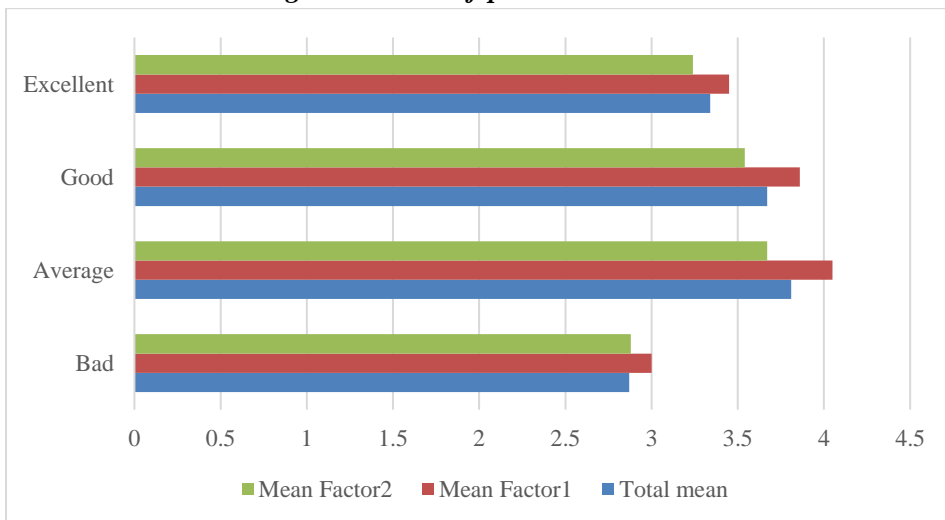*The nexus between training needs and self-perceived LAL levels*

Figure 4. The nexus between training needs and self-perceived LAL levels

Notable findings from the data include the low self-perceived LAL level combined with a low demand for additional training. This group consists of teachers with lower educational levels who have not met current standards and have over 20 years of experience. The reasons for this have been discussed in previous sections.

3. Conclusions

This exploratory study investigated secondary and high school English teachers' training needs for language testing and assessment in the context of Vietnam. 368 English teachers from three provinces in northern Vietnam completed a survey adapted from Fulcher (2012) [18]. Findings revealed that the teachers expected to be trained more on how to design language tests, especially the ones specific to distinctive teaching contexts. Moreover, there is certain relationship between the training needs and teaching experience, educational levels, current teaching levels, and self-perceived LAL of teachers.

Based on the data analysis, recommendations for organizing training and professional development programs in LAL, as well as for designing curriculum and materials, have been identified. First of all, regarding the selection of participants for training and professional development in LAL, consideration should be given to teachers with over 20 years of experience and lower educational levels. It is not advisable to organize training and workshops that combine language teachers from lower and upper secondary schools. Second of all, in terms of materials, it may be possible to exclude content related to "History of Language Testing," "Large-scale Testing," and "The Uses of Tests in Society" when providing training and professional development to lower and upper secondary school language teachers. There should also be an increase in the content related to the factor "Skills in Designing and Developing Language Tests" and provide detailed instructions and specific examples aligned with the current national curriculum for each grade level.

Nevertheless, this research only utilized one data-collecting instrument which was the survey. Further studies should collect data from more tools so that more in-depth data can be collected and analyzed.

REFERENCES

[1] Popham WJ, (2009). Assessment Literacy for Teachers: Faddish or Fundamental? *Theory Into Practice*, 48(1), 4–11. http://www.jstor.org/stable/40071570

[2] 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. Advanced online publication. https://doi.org/10.1080/15434303.2019.1674855

[3] Duong MT, (2019, November). *The high school exit examination in foreign languages in Vietnam: Current practices.* In Nguyen TMH, Ashton K, Duong TM, Sai CH, Nguyen TKA, High school graduation assessment practices for foreign languages: Trends and future directions. Paper presented at Vietnam Language Assessment Symposium (VLAS) 2019, Hanoi, Vietnam.

