



A FRAMEWORK FOR EXAMINING DIGITAL ASSESSMENT LITERACY OF LANGUAGE TEACHERS: THE EXPERTS' FEEDBACK

Nguyen Quynh Hoa*

*Office of Academic and Student Affairs, VNU University of Languages and International Studies,
No. 2 Pham Van Dong, Cau Giay, Hanoi, Vietnam*

Received 15 December 2025

Revised 15 March 2026; Accepted 22 July 2026

Abstract: Assessment has long been acknowledged as a crucial part of teaching practices, which serves multiple purposes including providing feedback and enhancing learners' performance (Brown, 2004). In the current era of technology and artificial intelligence (AI), teachers have employed a wide range of online assessment platforms and tools to deliver tests, maintain interaction, and provide feedback to students. In this context, teachers' language assessment literacy increasingly encompasses digital assessment literacy. Meanwhile, the term "digital assessment literacy" has not been clearly and consistently defined. In the existing literature, "digital assessment literacy" is often separated into "assessment literacy" and "digital literacy" or is defined through a simple or mechanical combination of these two literacies. This study proposes a conceptual framework for examining the digital assessment literacy (DAL) of language teachers and conducts an exploratory content validation through expert consultation. The framework is constructed on the basis of related models of assessment literacy and digital literacy (Davies, 2008; Fulcher, 2012; Pill & Harding, 2013; Taylor, 2013; UNESCO, 2018). Experts in the field of education, including those specialized in assessment and technology, were invited to provide feedback on the proposed framework, specifically its components and sub-components. Feedback on the plausibility and content relevance of the framework as well as other specific comments relating to necessary additions, omissions, or revisions were taken into consideration. The refined framework may serve as a conceptual foundation for further research on stakeholders' perceptions of language teachers' digital assessment literacy and its impact on teaching and assessment practices.

Keywords: assessment literacy, digital literacy, digital assessment literacy, conceptual framework, experts' feedback

* Corresponding author.

Email address: quynhhoa84@gmail.com

<https://doi.org/10.63023/2525-2445/jfs.ulis.5696>

KHUNG ĐÁNH GIÁ NĂNG LỰC KIỂM TRA ĐÁNH GIÁ SỐ CỦA GIẢNG VIÊN NGOẠI NGỮ: PHẢN HỒI TỪ CÁC CHUYÊN GIA

Nguyễn Quỳnh Hoa

*Phòng Đào tạo và Người học, Trường Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội,
Số 2 Phạm Văn Đồng, Cầu Giấy, Hà Nội, Việt Nam*

Nhận bài ngày 15 tháng 12 năm 2025

Chỉnh sửa ngày 15 tháng 3 năm 2026; Chấp nhận đăng ngày 22 tháng 7 năm 2026

Tóm tắt: Kiểm tra đánh giá từ lâu đã được công nhận là một phần quan trọng trong các hoạt động giảng dạy, nhằm phục vụ nhiều mục đích khác nhau bao gồm cung cấp phản hồi và nâng cao hiệu quả học tập (Brown, 2004). Trong thời đại công nghệ và trí tuệ nhân tạo (AI) hiện nay, giảng viên đã và đang sử dụng đa dạng các nền tảng và công cụ đánh giá trực tuyến để triển khai các bài kiểm tra, duy trì tương tác và đưa ra phản hồi cho người học. Trong bối cảnh đó, năng lực kiểm tra đánh giá của giảng viên ngoại ngữ dần bao hàm cả năng lực kiểm tra đánh giá số. Tuy nhiên, thuật ngữ "năng lực kiểm tra đánh giá số" chưa được định nghĩa một cách rõ ràng, thống nhất. Trên thực tế, "năng lực kiểm tra đánh giá số" thường được tách thành hai năng lực riêng biệt là: "năng lực kiểm tra đánh giá" và "năng lực số", hoặc được định nghĩa thông qua việc kết hợp một cách cơ học hai năng lực trên. Nghiên cứu này đề xuất một khung khái niệm nhằm xem xét năng lực đánh giá số của giảng viên ngoại ngữ và tiến hành đánh giá giá trị nội dung của khung ở mức độ thăm dò thông qua tham vấn chuyên gia. Khung được xây dựng trên cơ sở các mô hình liên quan về năng lực kiểm tra đánh giá và năng lực số (Davies, 2008; Fulcher, 2012; Pill & Harding, 2013; Taylor, 2013; UNESCO, 2018). Các chuyên gia trong lĩnh vực giáo dục, bao gồm cả chuyên gia trong lĩnh vực kiểm tra đánh giá và lĩnh vực công nghệ, được yêu cầu đưa ra phản hồi về khung được đề xuất, cụ thể là các thành tố chính và thành phần cấu thành của năng lực kiểm tra đánh giá số. Những phản hồi liên quan đến tính hợp lý và mức độ phù hợp về nội dung của khung, cũng như các ý kiến cụ thể về những nội dung cần bổ sung, lược bỏ hoặc điều chỉnh, đã được xem xét. Khung được điều chỉnh có thể được sử dụng làm khung khái niệm cho các nghiên cứu sâu hơn về nhận thức của các bên liên quan đối với năng lực kiểm tra đánh giá số của giảng viên ngoại ngữ và tác động của nó đến thực tiễn giảng dạy và kiểm tra đánh giá.

