

The Conceptualization of Language Assessment Literacy by Pre-service EFL Teachers in China

Juanjuan Geng

Foreign Language Colleges, Xianyang Normal University, Xianyang, China
mylovegjj123@gmail.com

Abstract: *The conceptualization of language assessment literacy (LAL) has been widely discussed from the perspectives of researchers and in-service teachers. However, the long-neglected group, the pre-service English as a Foreign Language (EFL) teachers, have been seldomly explored. Thus, in order to shed light on a more comprehensive picture of LAL, the present study purposively sampled six participants from a Normal University in China. A semi-structured interview was conducted and then it was transcribed and analyzed thematically. The findings revealed that 1) they tended to conceptualize it as an eight-dimensional: Discipline-related competence, person-related competence, and assessment-related competence. 2) They attached considerable importance to disciplinary competence and practical implementation of assessment methods while they tended to give scarce attention to pedagogical content knowledge and the potential influence of assessment. 3) LAL was dependent on the social cultural context. 4) Each participant's spider diagram of LAL conceptualization was unique and different from others.*

Keywords: Language assessment literacy, Pre-service teachers, Conceptualization.

1. Introduction

Since it was formulated to integrate assessment literacy (AL) with the field of language assessment by Brindley in 2001, language assessment literacy (LAL) has become the focal point in the research agenda (Babaii & Asadnia, 2019; Firoozi et al., 2019; Kavaklı & Arslan, 2019; Shahzamani & Tahririan, 2021). Issues in LAL have been widely addressed but not completely resolved yet (Giraldo, 2018). LAL is still in its infantile stage concerning the number of the theoretical and empirical studies in the field (Fulcher, 2012; Giraldo, 2018; Hakim, 2015; Jin, 2018). Hence, the theme LAL is under-explored and more research is warranted (Kim et al., 2020; Nimehchisalema & Bhatti, 2019; Yan & Fan, 2020).

Various attempts have been made to conceptualize LAL from the perspectives of professional researchers and in-service teachers (Butler et al., 2021). For instance, Shahzamani and Tahririan (2021) define LAL as language teachers' understanding of diverse purposes of assessment and knowledge of applying them accordingly. LAL is primarily conceptualized as a multi-dimensional concept shaped by the context where it is embedded, as illustrated in the existing types of LAL conceptualization frameworks: three-component models (Fulcher, 2012; Giraldo, 2018; Inbar-Lourie, 2008), scaled models (Bøhn & Tsagari, 2021; Kremmel & Harding, 2020).

While further research is still encouraged to take the perspectives from other stakeholders, especially the pre-service English as Foreign Language (EFL) teachers into consideration (Bøhn & Tsagari, 2021; Nimehchisalema & Bhatti, 2019; Taylor, 2013). It would be intriguing to find out whether the pre-service EFL teachers' LAL conceptualization is in alignment with the existing models. As their voice should be heard and addressed for a more comprehensive and deeper view of the conceptualization of LAL. Moreover, the elaboration of LAL conceptualization may lay a research foundation for subsequent studies, such as describing the LAL profile or selecting priorities in LAL improvement programmes for the target participants. To address the paucity

in research, the present study is to focus on pre-service EFL teachers to explore their conceptualization of LAL.

2. Literature Review

AL, first coined by Stiggins (1991) in general education, was defined from the description of the assessment literates who could discern the high-quality assessment from unsound ones. Nearly a decade later, one of the earliest attempts to integrate AL with language assessment was Brindley (2001), who argued for a curriculum-based LAL, although he did not specifically address LAL (Fulcher, 2012; Hildén & Fröjdendahl, 2018; Inbar-Lourie, 2017). Alongside the progression, the frameworks of LAL have also evolved from a mere focus on assessment-related knowledge and competence to a broader focus on the context-dependent interactive process.

As illustrated in Table 1, there are primarily two types of models focusing on LAL. The first type conceptualizes it as a three-tiered concept, which is mediated by a variety of factors. After examining the changes in language testing textbooks, Davies (2008) noted the emerging importance of principles and supplemented it into "Knowledge + Skills + Principles" model (p. 335), in which knowledge referred to the information about assessment and language; skills contained the assessment strategies or expertise; and principles entailed concepts underlying testing such as validity, reliability, and ethics as well as professionalism. This model has been still prevalent in more recent studies (e.g., Giraldo, 2018). Likewise, Inbar-Lourie (2008) established a framework for courses in language assessment, labeling "Why + What + How" model (p. 390). Based on the need analysis of in-service EFL teachers, Fulcher's (2012) hierarchy clarified the inter-related relationship among the three elements and addressed the issue of LAL in a much wider context. In his three-tiered Practices + Principles + Contexts framework, the bottom was the practice of language assessment (knowledge, skills, and abilities) with principles (processes, principles, and concepts) positioned in the middle serving as the guidance of the practice, and the top was historical, social, political, and

philosophical contexts (origins, reasons, and impacts).

Though helpful in identifying the key components in complex contexts, the three-element models had the deficiency that to what extent teachers should be assessment literate was not mentioned, just as what Taylor (2013) commented that the depth of LAL required for teachers was unclear. Thus, to address the gap, another type of models characterized by the multi-dimensional continuum emerged.

This updated type was featured by multi-dimensional scaled models with varying degree of expectations for teachers, specifically Pill and Harding's (2013) continuum scale, Taylor's (2013) eight-dimensional model, Baker and Riches' (2017) seven-dimensional model, Kremmel and Harding's (2020) nine-dimensional model, and Bøhn and Tzagari's (2021) ten-dimensional model.

