

RESEARCH

Open Access



Understanding university English instructors' assessment training needs in the Chinese context

Ling Gan^{1,2*}  and Ricky Lam¹

* Correspondence: Gloria-ganling@life.hkbu.edu.hk

¹Department of Education Studies, Hong Kong Baptist University, 15 Baptist University Road Campus, HK SAR, China

²School of Foreign Languages, Zhongnan University of Economics and Law, 182 Nanhu Avenue, Wuhan, China

Abstract

In the past decade, language assessment training has increasingly become a key research agenda in language testing and assessment, particularly within the Chinese context. Studies have explored the training of language instructors who provide language testing and assessment courses, the efficacy of assessment training courses, and assessment training needs of secondary school teachers. However, little attention has been paid to assessment training needs of university English instructors. To address the gap, we adopted a mixed-methods study to examine university English instructors' assessment training status, their training needs, and the factors that influence these needs. Based on the data from 68 questionnaire respondents, eight interviewees, and two national English course syllabi, we found that a majority of university English instructors lacked sufficient assessment training in classroom-based practices to meet the requirements set by the official documents. Despite this, they did not prefer having advanced training (e.g., workshops lasting at least three days) in assessment due to a variety of personal and contextual factors. The findings have implications for a need to revisit assessment training courses provided in teacher education programs and to review institutional policies that guide teacher professional development in China.

Keywords: Language assessment literacy, Assessment training needs, Personal factors, Contextual factors, University English instructors

Introduction

As the role of assessment in students' learning is broadly recognized, language assessment literacy (LAL) has become an increasingly crucial construct for assessment stakeholders, particularly for language teachers (Scarino, 2013; Taylor, 2013). As an entity from assessment literacy in general education, LAL generally refers to the knowledge, skills, principles, or the abilities stakeholders should acquire to conduct language assessment tasks (Fulcher, 2012; Inbar-Lourie, 2017). Despite the ongoing discussion on what specific LAL is needed for various stakeholders (Inbar-Lourie, 2017), it is widely agreed that teachers are those who need LAL most urgently (Harding & Kremmel, 2016). As the largest group of LAL consumers (Inbar-Lourie, 2017), language teachers are expected to acquire adequate LAL to conduct assessment activities that are aligned

with the intended learning outcomes. Insufficient LAL will be detrimental to assessment reliability and validity, hence to students' learning (Xu & Brown, 2017). Thus, the training to promote LAL appears to be pivotal to the development of language teachers' assessment expertise.

Until recently, much research has been conducted on various aspects of language assessment training, including language teachers' assessment training status quo (Fulcher, 2012), their training needs (Vogt & Tsagari, 2014), and online assessment training (Malone, 2008). Compared with the broad research in western educational contexts, assessment training is still underexplored in the Chinese setting (Jin, 2018), where the educational culture is prominently exam-oriented (Yan, Zhang, & Fan, 2018). Although recent LAL training studies in the Chinese context (Lam, 2019; Lan & Fan, 2019) have given insights into LAL from a local perspective, there is a dearth of research that addresses what assessment training language teachers have received, what training they wish to have, and what factors may exert an impact on their assessment training needs. This study attempts to address these thorny issues. We believe that helping university English instructors in China to understand their assessment training needs and the factors mediating these needs is of great importance, given that the instructors are currently taking enormous assessment responsibilities (e.g., decision-making) and handling intricate assessment problems (e.g., feedback, grading, rating, etc.) among millions of adult English learners (Xu & Brown, 2017).

Literature review

It is critically necessary for language teachers to acquire LAL since they are the direct test users (Scarino, 2013). However, most language teachers lacked sufficient knowledge and skills to conduct assessment-related tasks (Crusan, Plakans, & Gebriel, 2016; Vogt & Tsagari, 2014). The dissatisfying status quo of language teachers' LAL points to the urgent need for training in this respect (Fulcher, 2012; Malone, 2017), as language assessment training programs or assessment courses can both equip teachers with necessary LAL and promote their own teaching (Lee, 2019; Sultan, 2019).

Language teachers' inadequacy of assessment training has been well-documented in the field of language testing and assessment (Lam, 2015; Chung & Nam, 2018). Hasselgreen, Carlsenand Helness (2004) took the initiative to tap assessment training of language teachers in Europe. They observed that assessment training was absent among teachers who carried out language testing and assessment activities. Nonetheless, language teachers who received insufficient assessment training were found to have a perceived need for training (Guerin, 2010). The finding was also supported by the studies of Fulcher (2012) and Chung and Nam (2018). In their studies, most language teachers who lacked assessment training expressed strong training needs in practical areas such as test design and development. The lack of assessment training indicates a paucity of LAL among language teachers while their assessment needs affirm the effectiveness of assessment training in promoting teachers' LAL.

Teachers' assessment training needs have different priorities in different contexts. Built on the study by Hasselgreen, Carlsenand Helness (2004), the team of Vogt and Tsagari focused their research on assessment training needs of foreign language teachers across seven European countries (Tsagari & Vogt, 2017; Vogt & Tsagari, 2014). Their finding confirmed that language teachers in different countries had

different priorities in content areas of language testing and assessment in their training programs and that the extent of training desired by teachers depended on the educational contexts, specifically assessment culture in their countries. For example, Greek teachers needed more advanced training because both the Ministry of Education and the English curriculum in the country had a high demand on teachers in assessment practices. Differently, in Germany, language teachers desired moderate training needs, placing more focus on skill-based assessment such as listening or reading, because in this country, linguistic skills were intensively tested in high-stakes exams and highly required in national standards for foreign language teaching. Although these findings offer valuable insights into differing assessment training needs in Greece and Germany, such a simple comparison should be treated with caution as the external accountability demands from these countries differ significantly.

