v. 26, n. 2, e63855, 2026
ISSN: 1984-6398 · DOI: 10.1590/1984-6398202663855

Revista Brasileira de Linguística Aplicada

Article

Analysis of a Tailor-Made Language Assessment Course: Perspectives from Course Participants and Instructors

Análise de um curso de avaliação de línguas sob medida: perspectivas de participantes e formadores

Frank Giraldo*
*Universidad de Caldas | Manizales | COL
frank.giraldo@ucaldas.edu.co
https://orcid.org/0000-0001-5221-8245

Angela Patricia Velásquez Hoyos*
*Universidad de Caldas | Manizales | COL
angela.velasquez@ucaldas.edu.co
https://orcid.org/0000-0002-9743-5769

Luis Herney Villegas López*
*Universidad de Caldas | Manizales | COL
luis.villegas@ucaldas.edu.co
https://orcid.org/0000-0001-5313-9370

Abstract: In the field of language assessment, particularly language assessment literacy (LAL), there is a current need and call to advance pedagogies that engage teachers in language assessment training. As a possible contribution to this field, in this research article, we report on the design and impact of a language assessment course tailored for the specific contextual needs of five Colombian English language teachers. We collected data from teachers and course instructors through a multi-method approach involving five data collection instruments: non-participant observations, course instructors' reflective logs, participants' learning artifacts and reflective logs, and an individual interview. Upon data analysis, we found that the course contributed to teachers' contextual knowledge and practice in language assessment. Additionally, we highlight core course features (tasks, materials, and interactive reflection) and teachers' attitudes (active participation, peer feedback, and collegiality) that together formed a local LAL pedagogy. Based on these findings, we share conclusions and recommendations for context-sensitive LAL courses for language teachers.

Keywords: assessment literacy; language assessment; language assessment literacy; teacher learning; teachers' professional development.

Resumo: No campo da avaliação de línguas, particularmente na área de letramento em avaliação de línguas (LAL), há uma necessidade e uma demanda atual pelo avanço de pedagogias que engajem

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

professores em formação na área de avaliação. Como uma possível contribuição para este campo, neste artigo de pesquisa, relatamos o desenho e o impacto de um curso de avaliação sob medida para as necessidades contextuais de cinco professores colombianos de língua inglesa. Os dados foram coletados junto a professores e formadores do curso por meio de uma abordagem multimétodo, envolvendo cinco instrumentos de coleta: observações não participantes, diários reflexivos dos formadores, artefatos de aprendizagem e diários reflexivos dos participantes, além de uma entrevista individual. Após a análise dos dados, constatamos que o curso contribuiu para o conhecimento contextual e para a prática dos professores em avaliação de línguas. Além disso, destacamos características fundamentais do curso (tarefas, materiais e reflexão interativa) e atitudes dos professores (participação ativa, feedback e colegialidade) que, juntas, constituíram uma pedagogia de LAL. Compartilhamos conclusões e recomendações que podem fomentar futuros cursos de LAL sensíveis ao contexto para professores de línguas.

Palavras-chave: letramento em avaliação; avaliação de línguas; letramento em avaliação de línguas; aprendizagem de professores; desenvolvimento profissional docente.

1 Introduction

In the field of language assessment, language assessment literacy (henceforth LAL) has become a key area of scrutiny, with growing attention to teachers as a central stakeholder group in LAL research (Fulcher, 2012; Giraldo, 2018; Inbar-Lourie, 2008, 2013). Language teachers have participated in language assessment courses (Arias; Maturana; Restrepo, 2012; Montee *et al.*, 2013; Giraldo; Yan, 2025b). Precisely, the design and implementation of language assessment courses have emerged in LAL research, with scholars encouraging and igniting ideas for what may be called LAL pedagogies (Fulcher, 2020; Giraldo, 2024; Vogt, 2025). As research has shown, LAL courses generally impact several aspects of language teachers' professional development in testing and assessment (Giraldo, 2021; Moraes, 2023): increased awareness of what language assessment implies; better assessment design and use; and the connection among teaching, learning, and assessment.

Since LAL pedagogies are encouraged, the field may benefit from research-based practices around LAL courses for teachers. Therefore, to potentially fuel research into LAL

pedagogies, in this article we report on the impact of and insights from an online assessment course that was designed for five English language teachers from different educational institutions in Colombia. We characterize the impact of the course by relying on data coming from the five teachers and us, the three course instructors.[1] Since we designed, taught, and evaluated the course, our position as stakeholders may provide further feedback on LAL courses. Thus, with our findings and suggestions, we hope to contribute to LAL pedagogies by discussing the nature of the course and its impact, so that our insights may be transferable to and useful in other contexts (Guba, 1981; Merriam, 1996). Specifically, we aimed to answer the following research questions:

- How does a tailored language assessment course impact the LAL of five in-service English language teachers?
- What course characteristics impact the teachers' LAL and form a contextual LAL pedagogy?

2 Teachers' LAL in Context

LAL refers to the knowledge, skills, and principles teachers possess to assess students' language development (Fulcher, 2012; Inbar-Lourie, 2008). Therefore, LAL deals with the effective design and administration of assessment instruments, the valid interpretation of results from those instruments, and educational decisions made from such interpretation (Fulcher, 2012; Weng; Shen, 2022). Given its paramount role in ensuring reliable, authentic, and fair assessment practices, fostering LAL among teachers is essential. As Giraldo and Murcia (2018) explain, LAL cannot be applied uniformly for teacher training but must be adapted contextually so that such training is carefully planned and evaluated.