Taylor's (2013) eight-dimensional spider diagram might be the most impactful and subsequent studies still attempted to validate the classifications. Inspired by Pill and Harding's (2013) continuum of LAL in rejection of dichotomy of literate or illiterate with a preference to view it from a continuum, from "illiteracy", through "nominal literacy", "functional literacy", "procedural and conceptual literacy", to an expert level of knowledge (p. 183), Taylor (2013) hypothesized a model for teacher's LAL, involving (1) "knowledge of theory", (2) "technical skills", (3) "principles and concepts", (4) "language pedagogy", (5) "socio-cultural values", (6) "local practices", (7) "personal beliefs/attitudes", and (8) "scores and decision making" (p. 410). The teachers were suggested to be assessment literate to varying degrees among the eight aspects with the highest mastery level in language pedagogy and the lowest mastery level in three dimensions: knowledge of theory, principles & concepts, as well as scores and decision-making.

This model was recently reclassified and modified by Baker and Riches (2017), Kremmel and Harding (2020) and Bøhn and Tzagari (2021). According to Baker and Riches (2017), the dimension of socio-cultural values suggested by Taylor (2013) might best be considered as background or contextual influence implicitly subsumed within all other components instead of an independent one. They relabeled as (1) "theoretical and conceptual knowledge", (2) "task performance", (3) "language pedagogy", (4) "collaboration", (5) "awareness of local practices", (6) "awareness of personal beliefs/attitudes", and (7) "decision making", with all the dimensions positioned under the impact of the social-cultural values (p. 99). Another difference from Taylor's (2013) description was that language teachers were supposed to be most capable in three dimensions in total (i.e., language pedagogy, awareness of local practices, and awareness of personal beliefs/attitudes) and to be least capable in theoretical and conceptual knowledge.

Another attempt was made by Kremmel and Harding (2020) who intended to validate Taylor's (2013) LAL profile model, which was speculative in nature (Stabler-Havener, 2018). Kremmel and Harding (2020) revealed that there might exist a nine-component construct of LAL, which was generally in alignment with Taylor's (2013) diagram but adding some expansions: (1) "developing and administering language

assessments", (2) "assessment in language pedagogy", (3) "assessment policy and local practices", (4) "personal beliefs and attitudes", (5) "statistical and research methods", (6) "assessment principles and interpretation", (7) "language structure, use and development", (8) "washback and preparation", and (9) "scoring and rating" (p. 111). Consistently, in their proposal language teachers were described to be least capable in the dimension of statistical knowledge.

Likewise, Bøhn and Tzagari (2021) criticized Taylor's (2013) model for not specifying the components or dimensions in any great detail. Thus, they provided an extended version of Taylor's (2013) scaled model with great details to explain the meaning of each revised dimension. They suggested there were 10 dimensions in teachers' LAL: "knowledge of assessment theory, technical skills, principles, language pedagogy, social-cultural values, local practices, personal beliefs/attitudes, scoring, disciplinary competence, and collaboration competence" (p. 231). Teachers should be most proficient in three dimensions (i.e., principles, language pedagogy, and disciplinary competence) and least proficient in both technical skills and collaboration competence.

To sum up, the conceptualizations of LAL are clarified based on its evolvement from the traditional understanding as a combination of assessment-related knowledge plus skills to a dynamic, multi-layered and complex notion of a process in which LAL is not a static concept but is constantly embedded in the local contexts. Much have been done on the conceptualization of LAL from both in-service teachers' and assessment experts' perspective, however, what is still unclear is pre-service EFL teachers' conceptualization of LAL.

3. Method

Since the goal of the current study is to explore pre-service EFL teachers' LAL conceptualization, the semi-structured interview is adopted for its appropriateness to obtain a complicated and detailed understanding of the issue, to empower participants to voice their opinions, to share their experiences, and to understand the contexts or settings where participants resolve the issue (Creswell, 2007). The semi-structured interview with a general structure set beforehand covers the main questions to be asked and the detailed questions are left to be worked out during the flow of interviewing (Drever, 1995).

The interview protocol was designed by the researcher based on research objective, assessment practices in China, and prior studies by Bolívar (2020), and Yan and Fan (2020). Then it was refined by two of the researcher's colleagues in assessment and teacher education respectively to check its alignment with the research purpose as well as the structure and comprehensibility.

4. Participants

The participants selected in the study came from X Normal University in China. The familiarity with the contexts makes it more possible to establish a greater degree of rapport and trust between participants and researcher. People are usually

reluctant to do self-exploration of their experience and feelings to strangers, however, participants are more open to sharing if they trust the interviewer in most cases (Polkinghorne, 2007).

In this study, purposeful sampling is utilized. Purposeful sampling is defined as “researchers intentionally select individuals and sites to learn or understand the central phenomenon” (Creswell, 2012, p. 206). The participants are sampled on different dimensions of characteristics or traits (Creswell, 2012). Rather than attempting to be generalizable to other contexts, the maximal variation sampling method intends to elicit multiple perspectives, therefore participants are purposefully as different from each other as possible for a diversity of views (Creswell, 2007, 2012).

As to the number of participants, qualitative research does not have any specific rules for a study (deMarrais, 2004). The qualitative research tends to explore depth and details, usually focusing on a few participants is more proper (Fink, 2000; Miles & Huberman, 1994). Bearing this in mind, my study selected six participants to explore their LAL. To be specific, they were suggested by the counselors who were in charge of the senior students. Among the potential participants, the researcher purposefully chose six participants with varying background information, such as gender, class number, practicum field school, and English proficiency. The researcher contacted the six participants individually through text-messages for inviting to take part in the research voluntarily. Then the researcher organized a meeting at their convenient time to introduce the details of the research and get the consent form signed.