Assessment training needs are also individualized owing to the context-specific nature of LAL. That is, as teachers' LAL encapsulates their knowledge, skills, perceptions, and beliefs they bring to the teaching and assessment process (Scarino, 2013), their assessment training needs may vary individually. Yan, Zhang, and Fan (2018) discovered that EFL teachers at secondary schools in China had greater training needs for assessment practice than assessment theory. Similarly, the participants in Sheehan, Berry, and Munro's study (2016) reported showing no interest in language assessment theories. Moreover, these individualized assessment needs may vary across different career stages. For instance, pre-service teachers tended to long for assessment training courses with a focus on practical skills such as reporting results, developing test items, item reliability and validity and theories of assessment (DeLuca & Klinger, 2010). In a similar vein, novice teachers voiced fewer assessment training needs than their experienced colleagues in terms of language pedagogy, local educational policies, and technical skills (Xie & Tan, 2019).

Individualization of assessment training needs may be ascribed to a range of factors. Heavy workload discourages teachers from airing their training needs in assessment theories and principles because learning these theories and principles are difficult and time-consuming (Yan, Zhang, & Fan, 2018). Imbalanced training contents that teachers have received also lead to more varied individual desires for further training. As Lam (2015) pointed out, the inefficiency of assessment training for pre-service teachers in teacher education institutions may attribute to the fact that language assessment courses cannot provide sufficient assessment skills and theories for teachers. Besides, training time and resources may cause individualized assessment training needs. Those participants in Boyd and Donnarumma's study (2018) articulated insufficient time for them to reflect on their assessment activities and limited training resources of not helping them to have a fuller understanding of testing principles and practices. These constraints may push teachers to prioritize information they can understand during the training. The research endeavors that have been made so far demonstrate that factors influencing assessment training needs are highly contextualized. Different contexts may determine what factors and to what extent these factors constrain or facilitate assessment training.

The extant studies have yielded valuable insights into assessment training for language teachers. However, most of these studies centered on primary or secondary school language teachers and few were on EFL teachers at the tertiary level. Besides,

compared with the growing studies in western educational contexts, language assessment training research in the Chinese context is relatively scarce. Although assessment training has become an important issue in the Chinese academic community, most studies have focused on professional training of language instructors who provide language testing and assessment courses (Jin, 2010), assessment training landscape in Hong Kong (Lam, 2015), and assessment training needs of secondary school teachers (Yan, Zhang, & Fan, 2018). Yet, little research has been done targeting assessment training needs of university English instructors. To fill this gap, the paper aims to explore what assessment training university English instructors have received, what assessment training they aspire to have, and what factors affect their assessment training needs. The following research questions are addressed accordingly:

1. What assessment training have university English instructors received?
2. What assessment training needs have university English instructors expected?
3. What factors may influence university English instructors' assessment training needs?

Research method

Context

Before reporting the present research, it is important to unveil the challenges university English instructors face. In China, English learning is a mandatory component of education at all levels. The demand for English teachers is huge, as every year more than 225 million students receive formal education at the primary, secondary, and tertiary undergraduate and graduate levels (Qian & Cumming, 2017). As a result, university English instructors are recruited regardless of various educational backgrounds, with or without language testing and assessment learning or training. They will receive an orientation training of education and educational psychology in a short term (about 2 weeks) before the commencement of their teaching. However, testing and assessment are excluded from this two-week pre-service training program. Meanwhile, although there are no specialized assessment policies for higher education in China (Xu & Brown, 2017), the national documents about English language teaching such as *A Guide to College English Teaching (GCET)* (Chinese Ministry of Education, 2017) and *Teaching Guide for Undergraduate English Major (TGUEM)* (Chinese Ministry of Education, 2020) have elucidated assessment requirements. These documents explicitly require that instructors should conduct a variety of testing and assessment practices to evaluate students' language proficiency, explicitly stressing the parallel importance of formative and summative assessment in teaching. The documents attempt to promote the role of teachers in decision-making with respect to both formative and summative assessment processes. Other than being a decision-maker for classroom assessment, university English instructors also play the role of being a consultant for an increasing number of students who take the national College English Test Band 4 (CET4) and Band 6 (CET 6) as well as international standardized tests such as IELTS, TOEFL, and GRE. All these indicate that university English instructors should be empowered with adequate LAL to enable themselves to handle assessment-related activities in and off classes.

Design

The study adopted a mixed-methods design to provide an in-depth and nuanced understanding of university English instructors' assessment training situation and their

perceived training needs in the Chinese context. In this exploratory study, quantitative and qualitative data were used for the purposes of triangulation and complementarity. That is, the follow-up qualitative data explained, elaborated, and clarified the quantitative results. Quantitative data were collected from online questionnaires and qualitative data from semi-structured interviews and documents regarding assessment policies. The questionnaire was to investigate English instructors' current state of assessment training and training needs. Following up was the semi-structured interview and document analysis to dig further into teachers' training needs and address the issue of the influential factors on their needs. In this way, we attempted to paint a picture of assessment training landscape that is necessary for understanding the complex nature of Chinese university English instructors' LAL development.

Participants

The key participants are frontline instructors teaching university students English as a foreign language. We adopted a representative sampling in this study, taking into account the participants' various characteristics. Among 68 participants, 13 were male and 55 were female. Fifty-nine of the 68 participants teach non-English major students and nine teach English majors. The majority (75%) got a master's degree in different sub-disciplines, such as Linguistics, English Literature, Language Teaching or Translation, while nine of them (13%) had a doctoral degree in English Literature and Linguistics. In terms of teaching experience, 53% of respondents had 11–20 years and 19% over 20 years in English teaching respectively. Eight respondents come from Tier 1 universities, 40 from Tier 2, and 20 from Tier 3¹. After completing the questionnaires, eight participants (see Table 1) agreed to receive the follow-up interviews by leaving their email address or telephone number at the end of the questionnaire. Before data collection, all participants signed the informed consent form.