[4] MOET, (2008). *National Foreign Languages 2020 Project.* Hanoi. http://ngoainguquocgiaen.moet.gov.vn

[5] MOET, (2012). *Competency framework for English language teachers: User's guide.* Hanoi: Vietnam Education Publishing House

[6] Ngo TH, (2020). 'Standing between the flows': Interactions among neoliberalism, socialism and Confucianism in Vietnamese higher education. In Phan LH & Doan BN (Eds.), *Higher education in market-oriented socialist Vietnam* (pp. 41–62). Singapore: Springer.

[7] Tran TD, (2015). *An exploratory study of current assessment practices for improving the learning of English as a foreign language (EFL) in two Vietnamese universities* [Unpublished doctoral dissertation]. Queensland University of Technology.

[8] Weng, F., & Shen, B, (2022). Language Assessment Literacy of Teachers. *Front. Psychol*. 13:864582. DOI: 10.3389/fpsyg.2022.864582

[9] Berry V, Sheehan S, & Munro S, (2017). "Exploring teachers' language assessment literacy: a social constructivist approach to understanding effective practices." In ALTE (Ed.), *Learning and Assessment: Making the Connections - Proceedings of the ALTE 6th International Conference*. Association of Language Testers in Europe, 201–207.

[10] Xu Y, & Brown GTL, (2017). University English teacher assessment literacy: A survey-test report from China. *Pap. Lang. Test. Assess*. 6, 133–158.

[11] Jeong H, (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Lang. Test*. 30, 345–362. DOI: 10.1177/0265532213480334

[12] Popham WJ, (2006). Phony formative assessments: Buyer beware! *Educational Leadership*, 64(3), 86–87.

[13] Crusan D, Plakans L, & Gebril A, (2016). Writing assessment literacy: surveying second language teachers' knowledge, beliefs, and practices. *Assess. Writ*. 28, 43–56. DOI: 10.1016/j.asw.2016.03.001

[14] Coombe C, Troudi S, & Al-Hamly M, (2012). "Foreign and second language teacher assessment literacy: issues, challenges, and recommendations," in *The Cambridge Guide to Second Language Assessment*. eds. C. Coombe, P. Davidson, B. O'Sullivan and S. Stoynoff. Cambridge, UK: Cambridge University Press, 20–29.

[15] Chi NT, (2022). Challenges for college English teachers as assessors. *VNU Journal of Foreign Studies*, 38(5), 34-52. DOI: https://doi.org/10.25073/2525-2445/vnufs.4817.

[16] Stiggins RJ, (1991) 'Assessment literacy'. *Phi Delta Kappan*, 72 (7), 534–9.

[17] Stiggins, RJ, (1997). *Student-Centered Classroom Assessment* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

[18] Fulcher G, (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132.

[19] Stiggins RJ, (2002). Assessment Crisis: The Absence of Assessment for Learning. *Phi Delta Kappan*, 83(10), 758-765. https://doi.org/10.1177/003172170208301010

[20] McMillan JH, (2000). Fundamental Assessment Principles for Teachers and School Administrators. *Practical Assessment, Research & Evaluation*, 7, 78-93.

[21] Mertler CA & Campbell CS, (2004). *Assessing those who assess: Development of an instrument to measure teachers' assessment literacy*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, OH, October.

[22] Davies A, (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3): 327–347.

[23] Taylor L, (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-36.

[24] Boyles P, (2005). Assessment literacy. In Rosenbusch M. (Ed.). *National assessment summit papers* (pp. 11–15). Ames, IA: Iowa State University.

[25] Pill J, & Harding L, (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381-402. https://doi.org/10.1177/0265532213480337

[26] Elman C. Gerring J, & Mahoney J, (Eds.). (2020). *The production of knowledge : Enhancing progress in social science.* Cambridge University Press.

[27] Hair J, Anderson R, Tatham R & Black W, (1998). *Multivariate data analysis*. 5th Edition, Prentice Hall, New Jersey.