Từ khóa: năng lực kiểm tra đánh giá, năng lực số, năng lực kiểm tra đánh giá số, khung khái niệm, phản hồi của chuyên gia

1. Introduction

Assessment has long been acknowledged as an essential component of classroom practice. Brown (2004) argued that assessment could be used to identify students' strengths and weaknesses. Similarly, Rea-Dickins (2006) emphasized the crucial role of assessment as it helped identify the gap between learners' current performance and the learning target. Based on such information, educators could provide specialized academic support to individual learners. Accordingly, teachers are supposed to have adequate assessment literacy. In fact, the importance of developing assessment literacy among teachers has been widely highlighted (Rad, 2019; Coombe, Troudi & Al-Hamly, 2012).

During and after the emergence of COVID-19, the implementation of teaching and learning activities on various online platforms has increased significantly worldwide. In Vietnam, popular teaching platforms include Zoom, Microsoft Teams, Google Classroom,

Skype, and digital learning tools such as Canva, Moodle (Thai et al., 2023). Particularly, in the field of assessment, a variety of online assessment instruments are employed, among which tools like Google Forms, Quizizz, and Kahoot are commonly used. Teachers apply these tools to deliver tests, maintain interaction, and provide feedback to students. In this context, teachers' assessment literacy increasingly encompasses digital assessment literacy (DAL).

As an important form of literacy in modern education, DAL needs to be thoroughly and clearly defined and then used as a theoretical foundation for establishing policies, designing training courses, and guiding teaching and learning practices. However, there has been no clear definition of the term DAL. In practice, assessment literacy and digital literacy are defined either separately or combined in a rather mechanical way. To address this gap, the present study proposes a framework for examining the DAL of language teachers through a review of related theories and frameworks. Moreover, to evaluate the plausibility and content relevance of the proposed framework, experts in assessment and/or technology were invited to provide feedback.

Specifically, this research aims to answer two research questions:

1. *How do the experts evaluate the plausibility of the proposed framework, especially the relevance, representativeness, clarity and distinctiveness of its components and sub-components?*
2. *What changes are necessary to generate a more understandable and workable framework?*

By proposing and initially evaluating a conceptual framework of DAL, this study contributes to the theoretical clarification of the construct and provides a potential foundation for future empirical research and teacher professional development.

2. Literature Review

2.1. Assessment Literacy

The American Federation of Teachers (1990) provided the earliest conceptualization of assessment literacy for teachers, although the term “assessment literacy” was not used directly at the time. Instead, a list of seven assessment competencies in which teachers should be skilled was proposed, including choosing appropriate assessment methods; developing appropriate assessment instruments; administering, scoring, and interpreting the results; using assessment results when making decisions; developing valid grading procedures; communicating assessment results to stakeholders; and recognizing unethical, illegal, and inappropriate assessment methods and uses of assessment information.

In 1991, Stiggins was the first author to coin the term “assessment literacy”. He identified an assessment literate person as one who could identify quality assessments and critically evaluate assessment data to make inferences about learners' achievement. In other words, the author focused on two competencies – evaluating assessment tools and interpreting assessment data.

Since then, several models have been proposed to conceptualize assessment literacy or language assessment literacy, the narrower term. Davies (2008) applied a “skills and knowledge” approach to define assessment literacy. “Skills” referred to the practical know-how in test analysis and construction, while “knowledge” meant the “relevant background in measurement and language description” (Davies, 2008, p. 328). Such an approach results in a broader definition compared to that of Stiggins (1991) as it consisted not only of practical knowledge and skills but also focused on “principles”, i.e. the reasons for assessment.