Data collection

We adapted the questionnaire used by Vogt and Tsagari (2014) for the current study. This questionnaire was used because it had been attested in assessment training need research (Lan & Fan, 2019). The questionnaire comprised three parts. The first part was about respondents' background information including gender, academic degree, teaching experience, student types, the institution they work in, and research areas. The second part was to investigate assessment training respondents had received and wished to have. This main part contained three thematic areas: (1) classroom-based assessment practices, (2) purposes of testing and assessment, and (3) content and concepts of language testing and assessment. In this study, the sequence of area (1) and area (3) was reversed because we hold that, as the knowledge base for assessment practices, content and concepts should come first in the sequence. Then, each thematic area was divided into two sections: received training and needed training. A three-point Likert scale was adopted for each item with "not at all", "a little", and "advanced" options for received training and "none", "basic", and "advanced" for needed training. The ambiguous terms like "a little", "basic", and "advanced" in Vogt and Tsagari's

¹Note: There are three tiers of universities in China. Tier 1 includes high-ranking elite universities. Tier 2 are 211 semi-elite universities, and Tier 3 are ordinary universities (or non-211 universities).

Table 1 Basic information for interview participants

Interviewees	Gender	Teaching experience (year)	Degree/research area	Student type	University type
T1	M	3	Master/linguistics	Non-English major	Tier 2
T2	M	14	Master/TESOL	Non-English major	Tier 2
T4	F	15	Master/linguistics	Non-English major	Tier 2
T6	F	7	Master/psychology	Non-English major	Tier2
T7	M	4	Master/pedagogy	Non-English major	Tier2
T8	M	20	Doctor/psycholinguistics	Non-English major	Tier2
T3	F	17	Bachelor/English teaching	English major	Tier3
T5	M	17	Doctor/linguistics	Non-English major	

questionnaire (2014) were quantified in the current study by specifying “a little” and “basic” as being one to two days and “advanced” being three or more than three days. In the last part of the questionnaire, two open-ended questions were added to get teachers’ perceptions about assessment training and the factors that may influence their training needs.

The questionnaire was piloted with five university English instructors. Then, changes were made to it, including rewording technical terms, removing similar items, reinterpreting unclear expressions, and readjusting sequence of items. The final questionnaire consisted of 23 items with 10 items for thematic area (1), four items for thematic area (2), and nine items for the third thematic area. The questionnaire was distributed through Wenjuanwang, an online survey website, to university English instructors in the central part of China. One month later, altogether 70 respondents returned the questionnaires. After reviewing, two submissions were discarded due to incomplete responses. The remaining 68 valid responses were exported from the website and loaded into SPSS 25.0 for analysis.

Following up the questionnaire, semi-structured interviews were conducted. Based on the questionnaire data, we constructed an interview protocol focusing on teachers’ assessment training experiences, assessment learning resources, classroom assessment practices, perception of assessment training needs, and factors influencing training needs. Probing questions were asked according to interviewees’ responses for clarification and elaboration. For easy understanding and communication, interviews were conducted in Chinese, with each interview lasting from 30 to 50 min. All the interviews were audio-recorded and transcribed verbatim, but only those used for excerpts in the result section were translated into English. Finally, the interview transcripts were sent to the interviewees for “member checking” (Creswell & Poth, 2018, p.261). We also examined national university English course syllabi, including GCET (Chinese Ministry of Education, 2017) and TGUEM (Chinese Ministry of Education, 2020) to gain insights into the assessment requirements for university English instructors.

Data analysis

The study used SPSS 25.0 to analyze questionnaire data. The internal consistency reliability of the questionnaire reached .967, indicating a strong internal consistency. The questionnaire data aimed at describing a general trend, so descriptive statistics of the

23 items were reported, including frequency and percentage, to align with the purpose of the study. As for the open-ended questions, frequency of recurring themes in the responses was calculated. The interview data were analyzed via both an inductive approach, namely, code-labeling and identifying recurring themes, and a deductive approach by referring to questionnaire items as categories. Finally, two national university English course syllabi were analyzed through content analysis focusing on assessment requirements and training. To ensure coding reliability, coding and thematizing were completed by the first author and then verified by the second author.

Results

Results from questionnaires

The assessment training situation and training needs were discussed according to the three thematic areas defined in the questionnaire (see the “Data collection” and “Analysis” sections).

Content and concepts

As shown in Table 2, more than half of respondents received no training in the content and concepts regarding the skills, quality, and impact of tests. Among these content and concepts, “social dimension” seems to be the most neglected area in training, with three quarters of respondents reporting no training in this aspect. “Statistics” is the second least trained area among respondents, 66.2% of whom claimed no training at all. In

Table 2 Respondents’ assessment training received and needed in *content and concepts*

	Training received			Training needed		
		Frequency	Percentage(%)		Frequency	Percentage(%)
Receptive skills (listening /reading)	Not at all	44	64.7	None	8	11.8
	A little	18	26.5	Basic	36	52.9
	Advanced	6	8.8	Advanced	24	35.3
Productive skills (writing/ speaking)	Not at all	44	64.7	None	6	8.8
	A little	20	29.4	Basic	36	52.9
	Advanced	4	5.8	Advanced	26	38.3
Integrated language skills	Not at all	37	54.4	None	6	8.8
	A little	23	38.8	Basic	35	51.5
	Advanced	8	11.8	Advanced	27	39.7
Validity	Not at all	37	54.4	None	6	8.8
	A little	25	36.8	Basic	41	60.3
	Advanced	6	8.8	Advanced	21	30.9
Reliability	Not at all	38	55.9	None	6	8.8
	A little	24	35.3	Basic	40	58.8
	Advanced	6	8.8	Advanced	22	32.4
Discrimination	Not at all	40	58.8	None	8	11.8
	A little	22	32.4	Basic	39	57.4
	Advanced	6	8.8	Advanced	21	30.9
Difficulty	Not at all	39	57.3	None	6	8.8
	A little	22	32.4	Basic	42	61.8
	Advanced	7	10.3	Advanced	20	29.4
Statistics	Not at all	45	66.2	None	5	7.4
	A little	18	26.5	Basic	31	45.6
	Advanced	5	7.4	Advanced	32	47.1
Social dimension	Not at all	51	75.0	None	8	11.8
	A little	12	17.7	Basic	39	57.4
	Advanced	5	7.4	Advanced	21	30.9

terms of assessment training needs, above 50% of respondents expressed a need for basic training in the content and concepts and slightly more than one third desired for advanced training with exception of the “Statistics” area, with almost half of respondents asking for this type of training. Overall, the great majority of instructors lacked training in all content areas but expressed more needs for basic training than for advanced training in these areas except “statistics”.