5. Data Collection and Analysis

The interviews were arranged in a small meeting room at the available time of the participants. The 20-minute interviews were audio-taped under the consent of the participants. After the completion of the interview, the researcher transcribed the audio recordings verbatim in Chinese. The original interview transcripts were recorded and analyzed in Chinese in order not to lose meaning and avoid any misunderstandings because “some nuances of one language may never be adequately translated into another” (Esin et al., 2014, p. 208). The translated excerpts in English were presented in the data citations, as recommended by Van Nes et al. (2010) to stick to original language as long and as much as possible to avoid potential limitations of translation from source language to target language in qualitative data analysis.

Thematic analysis was adopted for it helping to theorize across a set of cases and to find thematic elements across participants and the events elicited from the participants (Riessman, 2005; Vaismoradi et al., 2013). The thematic analysis, like most qualitative analysis methods, is a non-linear process but iterative and recursive: the researcher moves back and forth among the different phases (Braun & Clarke, 2006). (1) The first phase is to familiarize with the data set, often beginning during data collection. The researcher actively engaged in the data by attentively reading the textual data to generate very early and provisional analytic ideas. (2) The second phase aims to generate codes. Coding means identifying relevant data and labeling them with a few

words that cover the meaning of the segment. Coding helps the researcher develop an insightful understanding of the data and provides a thorough basis for the analysis. The researcher embraced an open and flexible attitude to coding, refining, and revising codes throughout the entire research process. (3) The third procedure involves constructing themes. Guided by the research questions, the researcher examined, combined, and clustered the codes together into a possible theme. (4) The fourth procedure is to review the potential themes respectively. After all candidate themes were developed, the researcher reviewed the network of themes to further shape, clarify, or even reject the theme. (5) The fifth step is to define and name themes. Defining a theme means providing a short summary of the core concept and abstract of each theme. It is concerned about guaranteeing the clarity, cohesion, precision, and quality of developing thematic analysis. (6) The final phase is writing the report to summarise the findings.

6. Results

The investigation of the response elicited from the six participants yielded that LAL was composed of eight dimensions with varying importance owing to the shaping effect of the social cultural context. Meanwhile, each participant’s LAL conceptualization profile was unique from others. The eight dimensions in LAL were listed as followed:

- 1) Disciplinary & cross-disciplinary competence: means knowledge of language, English curriculum, and other related disciplines, as well as a command of English.
- 2) Pedagogical content knowledge: means knowledge of how to teach the curriculum-based content to learners and the competence to sustain learning.
- 3) Personal conception & attribute: means knowledge of how one’s own preconceptions, understandings and opinions may inform one’s conceptualizations, interpretations, judgments and decisions in assessment, and the individual’s personality.
- 4) Assessment principle & ethics: means knowledge of why to assess (i.e., formative, summative) and understanding legal and ethical responsibilities concerning the design, use, storage, and dissemination of assessment.
- 5) Assessment methods & implementation: means knowledge of a wide range of assessment strategies and competence of applying and carrying out them for the target learners.
- 6) Assessment content & criterion: means knowledge of how to assess the learning goals and specific content being learned (academic achievement or affective performance), and knowledge of rationale for grading or rubrics.
- 7) Assessment washback: means knowledge of potential influence of assessment, whether beneficial or damaging, on teaching and learning.
- 8) Assessment interpretation & communication: means knowledge of ways of interpreting evidence generated from assessment, and ways of communicating assessment results to stakeholders such as students, parents, managers/administrators, and the general public.

6.1 Conceptualization of LAL from Participant 1

From Amy's definition, she considered LAL as a comprehensive term covering all the assessment-related tasks conducted on four skills: listening, speaking, reading, and writing in English. She also provided a list of constituents: familiarity with English curriculum, competence of sustaining English learning, knowledge of Educational Psychology, character of teachers (patient, responsible, and adaptable), knowledge of assessment purpose & ethics, mastery of individualized assessment methods, and competence of communicating feedback to students. Among the list, knowledge of Educational Psychology was given the greatest priority as teachers were believed to know students and know how they learn first, which could be seen from the excerpt of the interview data.

[EFL teachers] should be familiar with English curriculum they are teaching, for example if they teach in primary school, they must know the entire framework from grade three to grade six...they should learn Educational Psychology, the most important among all the components, to better understand the mechanism of student learning and then they can design individualized assessment for student benefit as well as for students learning...they are supposed to be patient and responsible during the assessment process, adaptable to students' response...besides, communication of feedback is also important to convince student into believing their assessment results. ...Another dimension is to encourage students to sustain their passion and enthusiasm for English learning.

For a more vivid presentation, Amy's conceptualization of LAL was depicted in the spider diagram, with the most important dimension labelled with 2 and other dimensions labelled with 1 to demonstrate the scaled importance while 0 stood for dimensions without any reference (See Figure 1). Based on Amy's perspective, seven dimensions were involved in total, among which pedagogical content knowledge (PCK) was more critical than others whereas assessment washback was not explicitly mentioned.

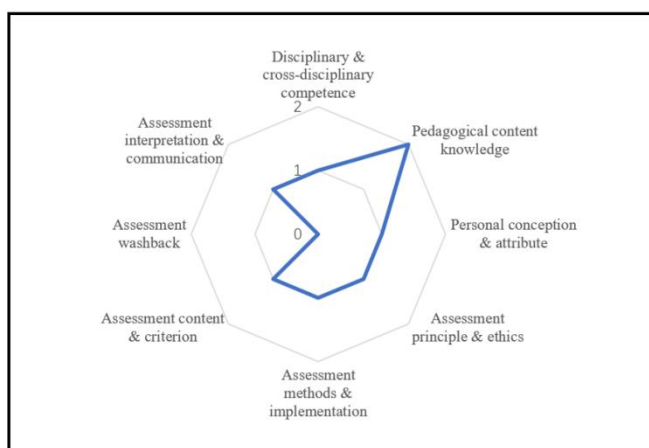


Figure 1: Conceptualization of LAL from P1 (Amy)

6.2 Conceptualization of LAL from Participant 2

In a more focused way, Betty, the second participant, defined LAL as the literacy on assessing language solely from linguistic aspects, such as grammar, pronunciation.