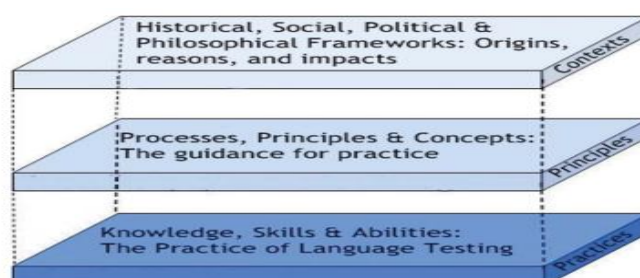
Fulcher (2012) expanded the definition even further and proposed a *Practices – Principles – Contexts* model. He defined assessment literacy as:

the knowledge, skills, and abilities required to design, develop, maintain, or evaluate large-scale standardized and/or classroom-based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. The ability to place knowledge, skills, processes, principles and concepts within wider historical, social, political and philosophical frameworks in order to understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals. (Fulcher, 2012, p. 125)

However, the author did not believe that the acquisition of the listed components was necessary for all stakeholders. Classroom teachers may focus on *Practices*, whereas *Principles* or *Contexts* may be more important to test writers, and university administrators. Besides, it is likely that different stakeholders need to achieve a different level of mastery. His model is illustrated in the following figure:

Figure 1

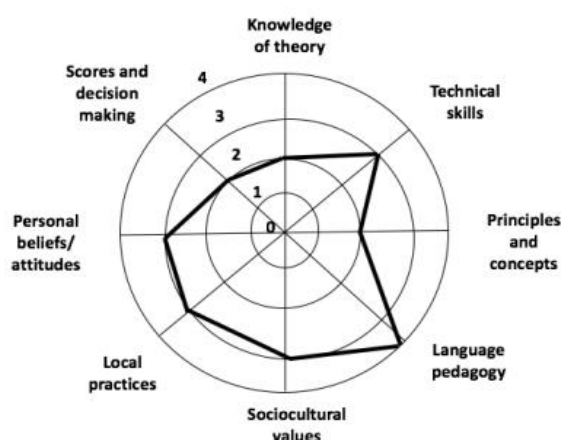
Fulcher's Model of Language Assessment Literacy



Taylor (2013) also created an outstanding model. He viewed language assessment literacy as comprising eight components – knowledge of theory, technical skills, principles and concepts, language pedagogy, sociocultural values, local practices, personal beliefs/attitudes, and scores and decision making – and mapped these components onto Pill and Harding's (2013) five levels to differentiate the language assessment literacy needs for different stakeholders. The following figure is Taylor's (2013, p. 410) language assessment literacy profile for classroom teachers.

Figure 2

Taylor's Language Assessment Literacy Profile for Classroom Teachers



The listed components in Taylor's (2013) model shared great similarity to the sub-components mentioned in Fulcher's (2012) model. For example, *Technical Skills*, *Principles*

and concepts, or Sociocultural values can be grouped under *Practices – Principles – Contexts* respectively. However, there was a new component which was added – *Personal beliefs and attitudes*. It can be understood as how stakeholders' preconceptions, viewpoints, and opinions may "inform [their] conceptualizations, interpretations, judgments, and decisions in assessment" (Scarino, 2013, p. 109).

Among the above-mentioned models, that of Fulcher (2012) provides a comprehensive overview of language assessment literacy. Specifically, it aligns with the common agreement that practices should be based on principles and theories, and skills and knowledge should be placed in specific contexts. However, the model does not specify the sub-components, which may lead to confusion such as the distinction between *Skills* and *Abilities*. Moreover, it does not include an influential component, namely *Personal beliefs and attitudes*, as in that of Taylor's (2013). This component may have a significant impact on teaching practices. Nelson and Guerra (2014, p. 70) claimed that beliefs were "powerful filters that shape how an individual sees the world, sees other people, and sees oneself". As a result, the component is worth considering as a key component within the *Principles* component. Another notable aspect of Fulcher's (2012) model is that the *Historical, Political, Social and Philosophical frameworks* in the *Contexts* component mainly refer to macro-level contexts, while overlooking the micro-level contexts such as *Local practices* in Taylor's (2013) model.

2.2. Digital Literacy

Digital literacy has been conceptualized in various ways in literature. Martin (2006) proposed a European framework of digital literacy in which he formulated a detailed definition. His definition put great emphasis on awareness and attitude as well as evaluative and reflective competences among digital citizens as follows:

Digital literacy is the awareness, attitude, and ability of individuals to appropriately use digital tools and facilities to identify, access, manage, integrate, evaluate, analyse and synthesize digital resources, construct new knowledge, create media expressions, and communicate with others, in the context of specific life situations, in order to enable constructive social action; and to reflect upon this process. (Martin, 2006, p. 155)

As for Churches (2008), he applied Bloom's Taxonomy to the digital world and created Bloom's Digital Taxonomy, in which he structured digitally focused activities into six levels of thinking.