Purpose of testing

In the four areas covered in this theme (see Table 3), more than one quarter of respondents received no training in “giving grades” (48.5%), “placing students” (45.6%), and “selecting students” (48.5%), and they wished to have more basic training than advanced training in this area. It is interesting that although 36.8% and 47.1% of instructors had no and little training respectively in “finding out what was needed in teaching and learning”, nearly half of them expected either basic or advanced training in this area (52.9% and 45.6% respectively). The result may indicate that instructors attach importance to the connection between instruction and assessment.

Classroom-based assessment practices

Table 4 shows that more than half of respondents have received training in these practices except for “using portfolio” and “understanding the relationship between assessment and *China’s Standards of English Language Ability*”. Yet, nearly 90% of respondents aired a desire for both basic and advanced training in all these areas and more basic training was wanted than advanced training. Specifically, respondents had comparatively least advanced training in “using portfolio” (10.3%), “understanding the relationship between assessment and *China’s Standards of English Language Ability*” (11.8%), and “selecting & using ready-made tests” (11.8%), while they need the least advanced training in these three areas (26.5%, 23.5% and 22.1% respectively). With respect to using alternative assessment methods, most respondents had no or little training in these methods, particularly in “using portfolio”. Yet more than half of them expressed a need for basic training in these practices and less than one third of them needed advanced training except in “using integrated assessment” (35.3%).

Table 3 Respondents’ assessment training received and needed in *purpose of testing*

	Training received			Training needed		
		Frequency	Percentage(%)		Frequency	Percentage(%)
Giving grades	Not at all	33	48.5	None	6	8.8
	A little	27	39.7	Basic	44	64.7
	Advanced	8	11.8	Advanced	18	26.5
Finding out what is needed in teaching and learning	Not at all	25	36.8	None	1	1.5
	A little	32	47.1	Basic	36	52.9
	Advanced	11	16.2	Advanced	31	45.6
Placing students	Not at all	31	45.6	None	9	13.2
	A little	28	41.2	Basic	44	64.7
	Advanced	9	13.2	Advanced	15	22.1
Selecting students	Not at all	33	48.5	None	11	16.2
	A little	25	36.8	Basic	40	58.8
	Advanced	10	14.7	Advanced	17	25.0

Table 4 Respondents' assessment training received and needed in *classroom-based assessment practices*

	Training received			Training needed		
		Frequency	Percentage (%)		Frequency	Percentage (%)
Developing classroom-based assessment	Not at all	34	50.0	None	9	13.2
	A little	24	35.3	Basic	37	54.4
	Advanced	10	14.7	Advanced	22	32.4
Selecting and using ready-made tests	Not at all	28	41.2	None	9	13.2
	A little	32	47.1	Basic	44	64.7
	Advanced	8	11.8	Advanced	15	22.1
Giving feedback based on assessment	Not at all	29	42.7	None	6	8.8
	A little	28	41.2	Basic	43	63.2
	Advanced	11	16.2	Advanced	19	27.9
Using self -assessment	Not at all	32	47.1	None	6	8.8
	A little	26	38.2	Basic	43	63.2
	Advanced	10	14.7	Advanced	19	27.9
Using peer-assessment	Not at all	30	44.1	None	6	8.8
	A little	25	36.8	Basic	43	63.2
	Advanced	13	19.1	Advanced	19	27.9
Using integrated assessment	Not at all	28	41.2	None	5	7.4
	A little	26	38.2	Basic	39	57.4
	Advanced	14	20.6	Advanced	24	35.3
Using portfolio	Not at all	40	58.8	None	8	11.8
	A little	21	30.9	Basic	42	61.8
	Advanced	7	10.3	Advanced	18	26.5
Understanding the relation between assessment and <i>China's Standards of English Language Ability</i>	Not at all	36	52.9	None	5	7.4
	A little	24	35.3	Basic	47	69.1
	Advanced	8	11.8	Advanced	16	23.5
Preparing students for external tests (IELTS, TOEFL)	Not at all	29	42.7	None	10	14.7
	A little	25	36.8	Basic	36	52.9
	Advanced	14	20.6	Advanced	22	32.4
Selecting appropriate assessment methods	Not at all	33	48.5	None	5	7.4
	A little	26	38.2	Basic	41	60.3
	Advanced	9	13.2	Advanced	22	32.4

Overall findings

As is shown in Table 5, most instructors received the least training in the areas of assessment content and concepts, and more than one third of them aired a need for advanced training in the corresponding areas. In contrast, a number of teachers would rather have LAL basic training in purposes of testing and classroom-based practices (60.3% and 61.0% respectively) than advanced training, even though nearly half of them

Table 5 Overall trend of assessment training

Areas of assessment	Training received		Training needed	
		Percentage		Percentage
Content and concepts	Not at all	61.3%	None	9.6%
	A little	30.1%	Basic	55.4%
	Advanced	8.7%	Advanced	35.0 %
Purpose of testing	Not at all	44.9%	None	9.9%
	A little	41.2%	Basic	60.3%
	Advanced	14.0%	Advanced	29.8%
Classroom-based assessment practice	Not at all	46.9%	None	10.2%
	A little	37.8%	Basic	61.0%
	Advanced	15.3%	Advanced	28.8%

received no training in these two areas (44.9% and 46.9% respectively). Overall, the current study revealed that among university English instructors, the three areas of language testing and assessment were all underdeveloped, of which “content and concepts” was the least developed and that more than half of respondents voiced a need only for basic training in all areas instead of advanced training.