From literal interpretation, I feel that language assessment literacy should focus on [assessing] language. For instance, assessing the students' English competence from the accuracy of grammar, sentences or pronunciation, just within the scope of language.

She also depicted a profile of a highly literate teacher in language assessment: equipped with high English proficiency, patient with assessing a large number of students, attentive to select a proper assessment method for each student's individualized needs, and able to interpret the assessment results for subsequent analysis. Compared with the rest of the four dimensions, personal conception & attribute (being attentive) was of greater importance, serving as the basis or pre-requisite in LAL from her point of view (See Figure 2). Three were not touched on: PCK, assessment principle & ethics, and assessment washback.

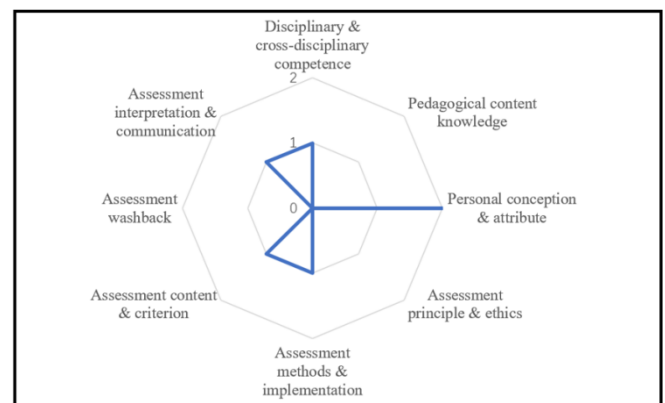


Figure 2: Conceptualization of LAL from P2 (Betty)

6.3 Conceptualization of LAL from Participant 3

Largely the same with Amy and Betty, Carol outlined LAL as the quality or competence to assess students' overall language proficiency. A qualified EFL teacher should know how to assess soundly rather than conducting assessment merely based on experience or instinct. Additionally, teachers should be aware of the potential effect of assessment imposed on students and courses. All these proposed components of LAL were equally important for EFL teachers just like "Cask effect explanation".

EFL teachers should know the possible effect of assessment on students and courses. If they are totally ignorant of the assessment-related knowledge and skills, they may either harm students or implement an unsound assessment unconsciously. Or maybe these behaviors are encouraged without any awareness...I think they [the components] are equally important without any discrimination, just like "cask effect". All of these should be advanced side by side.

From the description, it could be seen that Carol laid great emphasis on the assessment implementation, washback, ethics, and content, with no reference to other dimensions. (See Figure 3). Carol was the only participant who noticed the assessment washback, that means, the possible impact on students imposed by the assessment, either positive or negative.

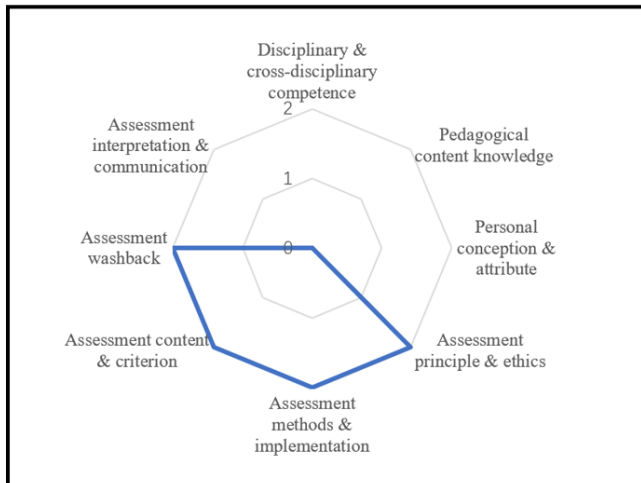


Figure 3: Conceptualization of LAL from P3 (Carol)

6.4 Conceptualization of LAL from Participant 4

Daisy, the fourth participant, conceptualized LAL more from the benefit of students as “capability of knowing how to assess students accurately and beneficially to students’ English learning”. The EFL teacher who possessed LAL should take the entire learning process and learning attitude into account rather than relying on the learning results (e.g., exam scores) exclusively. Also, the teacher should be competent in linguistic knowledge and skills apart from the ability of assessing students’ four language skills. Proficiency in English language should not lag behind. All these sub-categories were of equal importance but they were not in conflict.

My ideal EFL teacher should not be totally overwhelmed by exam scores, although in reality, this is the most common practice. Students should be assessed from the whole learning process and learning attitudes...teachers should design tasks in assessing students’ four skills to examine where the deficiency is. Moreover, the teachers’ language literacy, that is, linguistic knowledge and skills is also a must... I think they are equally important, but that does not mean they are conflictory with each other. All of them should be developed without a shortage in one aspect, only in this way, the one can become a professional EFL teacher. ...EFL teachers are suggested to notice the assessment towards students’ emotions. That means to notice the subtle changes in their emotions.

Daisy mentioned five dimensions of LAL: disciplinary competence, personal conception, assessment principle, assessment methods, and assessment content. All these elements mentioned were labelled with the same importance weight as displayed in Figure 4. Besides, there are two impressive highlights in Daisy’s conceptualization. One is she emphasized the assessment literate teacher should not solely rely on examination score to assess students. Instead, she advocated multiple assessment methods to obtain a more comprehensive picture of the whole learning process. The other one is she was the first participant who stressed the affective dimension in assessment content, which means not only the academic performance, but also the students’ attitudes and emotions should be incorporated into the assessment domain. In other words, she extended the coverage of the assessment domain to include the affective

dimension.