Figure 3

Bloom's Digital Taxonomy



Vuorikari et al. (2016, as cited in UNESCO, 2018) created the DigComp 2.0 framework, which identified five competence areas of digital literacy and the competences within each area. The framework is presented in Table 1.

Table 1

The Digcomp 2.0 Framework

Competence area	Competences
1. Information and data literacy	1.1 Browsing, searching and filtering data, information and digital content 1.2 Evaluating data, information and digital content 1.3 Managing data, information and digital content
2. Communication and collaboration	2.1 Interacting through digital technologies 2.2 Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.5 Netiquette 2.6 Managing digital identity
3. Digital content creation	3.1 Developing digital content 3.2 Integrating and re-elaborating digital content 3.3 Copyright and licenses 3.4 Programming
4. Safety	4.1 Protecting devices 4.2 Protecting personal data and privacy 4.3 Protecting health and well-being 4.4 Protecting the environment
5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.3 Creatively using digital technologies 5.4 Identifying digital competence gaps

An updated version of this framework was proposed by UNESCO (2018) in which *Devices and software operations*, *Career-related competences*, and *Computational thinking* were added. The addition of *Career-related competences* reflects the need to situate digital literacy within specific professional fields such as education in general and testing and assessment more specifically.

Based on the review and synthesis of the above-mentioned models of assessment literacy and digital literacy, it can be observed that there is a great resemblance in the ways these two terms have been conceptualized. Both literacies are often conceptualized as areas of competence. Besides, componential skills and knowledge can be organized into several levels. In both groups of frameworks, such factors as awareness, beliefs, and attitudes are emphasized along with the need to place assessment or digital activities into specific contexts. However, there has been no clear definition for DAL. The existing models generally address assessment literacy and digital literacy separately or combine them only in a rather mechanical manner, without offering a coherent framework that captures the integrated nature of DAL.

To address this gap, the present study proposes a conceptual framework of DAL adapted from Fulcher's (2012) *Practices – Principles – Contexts* model. The proposed framework consists of three main components and nine sub-components that integrate both assessment-related and digital dimensions. Each component and sub-component is clearly defined to enhance clarity and distinctiveness and to minimize potential overlap. The proposed framework of DAL is presented in Table 2.

Table 2*Proposed Framework of Digital Assessment Literacy (DAL)*

Components	Sub-components
1. Practices The foundational component that realizes digital assessment activities	1.1 Knowledge: Understanding of digital tools, content, and data used in assessment 1.2 Skills: Practical abilities to create, deliver, score, and analyze assessments using digital tools 1.3 Abilities: Higher-level capacities to adapt, solve problems, act ethically, and support others in digital assessment contexts
2. Principles The conceptual and theoretical component that guides digital assessment activities	2.1 Principles & Processes: Awareness and application of ethical, fair, and effective assessment practices 2.2 Concepts: Understanding of key digital assessment concepts (e.g., washback, privacy, validity) 2.3 Personal Beliefs & Attitudes: Teachers' openness, motivation, and attitudes toward using digital tools
3. Contexts The evaluative and reflective component that shapes digital assessment activities in a particular context	3.1 Historical, Political & Philosophical Context: Awareness of how policies, history, and ideas shape practice 3.2 Social & Cultural Context: Understanding of how social norms and learner backgrounds affect assessment 3.3 Institutional Context: Knowledge of school or organizational policies and digital support systems

In this framework, the *Practices* component represents the operational dimension of digital assessment literacy, referring to the knowledge, skills, and abilities required to implement digital assessment activities. The *Principles* component captures the conceptual foundations that guide assessment decisions, including theoretical understanding and teachers' beliefs and attitudes. The last component, *Contexts*, refers to the understanding of broader environments that shape digital assessment practices, including institutional, sociocultural, historical, and policy-related factors.

3. Research Method

3.1. Participants

Ten experts took part in the research. The expert panel was recruited through purposive sampling. To ensure the relevance of participants' expertise to the study, three main criteria were considered in the selection process. Specifically, the experts were required to (1) have substantial knowledge in language assessment, (2) have experience in teacher training or professional development related to language assessment and/or technological application in education, and (3) demonstrate deep understanding of educational contexts in which digital language assessment occurs.