As Hidri (2020) argues, assessment leads to learning, reinforcing why developing LAL is crucial in language education. Without strong assessment literacy, teachers can face difficulties in supporting students' language learning process, designing meaningful assessments, interpreting results accurately, or providing feedback that supports language growth. According to Levi and Inbar-Lourie (2019), the engagement of teachers in LAL equips them to make informed decisions based on assessment results, which ensures valid and meaningful assessment of students' language development. For these authors, LAL empowers teachers to design, use, and implement diverse assessment strategies that improve language learning.

Moreover, LAL equips teachers with the necessary tools to address diverse classroom realities through the connection of individual and contextual factors (Villa-Larenas; Brunfaut, 2023; Weng; Shen, 2022). On an individual level, teachers' linguistic backgrounds, training experiences, and academic preparation help them build their assessment competence (Lam, 2015; Yan; Zhang; Fan; 2018). Additionally, LAL allows language teachers to adapt themselves

[1] The three authors work for the Foreign Languages Department of Universidad de Caldas in Manizales, Colombia.

to external factors such as institutional policies, national assessment cultures, and resource constraints (Giraldo, 2023; Mohammadi, 2024).

Various scholars (Inbar-Lourie, 2024; Scarino, 2013; Taylor, 2013) have strongly emphasized that LAL is fundamentally shaped by its specific context and influenced by its sociolinguistic and assessment environments. Considering these authors' ideas, LAL is not just about mastering technical skills but also about developing the capacity to make context-sensitive judgments that account for local learning cultures, institutional policies, and students' sociolinguistic backgrounds. The scholars further argue that effective LAL enables language teachers to construct assessment events that are methodologically appropriate, pedagogically relevant, and socially meaningful to teachers' educational institutions.

Given this sociocultural dimension, scholars advocate for context-sensitive LAL training (Inbar-Lourie, 2024; Scarino, 2013, 2017; Taylor, 2013). According to these authors, such training should prepare teachers to design assessments grounded in teachers' contextual needs, which will help them critically evaluate how their assessment events and decisions align with their specific educational environments. In the next section, correspondingly, we review existing research on teachers' LAL development.

2.1 Related Research

A major research trend regarding teachers' LAL is the impact that LAL training has on their profiles. First and foremost, research studies report on teachers' increased knowledge of language assessment concepts and better designed assessment instruments, whether for classroom purposes (Arias; Maturana; Restrepo, 2012; Koh *et al*., 2018; Levi; Inbar-Lourie, 2019) or for large-scale national language tests (Baker; Riches, 2018; Kremmel *et al*., 2018). These studies, collectively, point towards practical hands-on tasks and interactive training sessions as catalysts to address both language assessment theory and practice in LAL training. In fact, such a practical approach to teachers' LAL development is widely encouraged by LAL scholars (Fulcher, 2012; Giraldo, 2024; Montee *et al*., 2013; Stabler-Havener, 2025).

Naturally, LAL training initiatives are meant to foster teachers' LAL, but the aspects that make LAL training fruitful may not be explicitly articulated in existing research. Upon our review of studies on assessment courses, we noticed that successful experiences have these core traits:

- *Diagnostic stages*: Studies have shown that useful LAL training is based upon rigorous diagnostic stages that, mostly, have used mixed-methods approaches to thoroughly describe teachers' LAL before course implementation (Arias; Maturana; Restrepo, 2012; Giraldo, 2024; Giraldo; Yan, 2025b; Koh *et al*., 2018; Montee *et al*., 2013; Nier; Donovan; Malone, 2009).
- *Collaborative and interactive spaces for LAL development*: Other studies have suggested that teachers need LAL activities so they can explicitly problematize language assessment issues that are pertinent to their practice. As these spaces occur in LAL training, teachers contribute to their LAL through collegiality and even feedback on their LAL (Baker; Riches, 2018; Giraldo, 2023; Spiliotopoulos *et al*., 2024).
- *Relevant, context-based materials*: An emerging issue in teachers' LAL training is the role of materials for LAL development. Various studies have indicated that materials

> for these stakeholders should be as contextualized as possible so that they relate to teachers' explicit needs (Giraldo, 2024; Fulcher, 2012; Sheehan; Thai, 2025; Tsagari; Armostis, 2025; Vogt, 2025). The authors in these studies collectively recommend that such materials be friendly and practical, without losing their technical nature.

These studies have mostly focused on the perceptions of teachers towards their training in language assessment. However, little is known about how other stakeholders, e.g., LAL course instructors, can contribute to understanding the impact of LAL training on teachers' professional development. Baker and Riches (2018), for example, studied the perceptions of course facilitators on their implementation of assessment workshops for EFL teachers in Haiti. The authors report that facilitators gained an understanding of the social and cultural context for EFL assessment in this country – notice, then, that this finding focuses on facilitators' growth, not on teachers' professional development and what contributes to it.

In this study, we seek to contribute to related LAL research by examining the impact of a language assessment course as perceived by the teacher participants *and* us, the course instructors. The course we designed and co-taught was called Language Assessment in Context (LAiC). In this paper, we describe its nature and impact through a multi-method approach in the Methodology section below.