Results from interviews and open-ended questions

Assessment training and learning experiences

Language testing and assessment courses are the formal learning approach in undergraduate or postgraduate program (T1, 5, 6, 7), while a short period (one morning or afternoon) of training in rating translation, speaking, and writing for high-stakes tests is commonly seen among in-service instructors (T2, 3, 6, 7). Such concepts as validity, reliability, or rating and assessment purposes were learned from these courses or training programs. However, it seems that the knowledge and skills learned from assessment courses in the pre-service period failed to be applied and enacted in the instructors' class when they become in-service teachers (T1, 5, 6, 7). This is clearly demonstrated in T1's statement:

“[...] my understanding of language assessment still rests on what I learned from the testing and assessment courses in the postgraduate program. But I do not understand how to apply my understanding into my instructions. The application is zero. [...] Therefore, I lecture alone in my class with no assessment activities.”

Meanwhile, those instructors, who received short-term in-service assessment training, considered training helpful for them to prepare students for national high-stakes tests and concomitantly promote students' learning. T6 was quite aware of the role of assessment training in learning enhancement and well used assessment training to establish a connection between what she learned from training and how she taught in the class:

“I pay special attention to the training for CET4/6 provided by Shanghai Foreign Language Education Press. I find the topics used in CET4/6 are closely related to the topics covered in our textbooks. Then I integrate these topics with my classroom teaching, establishing a mindset for students that if they learn well in the class, good test results come naturally.”

Regarding the compensation for the lack of assessment training, most instructors revealed that they had never been exposed to any resources related to language testing and assessment because “my research area is not in this field and I am not interested in it” (T1, 2, 3, 4, 5, 8). The other instructors responded that they had learned language testing and assessment knowledge by reading books or papers and attending lectures or conferences, as they realized that “learning assessment is useful for teaching and learning” (T6, 7).

Classroom assessment practices

Integration of formative and summative assessment is conceptually rooted in English instructors' classes, but practically diverse assessment methods are rarely utilized for

formative assessment, with formal testing still being the principal assessment method. Most instructors admitted that oral presentation was frequently used to assess their students with teacher self-made criteria. Based on these criteria, they gave oral feedback to students' presentations. This showed that assessment on presentation was still teacher-centered. Peer- and self-assessments found no way in this learning activities (T1, 2, 3, 4, 5, 8), as "students complained peer-assessment was detrimental to their friendship with other fellow students" (T3) and "the beneficial effect of self-assessment was not evident due to a lack of explicit standards or criteria" (T1). Automated essay scoring system was recommended as another assessment method by instructors as "it is novel, arouses students' interests and saves a lot of time (T2)", and "it can increase the scoring validity by combing my rating with the machine scoring" (T6). As for the other alternative assessment methods like portfolios, all instructors pointed out that they had the least idea of how to conduct, which is consistent with the finding from the questionnaire that instructors received no or little training of alternative assessment methods.

As formal testing is still the primary assessment method at the end of the courses, all instructors had the experience of preparing tests for students. They all selected ready-made tests with minor alterations, although most instructors reported they had no training in this area, which was also verified by questionnaire data. In terms of how to select the ready-made tests, most instructors were concerned about students' proficiency (T8), item difficulty (T4), and orientation to standardized tests (CET4/6 in China) (T2).

Assessment training needs

88.2% of respondents ($n = 60$) who answered the open-ended questions aspired to have assessment training needs. Similarly, all interviewed instructors aired their needs to have assessment training because they believed assessment training contributed to their assessment practice, an integral part of teaching, hence to students' learning. Regarding specific training needs, while half of the instructors were hesitant to respond, half of them responded that they expected to have training in specific assessment skills, such as scale development, item writing, and result interpretation. T3 preferred to improve specific assessment knowledge and skills, such as interpreting test results, as she stressed:

"I want to know the basic knowledge about formative assessment and know why I should implement formative assessment and how I can get the data from formative assessment and interpret them to my students."

T5 longed for the need for the assessment skills that could be applied into the classroom practices:

"When my students were reading an article, if I could have designed some questions to lead them to discuss insightfully, I think the learning effect would have been better. But I have no assessment knowledge or skills to design such questions at all."

Both T3 and T5 seemed to focus more on the training need of testing-specific skills instead of assessment-specific ones. Differently, the other two instructors disclosed a

wish to learn assessment theories through regular seminars, workshops or learning/academic community (T6, 7) because “without the relevant theories, you can’t judge if your assessment practices are appropriate or not [...]” (T7)

Influential factors on assessment training

All instructors were vocal about the challenges they had encountered in taking assessment training. The open-ended questions in the questionnaire revealed that more than one third of instructors ($n = 26$) had tight teaching schedules or great family responsibilities that they could not devote much time to assessment training. “Family”, “kids”, and “heavy teaching load” were the words repeatedly mentioned in the interviews (T1, 2, 8). Nearly half of the open-ended question respondents ($n = 32$) also reported that they got no support from the university or the department in assessment training. This result could be echoed by T2, 3, 4, 7, and 8 in their interviews, of which T2’s excerpt is as follows:

“The university or the department will interfere instructors’ development in whatever aspects. We must obey the university/department rules. If we want to develop our LAL, we must get support from the university or the department. However, the university or department usually pays little attention to our assessment training (T2).”

About one third of the open-ended question respondents ($n = 20$) held that their own research background had influenced their assessment training needs. Similarly, interviewed instructors (T1, 2, 3, 4, 5, 8) admitted that owing to non-assessment research background, they paid little or no attention to the assessment area, and in turn they had no intent of taking assessment training. This claim can also be seen from T1 and T8:

“My research interest is corpus linguistics and I show no interest in language testing and assessment at all. I don’t think it’s necessary to have training related to this area because I can manage the assessment with my current pedagogical knowledge (T1).”

“[...] my research area is in psycholinguistics. Though it is relative to testing and assessment to some extent, I show little interest in LAL, not to mention assessment training (T8).”