The following section presents general information about the experts, based on the data gathered from the survey questionnaire:

Table 3*Demographic Information of Participants*

Number of years working in the language education field	5 - < 10 years: 01 10 - < 15 years: 06 15 years: 03
---	---

Current working place	A public university: 10
Place of residence	In a city in the north of Vietnam: 10
Common assessment activities	A wide range of activities, the most common ones include: - Writing test items: 06/10 - Marking students' work: 06/10 - Proofreading test items: 05/10 - Reviewing marking across different markers: 05/10 - Analyzing test results: 05/10
Certifications	- Item writer certificate: 07/10 - Speaking and Writing Examiner Certificate: 09/10 - Data analysis certificate: 05/10 - MIE Expert: 03/10
Training experience	- Trainer of courses on testing and assessment: 05/10 - Years of experience: 12-15 years - Trainer of courses on technological application in education: 03/10 - Years of experience: 7-10 years
Level of literacy in assessment field (self-reported level)	- Functional literacy (sound, basic understanding): 02/10 - Procedural and conceptual literacy (understanding and practice of central concepts): 06/10 - Multidimensional literacy (extended knowledge beyond ordinary concepts): 02/10
Level of literacy in digital field (self-reported level)	- Functional literacy (sound, basic understanding): 06/10 - Procedural and conceptual literacy (understanding and practice of central concepts): 03/10 - Multidimensional literacy (extended knowledge beyond ordinary concepts): 01/10

* *MIE Expert (Microsoft Innovative Educator Expert)* is a global award given by Microsoft to educators who are pioneers and innovators in applying technology to teaching.

The demographic information shows that experts are all experienced teachers with expertise in assessment and/or digital fields, but their strength tends to lie more in the assessment field rather than in the digital field. Specifically, five experts are often trainers for courses that aim at enhancing teachers' assessment literacy. Three experts frequently carry out digital assessment activities in their classes. Three experts are trainers for courses relating to technological and digital applications in education. With such background and experience, these experts are expected to provide valuable feedback on the proposed framework of DAL. As this study represents an exploratory phase in developing the framework, a relatively small expert panel with ten participants was considered sufficient for initial content validation.

3.2. Research Tool

To answer the two research questions, a questionnaire was designed and delivered to ten experts as mentioned above. Besides, to provide background information to experts, a PowerPoint file was created, explaining briefly the research goal, related theories, the proposed framework and task description.

The questionnaire has two main parts. The first part focuses on gathering demographic information of participants, including their years of working experience, the university where they currently work, common assessment activities that they take part in, certifications, training experience, and their self-reported level of literacy in the assessment and digital fields. Another question that is added to this part is their general opinion on whether it is necessary to define DAL on its own.

The second part – the main part of the questionnaire – targets the experts' feedback on the proposed framework of DAL. In this part, these experts were asked to give both quantitative and qualitative feedback by rating the components and sub-components in terms of their relevance, representativeness, clarity and distinctiveness; and giving detailed comments including necessary omission, addition or revision. Most of the questions in the survey questionnaire are in the form of Likert scale, ranging from 1 to 5 corresponding to the level of agreement (from (1) *Strongly disagree* to (5) *Strongly agree*). In addition, the questionnaire also has open-ended questions to exploit more detailed information from participants.

3.3. Data Collection

The data collection was carried out after obtaining consent from the participants. The questionnaire was created as a Google Form, and the link was sent via email to the selected participants together with the briefing PowerPoint file. The briefing file includes background information, research goal, related theories, the proposed framework and task description. In the task description section, the expected feedback was clarified together with short definitions of the four criteria, i.e. relevance, representativeness, clarity, and distinctiveness.

- Relevance refers to the extent to which a component or sub-component appropriately represents the construct of DAL.

- Representativeness refers to the degree to which the component or sub-component reflects key aspects of DAL.

- Clarity refers to whether the component or sub-component is clearly defined and understandable.

- Distinctiveness refers to whether the component or sub-component is conceptually separate from other components/ sub-components without substantial overlap.

The experts were asked to go through the briefing file to get background information before completing the questionnaire. This also served as a pilot phase before the questionnaire was administered on a larger scale for data collection. Therefore, in addition to answering the questionnaire, experts were invited to describe their experiences with it – including whether any items or sections were unclear, ambiguous, or difficult to respond to.

3.4. Data Analysis

All the ten experts accepted the invitation and did the survey. They were coded as E1 to E10. At this pilot phase, the number of participants is small – only ten. As a result, no data analysis software was used to process data collected from the questionnaire. Instead, the quantitative feedback was analysed using descriptive statistics. In addition, the Item-level Content Validity Index (I-CVI) was calculated as the proportion of experts rating an item as either 4 or 5 on the Likert scale, indicating agreement that components or sub-components are relevant, representative, clear, or distinct. The Scale-level Content Validity Index (S-CVI/Ave) was calculated as the average of the I-CVI values across items, indicating the overall content validity of the framework.

As for the qualitative data, the experts' detailed feedback was analysed using simple thematic categorization, in which responses were grouped into four categories: omissions, additions, revisions, and other comments.