3 Methodology

This research was framed under a qualitative approach to explore the experiences and perceptions of five English language teachers and three course instructors regarding the LAiC course. Adopting a descriptive exploratory case study design (Yin, 2014), the investigation sought an in-depth, contextualized understanding of how the course influenced the teachers' LAL. To develop the course, we conducted a diagnosis (see Giraldo; Velásquez-Hoyos; Villegas-López, in press, for the full report) that included a questionnaire, an interview, an examination of teachers' institutional documents, and an analysis of their assessments to determine their specific needs related to LAL. After data analysis, we concluded that even though teachers advocated for a communicative approach in their assessment, they faced some challenges for implementation: evidence of construct irrelevance in their assessments and constraining administrative policies, which created a disconnection between their beliefs and assessment practices. Therefore, this diagnosis revealed that the LAiC course needed to be highly context-sensitive to address identified needs and expectations.

3.1 Participants

The participants of this study were five in-service Colombian English language teachers and three course instructors. The in-service teachers all came from diverse academic, teaching, and regional backgrounds (see Table 1). For ethical considerations, each participant selected a pseudonym to help them express their reflections and experiences freely. Additionally, to decrease instructor-related biases in one data collection method (non-participant

observation), we validated this instrument through expert review and piloting it, so we could determine that our observations were as objective as possible.

The teachers' sample was assembled through convenience sampling, prioritizing accessibility. We contacted 107 English language teachers via email since these teachers had prior involvement in assessment training programs organized by Giraldo. The email stated the study's objectives, procedures, timeline, and required documentation. Of those contacted, five teachers agreed to enroll and participate.

Table 1 – Teachers' Backgrounds

Pseudonym	Region of Colombia	Institution	Type of Institution
Prudente	Northwest	University	Private
Teacher C Maleja Johnny Rotten	West Center	High school	Public
PFC-P	Southwest		

Source: The authors.

The three course instructors are language teacher educators at a Colombian public university. They all hold a master's degree in English language teaching and have experience training in-service teachers through language assessment. For this research, the three instructors facilitated sessions and conducted class observations at various stages of the LAiC course. Their roles were rotational, depending on a week's LAL contents (see Table 2). For example, in one session one instructor would lead the class while another observed, which facilitated a dynamic and collaborative approach to data collection.

Table 2 – LAiC Contents and Instructors' Roles

Weeks	Week's Contents	Instructor	Observers
1-2	Communicative Language Testing	Researcher 1	Researchers 2 & 3
3-4	CLT & Washback		
5-6	Authenticity in assessment		
7-8	Construct (Ir)relevance & Score Pollution	Researcher 3	Researchers 1 & 2
9-10	Principles for an Assessment Culture	Researcher 2	Researchers 1 & 3

Source: The authors.

3.2 Course Design

The LAiC course had a total of forty learning hours, twenty of independent work and twenty of synchronous work. Before every weekly online workshop, the teachers read preparatory

material about the week's content. They also had to do tasks such as analyzing an assessment instrument given some guiding questions or reflecting on assessment challenges in their context. During the synchronous sessions, the course instructors provided LAL input through lectures and engaged the teachers in reflection and action on language assessment, based on the given content for a week. During the sessions, teachers and instructors used a variety of materials: a collaborative and editable Google Doc where teachers shared ideas and practices; handouts with reflection tasks or samples; slides for the lectures; sample assessment instruments and test-related videos for analysis; and Google Meet as the chosen platform for the course.

3.3 Data Collection Instruments

Five data collection instruments were administered during the LAiC course: a non-participant observation format, an instructor's reflective log, learning artifacts, a teacher's reflective log, and an individual interview. To contribute to the quality and validity of these five instruments, we asked two language teaching and research experts to provide feedback on their nature and whether the instruments were fit for purpose. From the experts' comments, we made the following two major improvements: specified categories for the observation format and instructor's reflective log; refined questions in the individual interview so they tapped into contextual knowledge and use of assessment theory and practice.

Non-participant observations: The three researchers played the roles of course instructors and observers. One researcher delivered the lesson, and another observed it. This instrument was structured and focused on aspects such as participants' engagement with the course, input, roles, interaction, and reflections on teachers' LAL development, among others. In sum, the instrument centered on aspects during the lessons that could have an impact on teachers' LAL and offer feedback for LAL pedagogies (see Appendix A for the observation protocol). This instrument was piloted by researchers 2 and 3, as they observed two lessons conducted by researcher 1. The protocol was then examined to facilitate understanding, detect possible biases, and clarify key observation categories, e.g., instructor's and participants' roles.

Instructor's reflective logs: This instrument gave each course instructor the opportunity to reflect and collect their feelings about each LAiC lesson, specifically positive aspects and elements to improve. Each instructor completed a log after delivering a lesson. Thus, in total, eight reflective logs were completed by the instructors throughout the LAiC course.

Learning artifacts: We collected and analyzed ten learning artifacts: a document with communicative learning objectives formulated by the teachers; a small-scale language assessment project proposal created by each teacher (five in total); and a document with a reflection on an assessment situation the teachers analyzed in class (four in total). To examine the artifacts, we used a grid consisting of predetermined categories related to major foci of the LAiC course. The following is an example of a criterion: the learning artifact shows evidence of communicative language testing principles.

Teachers' reflective logs: At the end of the course, the teachers were asked to share some insights and experiences about the LAiC course through a reflection log consisting of four questions. The questions asked about insights and concepts gained from the course, skills they felt they were developing, the connection between the course and their contexts,

and the overall impact on LAL. The information from this log was used as baseline data for the individual interview.

Individual interviews: The participants had the opportunity to express, in a Google Meet oral interview, their reflections and overall evaluation of the LAiC course. We asked the teachers whether the information from the course was useful for their assessment approach and whether they were implementing anything they learned in the course. Before the interviews were conducted, this instrument was piloted with one volunteer from the course whose data we are not reporting in this research article. During the piloting, the participant was asked how he felt during the interview, and if he had any suggestions or comments regarding the interview. With the information collected through the piloting and the analysis, the three researchers improved the interview by clarifying concepts such as *assessment approach*. See Appendix B for the interview protocol.