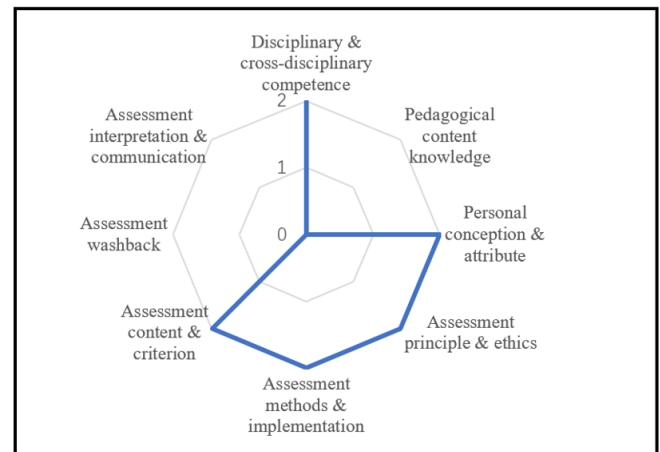


Figure 4: Conceptualization of LAL from P4 (Daisy)

6.5 Conceptualization of LAL from Participant 5

The fifth participant, Edwin, a male pre-service EFL teacher, was applying for TESOL (Teach English for Speakers of Other Languages) in England. He impressed the researcher with more technical terms in assessment when answering the questions in the interview, like diagnostic assessment, criterion-referenced exam. He conceptualized LAL from the following aspects: knowledge of assessment principle, skills, and methods; the inclusion of affective dimensions in assessment content; attitude towards assessment; way of feedback; disciplinary literacy and broad knowledge base in other related disciplines.

To my knowledge, I think you [the EFL teacher] must know some assessment methods, like group assessment, diagnostic assessment, formative assessment. Also, you should master some assessment principles, such as whether the assessment involves all students, whether it assesses various aspects of the students, including the learning attitude towards English whether the students are positive or not on English learning...then, you should also hold a positive attitude to assessment...At last, your own professional literacy should achieve a qualified level, with proficient English level and broad knowledge base not only in the discipline, but also in the related area to broaden students’ horizon and arouse their learning interests. ... You [EFL teachers] need to pay attention to the way of feedback, like oral or written feedback, direct or indirect feedback, and the timing of feedback, when to correct the mistakes made by students.

When asked about the sequence of these elements based on their significance, he provided a rather comprehensive opinion, “Ideally, these elements should be of the same significance, however, in reality, it is not the case”. He was the first participant who had pointed out the context-sensitive nature of LAL. From his observation in school, he attached greater importance to assessment methods.

As shown in Figure 5, his conceptualization covered six dimensions among which assessment methods & implementation was of utmost importance based on his own experience. All the dimensions identified by him were exposed to the social cultural context represented by the shadow in the diagram. In Edwin’s conceptualization of LAL, three salient features were obvious: the coverage of affective

dimension in assessment content, the explicit notice of the context-dependent nature of LAL, and the very emphasis on assessment methods.

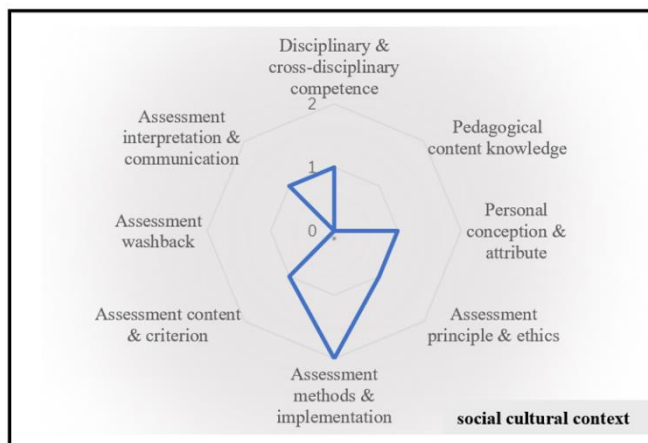


Figure 5: Conceptualization of LAL from P5 (Edwin)

6.6 Conceptualization of LAL from Participant 6

Likewise, Flora, perceived LAL as an integrated term encompassing everything related to language assessment. She also identified five dimensions of LAL with equal importance: English proficiency & linguistic competence, assessment implementation, assessment ethics, assessment criterion, and feedback communication.

At first, I feel like linguistic knowledge and related skills must be necessary and proficiency in English is also essential for the EFL teacher should be knowledgeable to teach...secondly, the teacher should be skilled in organizing and implementing assessment activities... at last, students' psychological state should also be given enough attention...meanwhile, the teacher is advised to deliver the assessment results or feedback in an appropriate way for students to better absorb. That is a big challenge for teacher to provide feedback in a more acceptable way for students. ...The EFL teacher is supposed to know the assessment criterion: how to distinguish the student with better academic achievement.

Flora also emphasized LAL's context-dependent nature by explaining the shaping effect of high-stake examinations thrust on LAL structure. Ideally, an EFL teacher should be competent in assessing the four linguistic skills, nonetheless, the teacher's LAL on assessing oral English was deficient to a large extent in the real society due to the absence of testing spoken English in most large-scale exams in China. Thus, LAL's construct was largely shaped by the construct and the corresponding weight of influential exams. As demonstrated in Figure 6, the background of LAL was shadowed to signify the shaping effect of the social cultural context. The five dimensions figured out by Flora were assigned with equivalent importance.