Both quantitative and qualitative data were interpreted for initial evaluation on plausibility and content relevance of the proposed framework and notable areas for revision in the questionnaire.

4. Findings

4.1. Components of DAL

The proposed framework of DAL consists of three main components, *Practices* – *Principles* and *Contexts*. Generally, experts considered the framework to be comprehensive (i.e. is complete and includes everything that is necessary) as eight out of the ten experts chose the option “All important components are fully covered”. Only one chose “Most important components are included” and one chose “Some important components are present, but not all”.

The following table shows the data collected through the questions in which experts evaluated three proposed components in terms of their relevance, representativeness, clarity and distinctiveness.

Table 4

Experts’ Rating Given to Three Components

Component/ criteria	The component is relevant – appropriate to represent DAL		The component is representative – reflecting the key aspects of DAL		The component is clear – precise and understandable		The component is distinct – not overlapping with other components	
	Mean	I-CVI	Mean	I-CVI	Mean	I-CVI	Mean	I-CVI
Practices	4.286	0.7	4.286	0.9	4.429	0.7	4.429	0.9
Principles	4.286	0.7	4.143	0.7	4.286	0.7	4.429	0.9
Contexts	4.571	1.0	4.714	1.0	4.714	1.0	4.857	1.0

The quantitative data showed that three proposed components taken from Fulcher’s (2012) model were highly evaluated as crucial components of DAL, especially *Contexts* with the I-CVI value of 1.0, indicating that all experts considered the component relevant, representative, clear and sufficiently distinct. *Principles* is the component that may require reconsideration in terms of its relevance, representativeness and clarity as its I-CVI values for these criteria were 0.7. According to Polit and Beck (2006), for a panel of ten experts, an I-CVI value of 0.78 or higher is generally considered acceptable. In the open-ended question, E5 expressed the feeling that *Practices* and *Principles* were somehow overlapping and should be distinguished more clearly.

4.2. Sub-Components of DAL

Each component in the proposed framework has three sub-components. Accordingly, there are nine sub-components. Generally, experts thought that the framework is rather comprehensive as four in ten participants chose the option “All important sub-components are fully covered”, four experts chose “Most important sub-components are included” and two experts chose “Some important sub-components are present, but not all”.

Like the above rating, experts evaluated nine proposed sub-components in terms of their relevance, representativeness, clarity and distinctiveness by choosing options from (1) *Strongly disagree* to (5) *Strongly agree*.

Table 5*Experts' Rating Given to Nine Sub-Components*

Sub-component/ criteria	The sub- component is relevant – appropriate to represent DAL		The sub- component is representative – reflecting the key aspect of DAL		The sub- component is clear – precise and understandable		The sub- component is distinct – not overlapping with other components	
	Mean	I-CVI	Mean	I-CVI	Mean	I-CVI	Mean	I-CVI
Skills	4.571	0.9	4.429	0.7	4.286	0.9	4.143	0.9
Knowledge	4.571	0.9	4.143	0.9	4.429	0.7	4.286	0.7
Abilities	4.571	0.9	4.429	0.9	4.429	0.7	4.286	0.9
Principles & Processes	4.429	0.7	4.286	0.7	4.286	0.7	4.143	0.7
Concepts	4.429	0.7	4.000	0.6	4.286	0.7	4.286	0.9
Personal Beliefs & Attitudes	4.000	0.7	3.857	0.7	4.286	0.9	4.286	0.9
Historical, Political & Philosophical Context	4.714	1.0	4.714	1.0	4.571	1.0	4.857	1.0
Social & Cultural Context	4.714	1.0	4.714	1.0	4.714	1.0	4.857	1.0
Institutional Context	4.714	1.0	4.714	1.0	4.714	1.0	4.714	1.0
Mean	4.524	0.9	4.365	0.8	4.445	0.8	4.429	0.9

The collected data showed a rather positive evaluation from experts as the average rating for all sub-components was 4.44 out of 5 on the Likert scale. The I-CVI values for the framework sub-components ranged from 0.6 to 1.0. Such values show that a large proportion of experts considered the sub-components as relevant, representative, clear, and distinct. Most items achieved I-CVI values above the commonly recommended threshold of 0.78 for panels of ten experts. At the scale level, the S-CVI/Ave was 0.86, suggesting good overall content validity of the proposed framework. However, the research results also highlighted several sub-components that may need further refinement.

Noticeably, *Principles & Processes* got the I-CVI values of 0.7 for all four criteria, which indicated that certain experts questioned the relevance, representativeness, and clarity of this sub-component. Moreover, this may be overlapping with other sub-components. Meanwhile, *Concepts* received the lowest I-CVI values for representativeness (0.6) and relevance (0.7). The other two sub-components had the I-CVI values of 0.7 for two over four criteria are *Knowledge* and *Personal beliefs & Attitudes*. In contrast, all three sub-components of Contexts achieved the I-CVI values of 1.0, indicating full agreement among the experts.