3.4 Data Collection

We collected data in this study in three ways during and after the ten weeks of the LAiC course. First, we conducted ten non-participant class observations, two of which were used for piloting the observation format and eight for data analysis. Along with these observations, we each wrote one entry in our instructor's logs: instructor one wrote six entries, whereas instructors two and three wrote two each. Second, we collected the ten learning artifacts during various weeks in the course: document with communicative objectives during week four, reflection document during week six, and the assessment project during week ten. During week ten, each teacher wrote their entry in the teacher's learning log.

Finally, the individual interviews lasted between 16 and 28 minutes (*Mdn* = 27) and were conducted either in English or Spanish, depending on the teachers' preference. The interviews took place in September 2025, four months after the LAiC course was finished. By doing the interviews at this time, our aim was to collect information on teachers' actual use (or lack thereof) of LAL course contents, which scholars have noted as a gap in LAL research (Gan; Lam, 2022; Giraldo, 2021).

3.5 Data Analysis

We applied thematic analysis to the data collected from the five instruments we employed. We followed the steps proposed by Braun and Clarke (2006), which are familiarization with data, generating preliminary codes, searching for themes, reviewing themes, defining and naming themes, and writing the report.

We used a spreadsheet to analyze the first 30% of the data across all five instruments (observations, instructor's reflective logs, teacher's reflective log, learning artifacts, and interviews) to identify preliminary patterns. Next, we held four meetings, lasting two hours each, to assess inter-rater reliability using Fleiss' kappa. In the first iteration, we achieved 33% agreement. Then, we refined the coding framework and defined concepts for the codes; for instance, we came to an agreement on the concepts of *tasks* and *input*. In the second iteration, the remaining 70% of the data was analyzed, resulting in a moderate inter-rater agreement

of 72%. Through four discussions, we consolidated 46 open codes and turned them into 11 broader categories. We then discussed these categories, from which three major themes emerged. The themes, which we describe next, captured the impact of the LAiC course, its pedagogical nature, and teachers' attitudes towards it.

4 Findings

4.1 Impacts of the LAiC Course on the Teachers' LAL

After data analysis, it was clear that the LAiC course impacted the teachers' LAL in two major areas: contextual application of course contents and an expansion of knowledge about language assessment, especially regarding principles of communicative language testing (CLT).

4.1.1 Contextual Application of LAiC Course Contents

First and foremost, our analysis of interview answers, learning artifacts, and reflective logs indicated teachers' use of LAiC contents for their teaching and assessment contexts. For instance, in the interview, PFC-P comments on her implementation of communicative assessment:

> What I have implemented has to do with the three subcompetences of communicative competence. The course allows you to realize that it is not necessary to think about these subcompetences as different [...] but can be [...] an integrated objective through which the three [linguistic, pragmatic, and sociolinguistic competences] can be evidenced. I am currently making adjustments to the learning objectives so as to integrate and see the connection among the three for assessment.

PFC-P's answers can be related to the way the LAiC course was designed. During the course, CLT was addressed explicitly during two workshops (4 course hours on weeks one and two) and permeated six more, e.g., washback and CLT during weeks three and four. Additionally, one way in which CLT was addressed in the course was through the design of communicative objectives, which is something PFC-P discusses in her answer.

In the following learning artifact, Maleja described a communicative task which she intended to implement in her school. Maleja's plan starts with a task-based objective, a major focus in the course. Like PFC-P's answer, Maleja is operationalizing communicative competence by integrating the structural, functional, and social aspects of language needed for task accomplishment. Finally, in the following sample from a learning artifact, the small-scale assessment project, Maleja proposes the use of two alternative assessment instruments, which the course highlighted as representative of CLT.

> Grade: 8th graders
> General objective: Students will be able to create a travel guide introducing their towns, using...
>
> *Linguistic*:

> Nouns: vocabulary and descriptive adjectives; syntax: simple sentences using present tenses
>
> *Pragmatic*:
>
> describe, adapt language for visitors, simple explanation and use a functional language highlighting the positive aspects
>
> *Sociolinguistic*:
>
> cultural representation
>
> community values
>
> Assessment instruments: Self-assessment through a checklist; peer feedback prompt.

In their reflective log, Maleja commented on the need to address authenticity in assessment, a core CLT principle she is now considering for her context. One of the log prompts asked the teachers what insights of concepts they had gained from the course as applicable to their context, to which Maleja wrote: "I noticed the importance of designing authentic tasks that mirror real-world communication".

4.1.2 Teachers' Reported Expansion of Language Assessment Knowledge

Regarding teachers' expansion of knowledge about language assessment, we used data from both an *emic* and an *etic* perspective. *Emic* data came from teachers' reflective logs and interview answers. *Etic* data, on the other hand, came from our instructors' logs and class observations. In this finding, we illustrate reported LAL knowledge that was evident across all five teachers, from these four data collection methods. For instance, in their reflective log entry, Teacher C commented that

> The course had a great impact not only on my theoretical understanding of language assessment but also on my practical and ethical approach to it. In terms of knowledge, I gained a clearer picture of what makes assessment valid, reliable, and authentic. Which I hope to apply in my daily practice here at school.

Teacher C's answers reflect what are now widely used concepts in communicative language testing, namely validity and authenticity (Bachman; Palmer, 1996; Green, 2014). We also note Teacher C's willingness to implement this knowledge through practice, an idea explored in the previous section of this finding.