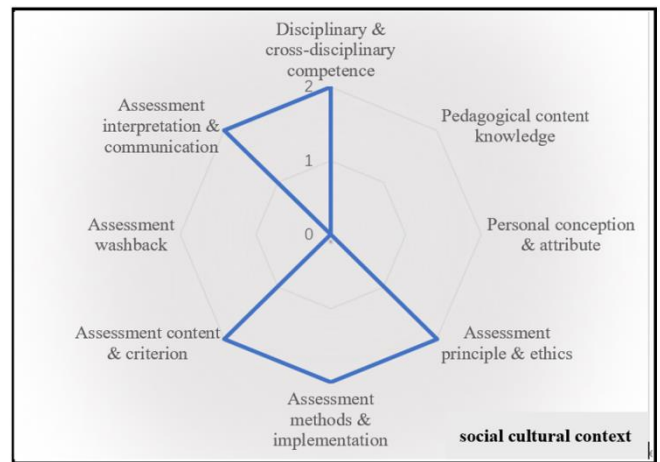


Figure 6: Conceptualization of LAL from P6 (Flora)

6.7 Conceptualization of LAL from All Participants

All of the six participants seemed to reach a consensus that LAL was a quality or competence of EFL teachers to assess students' English proficiency to facilitate learning by adopting various methods, though varying emphasis was laid on different dimensions in their references. The eight dimensions are different in frequency of being mentioned by the participants. Pedagogical content knowledge and assessment washback are least frequently mentioned except by only one participant. It means the six participants least focus on these two dimensions. By contrast, another two dimensions are covered by six participants though with varying importance: assessment methods & implementation and assessment content & criterion, which are considered by all the participants to be the very core component of LAL.

For a more vivid presentation, the overall conceptualization of LAL from the six participants was summarised. As shown in Figure 7, the synthesized conceptualization of LAL from the six pre-service EFL teachers in China was characterized by four salient features: multi-dimension, graded importance, context-dependence and uniqueness.

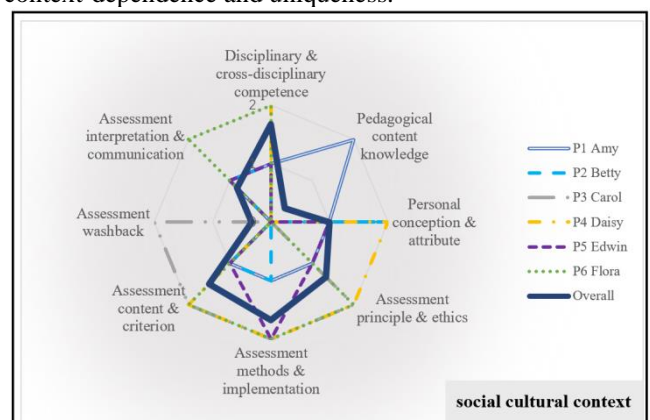


Figure 7: The combined overall conceptualization of LAL

1) LAL was perceived as a multi-layered concept, encompassing eight dimensions, which could be roughly classified into three broad domains. Discipline-related competence (disciplinary & cross disciplinary competence, pedagogical content knowledge (PCK)), person-related competence (personal conception & attribute), and assessment-related competence (assessment principle & ethics, methods & implementation, content & criterion, washback, interpretation & communication).

2) LAL's constructs were not equally important. The scaled importance from the most important to the least was: Disciplinary & cross-disciplinary competence and Assessment methods & implementation > Assessment content & criterion > Assessment principle & ethics > Personal conception & attribute > Assessment interpretation & communication > PCK and Assessment washback. The participants in the present study attached considerable importance to disciplinary competence and practical implementation of assessment methods while they tended to give scarce attention to PCK and the potential influence of assessment.

3) LAL was dependent on the social cultural context. The constructs of LAL were shaped by the context where LAL was embedded. They were oriented explicitly or implicitly by the social cultural context, including the macro exam culture and the micro teaching context, which were highlighted by two of the participants (See Edwin and Flora). Thus, LAL was sensitive to the local context and was printed with the distinctive feature of the context.

4) Each participant's spider diagram of LAL conceptualization was unique and different from others. None of them was identical.

7. Discussions

1) multi-dimension: The conceptualization of LAL from the participants' perspective overlapped greatly with the existing frameworks derived from other stakeholders. The participants in the study perceive LAL as multi-dimensional which is widely supported by the literature (Baker & Riches, 2017; Bøhn & Tsagari, 2021; Davies, 2008; Fulcher, 2012; Kremmel & Harding, 2020; Taylor, 2013). Most of the dimensions identified by the EFL teacher candidates are greatly overlapped with the previous models.

Nonetheless, a few dimensions mentioned in the present study are seldom referred to in other models, and vice versa. On the one hand, what the pre-service EFL teachers has included but not mentioned in prior models lies in two aspects: (1) cross-disciplinary competence (knowledge in other related discipline) and (2) personal attribute (patient, responsible, attentive). Few models in the literature seem to take the knowledge in related area and teacher's own characters into account. This is a new contribution to the existing literature through enriching the conceptualization of LAL from the neglected stakeholders, pre-service EFL teachers.

On the other hand, what the pre-service EFL teachers has not paid attention to but mentioned in prior models exists in two aspects: assessment policies and local practices (Baker & Riches, 2017; Bøhn & Tsagari, 2021; Kremmel & Harding,

2020; Taylor, 2013); and collaboration (Baker & Riches, 2017; Bøhn & Tsagari, 2021). One of the possible reasons maybe that these pre-service EFL teachers have few opportunities to implement or be involved in-depth in real assessment practices in school contexts even in teaching practicum, where they are just assigned the task of scoring the objective-tasks on most occasions. Therefore, they appear to be not so concerned about the practical factors in conducting an assessment such as assessment policies, local practices in schools and cooperation in the assessment.

2) priority: The pre-service EFL teachers' conceptualization in the present study share something in common with other stakeholders' (language teachers, teacher educators) perception of the priority in LAL, but interestingly, they have their own extraordinary classification in nonpriority along the LAL priority continuum.