In addition to quantitative data, valuable qualitative data were also given to the proposed framework. No comment fell under *Omission* group. In *Addition* group, E3 proposed that “caution” should be added to *Personal Beliefs and Attitudes*. It’s reasonable as online tools and applications must be used with prudence to minimize adverse effects. As for E10, she suggested adding “digital culture” to *Historical, Political & Philosophical Context* and “professional development support” to *Institutional Context*, which meaningfully accomplished the sub-components. In *Revision* group, E5 suggested changing *Principles & Processes* to *Operational*

& *Ethical Principles* to better fit in the *Principles* component.

As for *Other comments*, E8 wrote that three sub-components in *Practices* could be overlapping to some extent. This issue may be resolved by suggestions given by E1 “A list of indicators in the form of “Can do” statements should be created to clarify each sub-component.” Another considerable comment from E4 was that the proposed framework should be written in “simplified language” and “examples should be added to make the framework more understandable”, especially for teachers as practitioners. Besides, both E2 and E3 wondered why *Personal Beliefs and Attitudes* was placed under *Principles* component, which may explain why this sub-component got the low rating for both relevance and representativeness as discussed above. In their opinion, this sub-component should be either placed under *Contexts* or stand on its own as the fourth component. As for E7, she wondered why *Principles* and *Processes* were combined into one sub-component.

Despite thoughtful feedback, most experts claimed that it was challenging to give rating or comments on the proposed framework of DAL as it was conceptually complex. Two participants even expressed modest concerns about their level of expertise (E6 and E9). Such concerns raise the issue of either reconsidering the criteria for selecting experts or revising the research tool to better exploit the selected experts. The possible solutions can be simplifying the framework as suggested by E4, creating a list of Can-do statements as suggested by E1 or adopting a bottom-up approach instead of top-down, i.e. the participants can be English lecturers who clarify what they need to carry out online assessment. Such information can be the input to build up the framework of DAL.

5. Discussions

The study represents an initial step in evaluating the proposed framework of DAL through expert consultation and identifying necessary changes for refinement. After related theories and frameworks were reviewed, a framework was proposed with three main components, namely *Practices*, *Principles* and *Contexts*. Each component in turn contains three sub-components that reflect the construct of DAL. This proposed framework adopts and adapts the model used by Fulcher (2012). Specifically, the sub-component *Personal Beliefs and Attitudes* is added to the second component *Principles*, while the third component *Contexts* is extended to cover both macro-, meso-, and micro-level contexts. Besides, digital factors drawn from the Digital Literacy Global Framework are interwoven to transfer language assessment literacy into DAL.

The experts’ opinions as mentioned in the Findings section above reveal that the proposed framework of DAL appears conceptually relevant but requires further refinement. Certain components and sub-components may overlap to some extent. As for three components, experts pointed out that *Practices* and *Principles* should be distinguished more clearly. In terms of nine sub-components, three sub-components of *Practices* seem to be the hardest to separate. The sub-component that attracted the greatest concern is *Personal Beliefs and Attitudes*, the one which was added to Fulcher’s (2012) original model. The research results show that relevance and representativeness of this sub-component should be either reconsidered or more clearly justified to be included in the proposed framework.

Regarding the detailed comments on the proposed framework, certain additions and revisions should be considered. For example, *Principles and Processes* should be reworded or separated into two sub-components. However, the most common feedback was that the framework should be explained in simplified language or through specific indicators in the form

of “Can do” statements so that it is more understandable to readers. These operational indicators are associated with each sub-component and specify observable competencies among language teachers (e.g., Can design digital assessment tasks using appropriate online platforms), thereby facilitating self-assessment and supporting the design of professional development programs.

However, the results of this research must be considered with caution. The study used a relatively small sample (ten participants) and represented a pilot phase of the research. The collected data and interpreted information may signal certain areas or issues that should be reconsidered both in the proposed framework and in the research process.

As mentioned above, an additional aim of the research was to pilot the questionnaire. The concern from two experts about whether they were really “experts” and whether they were able to provide valuable feedback can be a warning to the researcher. Certainly, it would be ideal to find experts who have expertise in both assessment and digital fields. However, the number of such experts is limited, or the worst scenario is that there are no real experts in both fields. Selecting experts in either assessment or digital field, as in this research, may still present limitations, as the missing domain of expertise may affect the precision and scope of the feedback provided. Consequently, the researcher may need to adopt a bottom-up approach, starting from specific indicators of DAL rather than a purely top-down framework. Moreover, participants should be treated as language teachers with varied backgrounds and experiences who can share insights into what they need to implement digital assessment practices in their own contexts.