Students’ language proficiency is perceived as a major challenge on instructors’ assessment training, as T8 said:

“The students I teach have very low English proficiency. I find it difficult implementing formative assessment activities in the classroom. Therefore, I have never used the so-called self-assessment or peer-assessment, hence needless for assessment training.”

Results from document analysis

Assessment requirements and assessment training analyzed from the national university English course syllabi were presented as follows.

In GCET (Chinese Ministry of Education, 2017), an official document for non-English majors, the assessment requirements can be identified as assessment purposes and assessment methods. The purposes of assessment refer to assessment for learning (formative) and assessment of learning (summative). Assessment for learning appears when information obtained from testing and assessment is used to promote learning and instruction. For example, “assessment and testing should be used to provide diagnosis and feedback for teaching and learning so as to improve students’ language proficiency and competency (GCET, p.9).” Assessment of learning means assessment is used to assess students’ language ability. For instance, the guide states that “...testing content must be closely relevant to learning content...and...a diversity of tests should be used to evaluate comprehensively students’ language ability... (GCET, p.10)”. To serve the two purposes, a variety of testing and assessment methods were listed to assess students including formative and summative, integrated-skills and single-skill, external and internal, students’ portfolios and peer/self-assessment. However, the document does not explain why and how to use these methods in teaching practice.

With respect to assessment training, GCET (Chinese Ministry of Education, 2017) requires that the administration at the university or the department level should provide financial and material support to improve instructors’ assessment competence by training them of assessment knowledge and skills, particularly the theories and practices related to formative assessment. In this way, “the instructors are able to use various methods to analyze assessment data and interpret the testing/assessment results based on which teaching and learning can be promoted (GCET, p.14).” According to GCET, assessment training is viewed as a guarantee to the quality of testing and assessment for university English teaching through enhancing instructors’ LAL.

When examining TGUEM (Chinese Ministry of Education, 2020), we found that the guide highlighted the use of both formative assessment and summative assessment and of various assessment methods for the purpose of promoting students’ learning. Yet, no specific assessment methods were provided and no assessment training for instructors were covered.

Discussion

The current study aims to investigate assessment training of university English instructors in China. It revealed that most university English instructors received insufficient assessment training to qualify themselves for assessment tasks and as a result, they had a concerningly low LAL. The finding has been confirmed in other contexts (e.g., Fulcher, 2012; Vogt & Tsagari, 2014), whereas it differs from that of Lam’s study (2019), in which the participants were reported to have a higher profile of professional training in language assessment in Hong Kong. The difference may be explained by the facts that in Hong Kong, teachers need professional training to pass the Language Proficiency Assessment for Teachers and to obtain equivalent qualifications to start a teaching career (Coniam & Falvey, 2013) and that language assessment training was widely provided for pre-service teachers in this region (Lam, 2015). English teachers at either the university or the school level in mainland China are not mandated to get such a professional qualification. They simply take exit exams of education and educational psychology in the short-term orientation training to obtain an accreditation certificate signaling the commencement of teaching profession (Xu, 2017).

Despite lack of assessment training, university English instructors did not resort to other resources to make up for the void. This finding is incongruent with that reported in Tsagari and Vogt's study (2017) where teachers turned to books, papers, or colleagues to compensate for their insufficient LAL. Although these compensation strategies on the job are considered detrimental to the use of innovative assessment (Vogt & Tsagari, 2014; Lan & Fan, 2019), they are the effective methods to enhance teachers' LAL temporarily, given the absence of formal training and the lack of practical guidelines from the national documents, as these strategies can expose teachers to basic assessment knowledge or theories needed in language assessment.

The findings also indicated that assessment training across varied career stages took different effect in English instruction. The short-term assessment training for high-stakes tests with a focus on assessment practical skills was considered beneficial for instructors' LAL development and classroom assessment practices. Notably, language assessment courses instructors took in their pre-service program contributed little to the application of their assessment knowledge or skills into practices. The failure of application is possibly related to the substantially smaller portion of instruction time for classroom practice in the language testing and assessment courses (Jin, 2010). Without hands-on activities in classroom, it is hard for learners to transmit their explicit assessment knowledge into practices. This underscores the balanced coverage of assessment theories and practices in assessment training.

Another important finding was that a majority of university English instructors expressed assessment training needs and prioritized their needs in certain assessment areas. Specifically, they had the least training in the "content and concepts" area and expressed their desire for more advanced training in this area. The area of "classroom-based assessment practices" was the second least trained aspect, but very few teachers wanted to receive advanced training in this area. It is surprising to note that most university English instructors attached more importance to assessment training in "content and concepts" than in "classroom-based assessment practices", the skills most needed for their daily teaching. The result corroborates that of Vogt and Tsagari's study (2014). This higher need for advanced training in the "content and concepts" area may be related to the effect of learning on the job (Vogt & Tsagari, 2014), suggesting that teachers believe understanding the theoretical terms may enhance their confidence in critically evaluating the quality of assessment practices (Lan & Fan, 2019). Interestingly, among the subskills of the "content and concepts" area, nearly half of the respondents put forward a need for advanced training in "statistics". The similar finding can be found in the study by Sheehan, Berry, and Munro (2016), showing that statistics anxiety is usually considered a major barrier for most language instructors to be involved in language testing and assessment (Brown, 2013).

With respect to assessment training needs for classroom-based assessment practices, a pattern seems to emerge. A great majority of instructors received no or a little training in classroom-based practices, especially in alternative assessment methods such as "using portfolio", "using peer/self-assessment", and "using integrated-skill assessment", but only desired for less advanced training in these innovative methods, despite instructors' little know-how of these methods and the reiteration of using these methods by the national documents. The need for basic assessment training can be possibly accounted for by personal factors (e.g., research interest) or contextual constraints (e.g.,

insufficient support from the department). However, it is noted that, during the interviews, some instructors struggled when asked to name assessment aspects they needed to be trained. This is consistent with Tsagari and Vogt's (2017) finding that language teachers had difficulty in specifying their professional development needs. The difficulty appears when they fail to practice their preached assessment in the classroom or when teachers lack means or skills to reflect systematically on the nature and quality of their classroom assessment practices (Hill, 2017).