After analyzing our reflections as course instructors, we found evidence of solid teacher LAL learning. For example, the following instructor's journal entry reiterates teachers' LAL knowledge development and shows the connection among CLT principles, specifically authenticity and task-based assessment: "I believe teachers solidified the idea that authenticity is about the curriculum and how assessments should reflect that through tasks and items. Also, they reinforced, I think, the three ways to make objectives more authentic and communicative: task- and theme-based and functional." The LAiC course remarked on the idea of establishing authentic learning objectives, which the entry highlights as 'the curriculum.' Also, teachers learned about three ways to write such objectives, which are presented at the end of the entry.

In conclusion, across all five teachers, data from various methods suggest that the impact on teachers' LAL was, first and foremost, practical, e.g., implementing communication-based objectives, and theoretical, e.g., principles in CLT. The data also seem to indicate that

the practical and theoretical dimensions of language assessment were mixed. This notion is something we revisit in the discussion section of our article.

4.2 Instructional Design of the LAiC Course for an LAL Pedagogy

Data analysis suggests a clear contribution of the LAiC course instructional design to the development of an LAL pedagogy. This contribution was evident through the main components of the course: contextualized task design and implementation, materials relevant to teachers' specific contexts, and interactive and reflective sessions. We believe that these components, collectively, shaped a cohesive framework that seemed to foster participants' LAL.

4.2.1 Contextual Tasks for LAL

Regarding the first component, contextualized task design, the instructor's reflective log provides evidence of context-sensitive tasks that allowed teachers to bridge theory and practice in language assessment. In the LAiC course, the tasks were designed around issues that participants encountered in their daily practice as teachers, which motivated them to analyze, adapt, and apply assessment concepts and principles in their institutional realities. In workshop 4, participants examined *Camila's case*, which involved reading a real situation in which a teacher had to reformulate the school's syllabus and incorporate communicative objectives to assess her students. The contextualization of tasks was evident in the instructor 1's reflective log for workshop 4, a trend that was consistent in this data collection instrument:

> The content tonight was about how to integrate curriculum, assessments, and assessment for learning. In this regard, I believed teachers' reaction was positive, especially in Camila's case; I feel like this example helped them to contextualize washback a bit more, when it comes to the three elements already mentioned.

This contextualization of tasks was also evident in teachers' reflective logs and non-participant observation, as noted in the following excerpt from Teacher C's answer:

> the course was designed based on the input I provided in the interview. The practical activities were always centered on my setting, as well as the examples and tasks [...] during the LAiC course. When we discussed the issue of large-scale tests, it was the test I had worked on in a session with my students.

Additionally, a non-participant observation reinforced this finding by demonstrating how contextualized task design, such as analyzing real teachers' cases or examining the influence of large-scale tests, effectively fostered meaningful connections between theoretical constructs and teachers' lived classroom experiences. As researcher 3 states in the log for workshop 3, some teachers reflected on what can be done with large-scale tests (e.g., the TOEFL) in their contexts to meet specific objectives: "Based on a task proposed in the workshop, participants (Johnny Rotten and Maleja) mentioned that their schools have used large-scale tests. They reflected on the impact of washback on the application of these large-scale tests like the TOEFL".

4.2.2 Materials for Teachers' LAL Development

The second LAiC course component that contributed to the construction of an LAL pedagogy was the design of materials and their relation to teachers' context. Materials such as slides, the workshop sequence, videos, and samples of tests, among others, enabled participants to internalize assessment principles, e.g., alignment among learning objectives, course activities, institutional curricula, and assessment instruments. The following data, which comes from an observation and represented a common trend throughout the course, describes this issue: "Contributions to LAL pedagogies: different materials such as videos, audios, the workshop sequence and the interaction with course material prior to the class prepared teachers for contextualizing, reflecting and discussing in the LAiC course".

The following reflection also aligns with the role of materials in fostering teachers' LAL development. As Johnny Rotten noted in the interview: "Regarding the materials of the course, let's say to cover a grammar topic. If I revise my exam and think about all the insights learned from the course and the handbook, I realize that that is not real communication or that is not connected to real communication".

These pieces of evidence suggest that material design in the LAiC course served as a guiding and mediating tool for teachers to contextualize assessment concepts, understand the importance of pedagogical coherence, and engage in reflective professional practice that helped teachers recognize inconsistencies between their stated learning objectives and the types of assessment tasks and instruments they were designing.

4.2.3 Engagement of Participants Through Interaction and Reflection

Finally, interactive and reflective sessions supported the construction of an LAL pedagogy. For instance, in the non-participant observation in workshop 4, the observer reflected on how different aspects of the sessions, such as participants' reflections, the interaction style, the role of the participants, and the use of different multimedia tools, created a positive classroom environment that facilitated an appropriate development of the course.

> Another important aspect that contributed to the construction of an LAL pedagogy in this session was participants' reflections and engagement with the course content […]. The session allowed teachers to be listeners, reflectors, writers, and collaborators by using different tools such as the chat box, the workshop sequence, and real case readings […]

In their reflective logs and interviews, teachers also explored how interaction and reflection impacted their LAL in context. The two data samples below provide evidence for this trend.

Johnny Rotten's entry in their Reflective log

This course always has a positive impact in my job since there is lots of reflection, raising awareness and revising concepts that are directly related to what we do with our students.

Maleja's answer in their Interview

> The course was useful because of all the reflective part, the practical activities we did and that we can apply in our schools.