On the most important extreme, the present conclusion is mostly similar to Kremmel and Harding (2020) and Bøhn and Tsagari (2021), whose studies also highlight discipline-related competence. This competence is considered by the participants as the most important requirement for EFL teachers' LAL. Meanwhile, the mastery of assessment methods and implementation of the assessment are of the same importance. Likewise, assessment in language pedagogy is also labelled as utmost important in Kremmel and Harding (2020)'s research. That means, the pre-service EFL teachers in China share roughly the same understanding of priority in the LAL conceptualization with language teachers in UK (See Kremmel & Harding, 2020) and teacher educators in Norway (See Bøhn & Tsagari, 2021).

On the contrary, the other extreme presents a totally different picture. On the least important extreme, the PCK and assessment washback summarised from the study are seldom classified as the comparatively less important dimensions in other frameworks, which consistently take theoretical knowledge in assessment as the less demanding requirement for language teachers' LAL. This discrepant classification may be explained by the well-acknowledged divergence among different stakeholders' conceptualization of LAL and the variance in each research context which LAL depends on.

3) context-dependence: The third prominent feature, context-dependence, means LAL is mediated by the contexts where it is rooted. This dependence has been agreed by a wide range of studies in the literature (Baker & Riches, 2017; Crusan et al., 2016; Davidson & Coombe, 2019; Farhady & Tavassoli, 2018; Han & Kaya, 2014; Jia et al., 2006; Portelli & O'Sullivan, 2016; Sultana, 2019). Instead of treating LAL in a vacuum, the pre-service EFL teachers in China has noticed the influential impact imposed by the social cultural contexts on LAL constructs.

4) uniqueness: The very core meaning of LAL conceptualization among the participants is rather similar whereas the peripheral aspects of the conceptualization tend to be more characterized by their own individual experiences. None of these conceptualizations of LAL is identical as displayed in Figure 7, which exhibits unique spider diagrams from each participant in the current study.

8. Conclusions and Limitations

The voices of pre-service EFL teachers have been presented to supplement the inadequate focus on this critical stakeholder in LAL research scope, especially in the context of China, where LAL research is comparatively limited (Gan & Jiang, 2020). The pre-service EFL teachers' conceptualizations of LAL have been explored to provide a holistic picture of LAL from the long-neglected group. They conceptualized LAL as an eight-dimensional concept with graded importance, embedded in the social cultural contexts. Two newly-found dimensions were added by the participants: cross-disciplinary competence (knowledge in other related discipline) and personal attribute (patient, responsible, attentive). Meanwhile, they seemed to pay less due attention to assessment policies and local practices compared with the previous conceptualizations proposed in the literature.

Though enlightening in the research findings, it also has a limitation. The limitation comes from the recruitment of only one male participant in this study. As revealed by previous studies (e.g., Alkharusi et al., 2012), gender may be a factor influencing the assessment knowledge and self-perception of assessment competence. Therefore, it would be much better to recruit the same number of male participants as female ones for not over-representation of female voices.

Concerning the LAL context-dependent nature, further studies are suggested to be carried out in the rest regions of China or outside China for a more comprehensive understanding of LAL in various assessment contexts. The present study was conducted in a pre-service teacher education programme situated in the western part of China, which was recognized as a less developed area with restricted assessment resources. The settings sampled in the current study may differ dramatically from teacher education programmes in other contexts, like other regions of China or other parts of the world.

Funding

This work was supported by Shaanxi Province's "14th Five-Year Plan" Education Science Planning under Grant [number SGH22Y1434]; Xianyang Normal University Teaching and Reform Project under Grant [number 2023YB55].

Statements and Declarations

Conflict of interest: The author declares no conflict of interests in current study.

Ethical approval: This study was conducted after the Institutional Review Board approval from X Normal University and obtained written consent from each participant.

Biographical note: Geng Juanjuan, PhD, associate professor in Xianyang Normal University. Her interest is in teacher education and assessment literacy.

References

- [1] Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2012). Educational assessment attitudes, competence, knowledge, and practices: An exploratory study of Muscat teachers in the Sultanate of Oman. *Journal of Education and Learning*, 1(2), 217-232. <http://dx.doi.org/10.5539/jel.v1n2p217>
- [2] Babaii, E., & Asadnia, F. (2019). A long walk to language assessment literacy: EFL teachers' reflection on language assessment research and practice. *Reflective Practice*, 20(6), 745-760. <https://doi.org/10.1080/14623943.2019.1688779>
- [3] Baker, B. A., & Riches, C. (2017). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557-581. <https://doi.org/10.1177/0265532217716732>
- [4] Bøhn, H., & Tsagari, D. (2021). Teacher educators' conceptions of language assessment literacy in Norway. *Journal of Language Teaching and Research*, 12(2), 222-233. <http://dx.doi.org/10.17507/jltr.1202.02>
- [5] Bolívar, E. M. R. (2020). Monitoring preservice teachers' language assessment literacy development through journal writing. *Malaysian Journal of ELT Research*, 17(1), 38-52.
- [6] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- [7] Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 126-136). Cambridge University Press.
- [8] Butler, Y. G., Peng, X., & Lee, J. (2021). Young learners' voices: Towards a learner-centered approach to understanding language assessment literacy. *Language Testing*, 38(3), 429-455. <https://doi.org/10.1177/0265532221992274>
- [9] Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Sage.
- [10] Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.
- [11] Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43-56. <https://doi.org/10.1016/j.asw.2016.03.001>
- [12] Davidson, P., & Coombe, C. (2019). Language assessment literacy in the Middle East and North Africa (MENA) region. *Arab Journal of Applied Linguistics*, 4(2), 1-23.
- [13] Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-348. <https://doi.org/10.1177/0265532208090156>
- [14] deMarrais, K. (2004). Qualitative interview studies: Learning through experience. In K. deMarrais & S. D. Lapan (Eds.), *Foundations for research: Methods of inquiry in education and the social sciences* (pp. 51-68). Lawrence Erlbaum Associates.