Multiple personal and contextual constraints have been found to justify the preference for basic assessment training among university English instructors. Non-language assessment research interest/area pushes university English instructors into a mindset that they simply need a basic understanding of language assessment instead of the expertise like an assessment expert. Balance between workload and family pressure is a crucial factor that constrains instructors from devoting time to advanced assessment training, which is also confirmed in Yan, Zhang, and Fan's study (2018). A lack of support from the departments or universities may discourage university English instructors from being enrolled in any long-term assessment training courses or workshops. In China, the evaluation of faculty members, including university English instructors, still focuses on their academic publications instead of their instructions (Mohrman, Geng, & Wang, 2011). Professional development regarding teaching skills, such as assessment training, is not a priority of the department or the university. In addition, students' low language proficiency may deter instructors from implementing alternative assessment methods in the class, which leads to teachers' belief that they need no profound training in assessment knowledge, skills, and practices.

All factors discussed above suggest that assessment training has become an inevitable dilemma for English instructors. They are quite aware of the vital role of assessment training in improving their assessment practices on the one hand, and on the other hand they feel trapped in the individual, institutional, or contextual quagmire, being unable to take advanced training. The dilemma emerges, as assessment policies are usually enforced in a top-down way (Sheehan & Munro, 2019) in many countries, including China, leaving teachers with little say in assessment training. Yet, we argue that university English instructors may exercise their agency in assessment to mitigate the constraints to achieve some degree of familiarity with the fundamental principles and practices of assessment. And this familiarity should "be embedded from an early stage in the training of language teachers" (Taylor, 2013, p.407).

The findings also implied that university English instructors were given few opportunities for assessment training, as it was greatly neglected by the administration at the department or university level, regardless of the official reiteration of providing training to promote instructors' LAL. This neglect confirms Lam (2015)'s finding that language assessment training got insufficient support to promote LAL from local institutions such as in Hong Kong. One of the reasons for the negligence is pertinent to the previously mentioned imbalance of requirements on instruction and research for instructors in mainland China. Another possible explanation is the assumption that teachers have no external motivation to become assessment-competent in their working contexts, because most textbook publishers have provided ready-made quizzes, unit tests, and even item banks that teachers need not create their own assessment tasks (Stiggins, 1993), hence needless for assessment training. In sum, like LAL development, language

assessment training is also context-dependent (Yan, Zhang, & Fan, 2018), confined to administrative mandates and institutional policies.

Conclusion

The study has explored assessment training needs of university English instructors in the Chinese context. Generally, university English instructors did not receive sufficient assessment training to well prepare themselves for assessment practices. Despite this situation, instructors prefer basic training lasting one to two days instead of advanced training (three or more than three days). Such preference may be attributed to various constraints from personal and contextual levels. These findings can add our existing knowledge of using assessment training to promote LAL by taking assessment training needs and the influential factors into consideration. Findings from the current study also call for further investigations into assessment training of university English instructors. For instance, it is worth exploring whether instructors of diverse teaching experiences or from different tiers of universities may have differing training needs.

Like other research in assessment training, the study is not without limitations. First, it is a small-scale research project with 68 questionnaire respondents and eight interviewees. Despite data triangulation employed in the study, the trustworthiness of the study would be reduced to some extent. Second, the terms like “basic” and “advanced” used in the questionnaire are still ambiguous, although we have quantified them by being day-based. The ambiguity of terms may cause different interpretations, hence different responses from participants. Notwithstanding these limitations, we believe a number of practical implications can be drawn from this study to improve assessment training for university English instructors in China.

First of all, the university administrators and department heads should be mindful of their instructors' assessment practices, prioritize assessment training for the in-service instructors, and provide more opportunities and financial support for them to receive assessment training to keep updated in language testing and assessment. The first change that can be made is to mandate language assessment as an integral part in the orientation training for the newly recruited university English instructors and as an exit benchmark of the training. Second, assessment training providers and training instructors should track instructors' assessment practices, notice their training needs, provide regular in-service training, and tailor the training contents, materials, and approaches to cater for different needs. For instance, our findings showed that most instructors had experience of selecting ready-made tests for students. Therefore, we suggest that the content of how to select ready-made tests should be prioritized in assessment training. Third, we propose that modes of assessment training can be flexible, given multiple constraints. Assessment training providers can create an online course, just like a flipped classroom, to upload and upgrade assessment training materials and provide online tutorials for those instructors who need help in assessment-related issues. Fourth, individual English instructors are advised to continue language testing and assessment learning to keep themselves abreast with current development in this field and to achieve self-instruction (Malone, 2017). They can peer observe colleagues who are skilled in assessment, as these teachers' assessment expertise can be a valuable source for assessment training (Wang, 2017) and LAL development of the instructors.

Abbreviations

CEFR: Common European Framework of Reference for Language; CET4/6: College English Test Band 4/6; EFL: English as a foreign language; IELTS: International English Language Testing System; GCET: A Guide to College English Teaching; LAL: Language assessment literacy; T1, 2, 3, 4, 5, 6, 7, 8: Teacher 1, 2, 3, 4, 5, 6, 7, 8; TGUEM: Teaching Guide for Undergraduate English Major; TOFEL: Test of English as a Foreign Language

Acknowledgments

We would like to thank Dr. CHEN Meilin from the Language Center, Hong Kong Baptist University, who kindly proofread the first version of this manuscript and Dr. JIANG Changsheng at Zhongnan University of Economics and Law, who shared his critical comments in the manuscript revision. We would also like to thank all participants for their devotion to this study.

Authors' contributions

The manuscript was written by the first author. The second author provided valuable feedback on different revisions of the manuscript. Both authors read and approved the final manuscript.