The data from non-participant observation, teachers' reflective log, and interview reveal that the design and implementation of the course session fostered teachers' ability to make connections between theoretical constructs and their teaching realities. The emphasis of the course on reflection, collaboration, and contextual application allowed participants to problematize assessment for their teaching contexts. To conclude, the information provided by all the participants illustrates how the instructional design of the LAiC course fostered a dynamic and context-responsive LAL pedagogy that valued their lived experiences.

4.3 Teachers' Attitudes Contributing to the LAiC Course

The third major finding drawn from data analysis was teachers' salient attitudes during the LAiC course. These attitudes were operationalized in three ways: teachers' reflection about their LAL contexts, high level of participation during the course, and collegiality. Together, these attitudes contributed to the course's LAL pedagogy, which we described in the previous finding.

4.3.1 Reflection on LAL Contexts

All five teachers' reflective stances were identified throughout the entire course and were triggered by either course contents or tasks. For instance, the non-participant observation in week 3 reports that Johnny Rotten participated by reflecting on the focus on the preparation of *Pruebas Saber* (a national test): "I have used a negative washback focusing on Pruebas Saber. Not only pruebas saber but everything. There was not a specific goal". Then, in the same instrument, the observer describes how other teachers reflected on their contexts, connecting them with course topics: "Prudente talks about the importance of positive washback with ethical and fair assessment practices". As Prudente explained in the lesson: "We are always focusing on objectives. We assessed what is good and not so good. If I am hesitating about a student grade, then I share the exam with another colleague to have better assessment practices and have reliable results".

Additionally, the following data piece, from an observation, also shows how course contents ignited self-reflection about the teachers' contexts. In workshop 3, Teacher C explained score pollution in his teaching context, describing how sometimes students are graded in aspects unrelated to language learning: "It is problematic for students on the grading scale. The positive washback is software management as students learn how to use software. However, this could lead to score pollution." The reflective stance in this data piece relies on how Teacher C recognizes some positive washback, learning how to use software. However, he also reflects on how that learning is regarded as "problematic" since it is not assessing language skills.

Finally, in the individual interview, Teacher C highlighted that one relevant aspect during the course was the opportunity that the teachers had to share their insights about their contexts, and how aware he is now about his assessment practices: "we were able to give

you some insight on how is our day-to-day, our institutional planning, how we work, and this was brought into the content of the course, which allowed me to become more aware of how I assess my students and if there is no assessment culture where I work".

To conclude, reflective spaces in the LAiC course stimulated teachers' enactment of LAL. Such a stance was possible because the course connected the cognitive side of LAL pedagogies (i.e., learning theory about assessment) with the sociocultural dimension of teachers' contextual LAL. Thus, it seems that reflection was a crucial aspect to invite teachers to problematize language assessment, rather than being passive recipients of knowledge in the LAiC course.

4.3.2 Active Participation

During the LAiC, the teachers' participation was active and multimodal. Their engagement and interaction were noticed as they typed comments and contributions, and when using emojis and reaction buttons from the Google Meet platform. Throughout the course, the participants could use their microphones, the workshop sequences, and the Google Meet chat to ask and answer questions, and to participate. To illustrate this finding, in non-participant observation for workshop 6, the teachers are described as "active participants" because they added comments about the material they read prior to the session. The teachers are also described as reflective practitioners who participated "by answering questions and connecting their reflections with their teaching contexts".

Participation also consisted of teachers providing feedback on class tasks based on their experiences. For instance, in the observation for workshop 5, Maleja commented on an example of a learning objective: "the reading is not very communicative, they lack real-life purpose, it's reading comprehension, not interaction".

Additionally, teachers gave feedback to each other throughout the course. For example, the following data, from a non-participant observation, shows that the teachers participated by sharing their knowledge about LAiC course topics. When discussing content validity in workshop 7, Teacher C gave an example: "The highest score would be the one that uses polite language in the test, but I never taught that in the course". Prudente contributed to Teacher C's comment by stating the following: "it is basically assessing something you have never taught".

4.3.3 Collegiality Among Course Participants

The LAiC course provided participants with an approachable environment where they could share their LAL experiences and insights. This atmosphere of academic community was reflected in their engagement throughout the LAiC course, and it was also recorded across different data collection methods, especially in the instructors' reflective log and non-participant observations. For instance, in their reflective log in workshop 3, the instructor comments on the participants' rapport: "The teachers showed respect to each other's contributions, which feels like a way to contribute to the small community we have in the course". In workshop 4, the non-participant observer reported that: "The use of friendly and accessible language by teachers and the instructor, not very technical, helped participants

better understand LAL concepts and see their relevance in their daily teaching. This made the learning experience of the course feel less tedious and more inviting".

In conclusion, thanks to our data analysis, it seems that the LAiC course was not just a course to develop LAL but to engage in respectful discussions about language assessment. This effect of the course was possible due to teachers' willingness to work with and for others based on common LAL goals. Thus, teachers' collegiality in the course instilled a friendly yet professional perspective to LAL development. As such, their attitudes contributed to course success by making the LAiC course a community of language assessment practitioners.

5 On the Use of Lectures in the LAiC Course

Unlike the findings we have presented thus far, thanks to the analysis of data from the non-participant observations and our instructor's journal entries, we found that lectures were an area for improvement throughout the LAiC course. Specifically, we noticed that the lectures we used were mostly transmission-oriented, instructor-dominated, and took longer than planned. Consequently, we determined that they did not seem to engage teachers in LAL development or enactment. For instance, in non-participant observation 8, the observer noted issues with content delivery in the lecture and time management: "During the lecture on score pollution, the instructor took 20 minutes longer than initially stated in the lesson plan for this workshop. In total, then, the lecture took 35 minutes of the instructor fronting the lesson and teachers, apparently, passively listening".