- [15] Drever, E. (1995). *Using semi-structured interviews in small-scale research: A teacher's guide*. Scottish Council for Research in Education.
- [16] Esin, C., Fathi, M., & Squire, C. (2014). Narrative analysis: The constructionist approach. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 203-216). Sage. <https://dx.doi.org/10.4135/9781446282243.n14>
- [17] Farhady, H., & Tavassoli, K. (2018). Developing a language assessment knowledge test for EFL teachers: A data-driven approach. *Iranian Journal of Language Teaching Research*, 6(3), 79-94.
- [18] Fink, A. S. (2000). The role of the researcher in the qualitative research process: A potential barrier to archiving qualitative data. *Forum: Qualitative Social Research*, 1(3), Article 4. <http://nbn-resolving.de/urn:nbn:de:0114-fqs000344>
- [19] Firoozi, T., Razavipour, K., & Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia*, 9(1), 1-14. <https://doi.org/10.1186/s40468-019-0078-7>
- [20] Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132. <https://doi.org/10.1080/15434303.2011.642041>
- [21] Gan, L., & Jiang, C. (2020). Jiaoshi ceping suyang guowai yanjiu shuping [Literature review of research on teacher assessment literacy abroad]. *Zhongguo kaoshi*, (1), 72-78.
- [22] Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile Issues in Teachers Professional Development*, 20(1), 179-195. <https://dx.doi.org/10.15446/profile.v20n1.62089>
- [23] Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education and Literacy Studies*, 3(4), 42-48. <http://dx.doi.org/10.7575/aiac.ijels.v.3n.4p.42>
- [24] Han, T., & Kaya, H. I. (2014). Turkish EFL teachers' assessment preferences and practices in the context of constructivist instruction. *Journal of Studies in Education*, 4(1), 77-93. <https://doi.org/10.5296/jse.v4i1.4873>
- [25] Hildén, R., & Fröjdendahl, B. (2018). The dawn of assessment literacy—exploring the conceptions of Finnish student teachers in foreign languages. *Language*, 12(1), 1-24. <https://doi.org/10.17011/apples/urn.201802201542>
- [26] Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402. <https://doi.org/10.1177/0265532208090158>
- [27] Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment* (3rd ed., pp. 257-270). Springer. https://doi.org/10.1007/978-3-319-02261-1_19
- [28] Jia, Y., Eslami, Z. R., & Burlbaw, L. M. (2006). ESL teachers' perceptions and factors influencing their use of classroom-based reading assessment. *Bilingual Research Journal*, 30(2), 407-430. <https://doi.org/10.1080/15235882.2006.10162883>
- [29] Jin, Y. (2018). Waiyu jiaoshi pingjia suyang fazhan: lilun kuangjia he lujing tansuo [Foreign language teachers' assessment literacy development: Theoretical framework and routine exploration]. *Waiyu jiaoyu yanjiu qianyan*, (2), 65-72.
- [30] Kavakh, N., & Arslan, S. (2019). Towards a continuum from know-how to show-how for developing EFL student-teachers' assessment literacy. *International Online Journal of Education and Teaching*, 6(1), 223-232. <http://www.iojet.org/index.php/IOJET/article/view/593>
- [31] Kim, A. A., Chapman, M., Kondo, A., & Wilmes, C. (2020). Examining the assessment literacy required for interpreting score reports: A focus on educators of K-12 English learners. *Language Testing*, 37(1):54-75. <https://doi.org/10.1177/0265532219859881>
- [32] 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. <https://doi.org/10.1080/15434303.2019.1674855>
- [33] Nimehchisalema, V., & Bhattib, N. (2019). A review of literature on language assessment literacy in last two decades (1999-2018). *Assessment*, 8(11), 44-59.
- [34] 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>
- [35] Polkinghorne, D. E. (2007). Validity issues in narrative research. *Qualitative Inquiry*, 13(4), 471-486. <https://doi.org/10.1177/1077800406297670>
- [36] Portelli, L., & O'Sullivan, K. A. (2016). Teachers' perceptions of the influence of assessment on their teaching of year 9 English. *English in Australia*, 51(1), 71-80.
- [37] Riessman, C. K. (2005) Narrative analysis. In N. Kelly, C. Horrocks, K. Milnes, B. Roberts, & D. Robinson (Eds.), *Narrative, memory and everyday life* (pp. 1-7). University of Huddersfield Press. <http://eprints.hud.ac.uk/id/eprint/4920>
- [38] Shahzamani, M., & Tahririan, M. H. (2021). Iranian medical ESP practitioners' reading comprehension assessment literacy: Perceptions and practices. *Iranian Journal of English for Academic Purposes*, 10(1), 1-15.
- [39] Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Working Papers in Applied Linguistics & TESOL*, 18(1), 1-22. <https://doi.org/10.7916/salt.v18i1.1195>
- [40] Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534-39.
- [41] Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1-14. <https://doi.org/10.1186/s40468-019-0077-8>
- [42] Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412. <https://doi.org/10.1177/0265532213480338>
- [43] Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, 15(3), 398-405. <https://doi.org/10.1111/nhs.12048>

- [44] Van Nes, F., Abma, T., Jonsson, H., & Deeg, D. (2010). Language differences in qualitative research: is meaning lost in translation?. *European Journal of Ageing*, 7(4), 313-316. <https://doi.org/10.1007/s10433-010-0168-y>
- [45] Yan, X., & Fan, J. (2020). "Am I qualified to be a language tester?": Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing*, 1-28. <https://doi.org/10.1177/0265532220929924>