6. Conclusion

This study proposed a conceptual framework for examining DAL of language teachers through a review of related theories and frameworks on assessment literacy and digital literacy. The proposed framework adapted Fulcher’s (2012) model, consisting of three main components and nine sub-components. Experienced language teachers with expertise in assessment and/or technology were invited to provide feedback on the proposed framework of DAL. The research results show positive evaluation from participants, which suggests preliminary support to the general plausibility and content relevance of the framework. However, participants suggested that the framework should be subjected to thorough revision to better reflect DAL. Besides, a simplified version of the framework should be created with specific indicators in the form of “Can do” statements so that it is more understandable to not only experts but general language teachers.

It should be noted that the limited sample size and geographical concentration may restrict the diversity and validity of perspectives presented regarding the proposed framework. The official validation stage will necessarily involve a larger and more diverse group of participants from different institutional and regional contexts to further refine and validate the framework. The expected outcome will be the refined framework for examining DAL of language teachers. This refined framework may serve as a conceptual foundation for future empirical studies examining teachers’ DAL and for designing professional development programs in language education. Future research could also explore language teachers’ perceptions of DAL and the impact of such perceptions on teaching and learning practices.

References

- American Federation of Teachers, National Council on Measurement in Education and National Education Association. (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, 9(4), 30. <https://doi.org/10.1111/j.1745-3992.1990.tb00391.x>

- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education. <https://evelintampubolon.wordpress.com/wp-content/uploads/2016/09/h-douglas-brown-language-assessment.pdf>
- Churches, A. (2008). *Bloom's digital taxonomy*. Australian School Library Association NSW Incorporated. <http://burtonslifelearning.pbworks.com/f/BloomDigitalTaxonomy2001.pdf>
- Coombe, C., Troudi, S., & Al-Hamly, M. (2012). Foreign and second language teacher assessment literacy: Issues, challenges and recommendations. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoyloff (Eds.), *The Cambridge guide to second language assessment* (pp. 20-29). Cambridge University Press. https://books.google.com.vn/books/about/The_Cambridge_Guide_to_Second_Language_A.html?id=7STeAgAAOBAJ&redir_esc=y
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347. <https://doi.org/10.1177/0265532208090156>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132. <https://doi.org/10.1080/15434303.2011.642041>
- Martin, A. (2006). A European framework for digital literacy. *Nordic Journal of Digital Literacy*, 1(2), 151-161. <https://doi.org/10.18261/ISSN1891-943X-2006-02-06>
- Nelson, S. W., & Guerra, P. (2014). Educator beliefs and cultural knowledge: Implications for school improvement efforts. *Educational Administration Quarterly*, 50(1), 67-95. <https://doi.org/10.1177/0013161X13488595>
- 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>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489-497. <https://doi.org/10.1002/nur.20147>
- Rad, M. R. (2019). The Impact of EFL Teachers' Assessment Literacy on Their Assessment Efficiency in Classroom. *Britain International of Linguistics, Arts and Education (BloLAE) Journal*, 1(1), 9-17. <https://doi.org/10.33258/biolae.v1i1.14>
- Rea-Dickins, P. (2006). Currents and eddies in the discourse of assessment: a learning-focused interpretation. *International Journal of Applied Linguistics*, 16(2), 163-188. <https://doi.org/10.1111/j.1473-4192.2006.00112.x>
- 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. <https://doi.org/10.1177/0265532213480128>
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534-539. <https://www.scirp.org/reference/referencespapers?referenceid=3968859>
- 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>
- Thai, T., Sheehan, S., Nguyen, Q. T. N., Nguyen, T. T. P., Nguyen, Y. T. Q., Tran, H. T. T., Nguyen, C. T., Nguyen, H. Q., & Bui, S. (2023). Digital assessment literacy: English assessment practice in Vietnam. In A. L. Gokturk-Saglam & A. E. Sevgi-Sole (Eds.), *Emerging practices for online language assessment, exams, evaluation, and feedback* (pp. 162-183). IGI Global. <https://www.igi-global.com/gateway/chapter/325669>
- UNESCO Institute for Statistics. (2018). *A Global Framework of Reference on Digital Literacy Skills for Indicator 4.4.2*. UNESCO Institute for Statistics. <https://www.uis.unesco.org/sites/default/files/medias/fichiers/2025/08/ip51-global-framework-reference-digital-literacy-skills-2018-en.pdf>

APPENDIX