Authors' information

Ling Gan is a PhD student in the Department of Education Studies at Hong Kong Baptist University. Her research interests include language assessment literacy and language teacher professional development. Ricky Lam is an associate professor in the Department of Education Studies at Hong Kong Baptist University. He is also the President of Hong Kong Association for Applied Linguistics (HAAL). His research interests include digital writing portfolios; portfolio assessment, second language writing assessment, teacher assessment literacy, assessment for learning and assessment as learning.

Funding

The study was supported by "the Fundamental Research Funds for the Central Universities", Zhongnan University of Economics and Law (2722019JCG042).

Availability of data and materials

The data are available upon request from the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Received: 9 July 2020 Accepted: 26 August 2020

Published online: 18 September 2020

References

- Boyd, E., & Donnarumma, D. (2018). Assessment literacy for teachers: a pilot study investigating the challenges, benefits and impact of assessment literacy training. In D. Xerri, & P. Vella Briffa (Eds.), *Teacher Involvement in High-Stakes Language Testing*, (pp. 105–126). Cham: Springer International Publishing AG.
- Brown, J. D. (2013). Teaching statistics in language testing courses. *Language Assessment Quarterly*, 10(3), 351–369.
- Chinese Ministry of Education. (2017). *A Guide to College English Teaching*. Retrieve from <https://wenku.baidu.com/view/a514b9865ebfc77da26925c52cc58bd63086936d.html?re=view>
- Chinese Ministry of Education (2020). *Teaching Guide for Undergraduate English Major*. Beijing: Foreign Language Teaching and Research Press.
- Chung, S. J., & Nam, Y. J. (2018). Language assessment literacy of Korean EFL teachers: an investigation of their training experiences and needs. *Modern English Education*, 19(1), 38–48.
- Coniam, D., & Falvey, P. (2013). Ten years on: The Hong Kong language proficiency assessment for teachers of English (LPAT E). *Language Testing*, 30(1), 147–155.
- Creswell, J., & Poth, C. (2018). *Qualitative inquiry and research design: choosing among five approaches*, (4th ed.). Washington D. C: SAGE.
- Crusan, D., Plakans, L., & Gebriel, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43–56.
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419–438.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132.
- Guerin, M. C. (2010). Initial findings from a pilot Italian study of foreign language teachers' stated language assessment knowledge-base and needs. *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching*, 4, 80–103
- 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). Berlin: Mouton de Gruyter.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). European survey of language testing and assessment needs. Part one: General findings. Retrieved from <http://www.ealta.eu.org/documents/resources/survey-report-pt1.pdf>
- Hill, K. (2017). Understanding classroom-based assessment practices: a precondition for teacher assessment literacy. *Papers in Language Testing and Assessment*, 6(1), 1–17.
- Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, & N. Hornberger (Eds.), *Language Testing and Assessment, Encyclopedia of Language and Education*, (3rd ed., pp. 257–286). Berlin: Springer.
- 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.
- Jin, Y. (2018). Defining and developing the assessment literacy of foreign language teachers. *Foreign Language Education in China*, 1(2), 65–72.

- Lan, C., & Fan, S. (2019). Developing classroom-based language assessment literacy for in-service EFL teachers: the gaps. *Studies in Educational Evaluation*, 61, 112–122.
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–97.
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, 81, 78–89.
- Lee, J. (2019). A training project to develop teachers' assessment literacy. In E. White, & T. Delaney (Eds.), *Handbook of research on assessment literacy and teacher-made testing in the language classroom*, (pp. 58–80). Hershey, PA: IGI Global.
- Malone, M. (2017). Training in language assessment. In E. Shohamy, & N. Hornberger (Eds.), *Language Testing and Assessment, Encyclopedia of Language and Education*, (3rd ed., pp. 225–239). Berlin: Springer.
- Malone, M. E. (2008). Training in language assessment. In E. Shohamy, & N. Hornberger (Eds.), *Language Testing and Assessment, Encyclopedia of Language and Education*, (2nd ed., pp. 225–240). New York: Springer.
- Mohrman, K., Geng, Y. Q., & Wang, Y. J. (2011). Faculty life in China. *The NEA 2011 Almanac of higher education*, 83–99. Retrieved from https://www.researchgate.net/publication/265004900_Faculty_Life_in_China/link/5672e6be08ae04d9b099b1af/download.
- Qian, D. D., & Cumming, A. (2017). Researching English language assessment in China: focusing on high-stakes testing. *Language Assessment Quarterly*, 14(2), 97–100.
- 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.
- Sheehan, S., Berry, V., & Munro, S. (2016, September 16–17). *Singing from the same hymn sheet? What language assessment literacy means to teachers*. In: Language Assessment Literacy Symposium, Lancaster University.
- Sheehan, S., & Munro, S. (2019). *Classroom assessment: the development of teacher's cognition*. London: British Council.
- Stiggins, R. J. (1993). Teacher training in assessment: overcoming the neglect. *Teacher Training in Measurement and Assessment Skills*. 4. Retrieved from <https://digitalcommons.unl.edu/burosteachertraining/4>
- Sultan, N. (2019). Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*. <https://doi.org/10.1186/s40468-019-0077-8>.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412.
- Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41–59.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402.
- Wang, X. (2017). A Chinese EFL teacher's classroom assessment practices. *Language Assessment Quarterly*, 14(4), 312–327.
- Xie, Q., & Tan, S. Y. (2019). Preparing primary English teachers in Hong Kong: focusing on language assessment literacy. *Journal of Asia TEFL*, 16(1), 653–673.
- Xu, Y. (2017). *Assessment literacy of university English instructors in China: a mixed-methods study (Unpublished Doctoral Thesis)*. Hong Kong, China: University of Hong Kong.
- Xu, Y., & Brown, G. T. L. (2017). University English teacher assessment literacy: a survey-test report from China. *Papers in Language Testing and Assessment*, 6(1), 133–156.
- Yan, X., Zhang, C., & Fan, J. S. (2018). "Assessment knowledge is important, but..." How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158–168.

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)