In the reflection log the instructor wrote for this workshop, the data signals time management as an issue for improvement. Specifically, in the prompt to reflect on materials use, the instructor stated that he "did not exploit the material completely because time ran out very quickly". Additionally, the instructor reflected on what did not go so well in this session: "Exploiting more the material in the lecture while reducing my talking time".

A similar sentiment was echoed in the entry of the instructor's journal for workshop 3. In this case, the instructor focused on a problem regarding the lecture section of this lesson.

> What did not go so well?
>
> The mini lecture, because I took longer than I considered. Also, I wonder whether teachers really got the meaning, especially, of validity. I did not get any feedback from them and could not make the lecture more engaging and/or interactive to check on their LAL knowledge of this topic.

To conclude, while the LAiC course was mostly successful pedagogically speaking, our analysis of data led us to notice that lecturing was an area that required attention from us three as course instructors and researchers. While lectures may be a staple activity or moment in LAL courses for language teachers, we have some evidence to problematize their use in two aspects: how to make lectures interactive so teachers can engage with their content and how to manage time well while using lectures.

6 Discussion

From the findings in our study, we reiterate the crucial role of teachers' contextual needs for professional development through LAL training (Baker; Riches, 2018; Giraldo, 2024; Scarino, 2013; Tsagari; Armostis, 2025). Taking teachers' needs into account may have been a major reason why the course was useful to them, both for applying LAL contents in their contexts and for advancing their knowledge about assessment. Thus, one major component of LAL pedagogies, for teacher training, is taking their assessment needs into consideration, which we did in two ways: describing their LAL as comprehensively as possible and devising an LAL course based on such description.

Regarding impactful LAiC course characteristics, we believe this may have happened primarily for two reasons. First, we based the course on a task-based approach to teachers' LAL development (Fulcher, 2020; Giraldo, 2024): context-based materials for input, interactive roles for participants and instructors, and contextualized tasks and activities. Second, we believe the course was impactful because we made explicit design decisions based on principles for language teachers' professional development (Cárdenas; González; Álvarez, 2010; Freeman; Johnson, 1998; Kumaravadivelu, 2012): developing the LAiC course in a bottom-up manner; engaging teachers in reflection and action for their teaching – and assessment – contexts; practice-based teacher learning; and collegial, collaborative activities and tasks.

Furthermore, we believe the main reason for teachers' attitudes in the course was self-selection. The teachers chose to participate in the course, given their interest in developing their LAL. This phenomenon of teachers' interest and motivation towards language assessment has been addressed in other studies (Fulcher, 2012; Giraldo, 2023). We might have had different findings to report had the teachers in the course been obligated to partake. Additionally, four of the five teachers had already participated in LAL training conducted by Giraldo, and the remaining teacher was tasked with teaching language assessment in an undergraduate degree. Thus, the five teachers were highly motivated to be in the course, which led them to perform the roles they did in favor of the course. Lastly, the five teachers in our study had similar concerns in their assessment contexts, e.g., test washback and authenticity in task design. We believe that these shared concerns, and the other reasons we discuss, sparked teachers' reflections, participation, and collegiality. Finally, regarding our finding about using lectures in the LAiC course, we consider that our unsuccessful use of this strategy stems from the idea that we used them only to provide content, following frameworks for LAL pedagogies in which providing input is an essential component (e.g., Fulcher, 2020). How to teach that input successfully through lectures, however, is not documented in scholarly LAL literature.

7 Limitations

Four limitations in our study must be addressed. Firstly, our findings are not generalizable, nor are they intended to be so. The findings are grounded contextually and socially and, as such, represent a case study that can contribute to understanding teachers' LAL *in situ* (Baker; Taylor, 2024; Giraldo; Yan, 2025a; Zhang; Yu; Browne, 2025). Secondly, the findings may have been caused by the Hawthorne effect: the five teachers knew they were to be observed and

their ideas about language assessment were to be studied, so they did their best in the LAiC course and when providing insights. This effect may have also applied to our roles in the study: we observed each other's LAiC lessons at least once and wrote personal reflections on the LAiC course and its participants. Our ideas, then, may have influenced unconsciously the way we behaved as instructors and, as such, should be analyzed carefully in other LAL contexts.

Third, we conducted the individual interviews four months after the LAiC course, which means we depended on teachers' memory to provide insights on the course and its relation to their teaching contexts. Finally, we consider that we could have gathered more robust results had we conducted class observations and documentary analysis to better account for post-course impact. Notwithstanding these limitations, we argue that there can be transferability of the insights we have gathered from the findings in our study (Guba, 1981; Merriam, 1996) so that they can be useful for other LAL teacher educators and researchers tasked with designing and/or teaching assessment courses for teachers.

8 Conclusions and Recommendations

Acknowledging that LAL training empowers teachers to navigate the complex connection between their individual expertise and their specific institutional, sociocultural, and sociolinguistic environments, this research responded to the call for context-sensitive LAL courses. In response to our research questions, this study demonstrates that the LAiC course effectively fostered the LAL of participants by bridging theoretical principles with contextualized practice. The course's instructional design (e.g., the use of contextualized tasks, varied materials, and reflective sessions) related directly to teachers' institutional realities and contributed to participants' LAL. Furthermore, the use of collegial and dialogic elements of the LAiC course, along with multimodal engagement, was fundamental for creating a positive classroom atmosphere, transforming the learning process into a collaborative endeavor where participants became active agents in their own professional development.

Additionally, the findings of this study corroborate that teachers' LAL development is not a homogenous set of technical skills, but a situated, context-dependent competence which is essential for teachers to make fair, valid, and pedagogically sound assessment decisions. Through the design and implementation of contextualized tasks and materials, this study contributes to understanding how LAL training helps teachers to connect communicative language assessment principles (e.g., washback and authenticity) with their teaching contexts. Hence, this study aims to inform that tailor-made LAL training courses empower teachers to further construct assessment events that are methodologically appropriate within their different educational contexts.

For future research, teacher educators developing LAL training programs should take into consideration the contextual realities of participants. These professional development courses need to be designed around the needs and interests of stakeholders. This exchange of experiences and reflections could also shape a collegial atmosphere in the LAL course, as participants might also be able to provide feedback and contribute to class discussions.

Furthermore, it is recommended that future studies involve a larger number of participants and include additional stakeholders, as research has shown that such initiatives have positive and sustained impacts. In addition, although we have limited data on this

matter, we consider that the role of lectures in teachers' LAL learning and development should be studied further. Finally, to evaluate the long-term effects of an LAL course for teachers, a post-course study should be conducted, incorporating classroom observations as data collection instruments to identify the program's actual impact and remaining limitations in real teaching contexts.

Funding

This article is part of a research study called *Diseño e Implementación de un Programa de Desarrollo Profesional Docente en Evaluación de Lenguas Extranjeras* (Design and Implementation of a Teachers' Professional Development Program on Language Assessment), which was sponsored by the research office (Vicerrectoría de Investigaciones y Posgrados) at Universidad de Caldas in Manizales, Colombia. The code for this research project is PRY-181.

Ethics Committee Report

We, the three authors, state that our research was conducted in accordance with internationally recognized ethical principles in the case of research involving human beings; specifically, all five participants in the study received a thorough information and consent letter which they all signed. This research project was approved by the Ethics Committee of the College of Arts & Humanities of Universidad de Caldas on September 12, 2023. This approval was then validated with our project code: PRY-181.

References

ARIAS, C.; MATURANA, L.; RESTREPO, M. Evaluación de los aprendizajes en lenguas extranjeras: hacia prácticas justas y democráticas. *Lenguaje*, Cali, v. 40, n. 1, p. 99-126, 2012.

BACHMAN, L.; PALMER, A. *Language Testing in Practice*: Designing and Developing Useful Language Tests. Oxford: Oxford University Press, 1996.

BAKER, B.; RICHES, C. The Development of EFL Examinations in Haiti: Collaboration and Language Assessment Literacy Development. *Language Testing*, v. 35, n. 4, p. 557-581, 2018.

BAKER, B.; TAYLOR, L. (ed.). *Language Assessment Literacy and Competence*: Case Studies from Around the World. Cambridge: Cambridge University Press & Assessment, 2024.

BRAUN, V.; CLARKE, V. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, v. 3, n. 2, p. 77-101, 2006.

CÁRDENAS, M.; GONZÁLEZ, A.; ÁLVAREZ, A. El desarrollo profesional de los docentes de inglés en ejercicio: algunas consideraciones conceptuales para Colombia. *Folios*, Bogotá, n. 31, p. 49-67, 2010.

FREEMAN, D.; JOHNSON, K. Reconceptualizing the Knowledge-Base of Language Teacher Education. *TESOL Quarterly*, v. 32, n. 3, p. 397-417, 1998.

FULCHER, G. Assessment Literacy for the Language Classroom. *Language Assessment Quarterly*, v. 9, n. 2, p. 113-132, 2012.

FULCHER, G. Operationalizing Assessment Literacy. *In*: TSAGARI, D. (ed.). *Language Assessment Literacy*: From Theory to Practice. Newcastle upon Tyne: Cambridge Scholars, 2020. p. 8-28.

GAN, L.; LAM, R. A Review on Language Assessment Literacy: Trends, Foci and Contributions. *Language Assessment Quarterly*, v. 19, n. 5, p. 503-525, 2022.

GIRALDO, F. *Fostering Pre-Service Teachers' Language Assessment Literacy*. Manizales: Editorial Universidad de Caldas, 2023.

GIRALDO, F. Language Assessment Literacy and Teachers' Professional Development: A Review of the Literature. *Profile*: Issues in Teachers' Professional Development, v. 23, n. 2, p. 265-279, 2021.

GIRALDO, F. *Language Assessment Literacy for In-Service Teachers*: Online Course Planning and Implementation. Manizales: Editorial Universidad de Caldas, 2024.

GIRALDO, F. Language Assessment Literacy: Implications for Language Teachers. *Profile*: Issues in Teachers' Professional Development, v. 20, n. 1, p. 179-195, 2018.

GIRALDO, F.; MURCIA, D. Language Assessment Literacy for Pre-Service Teachers: Course Expectations from Different Stakeholders. *GIST – Education and Learning Research Journal*, n. 16, p. 56-77, 2018.

GIRALDO, F.; VELÁSQUEZ-HOYOS, A.; VILLEGAS-LÓPEZ, L. Characterizing Teachers' Language Assessment Literacy to Design an Online Assessment Course. *Profile*: Issues in Teachers' Professional Development, v. 28, n. 2, 2026. In press.

GIRALDO, F.; YAN, X. (ed.). *Fostering Language Assessment Literacy*: Materials, Methods, and Emerging Issues. London: Routledge, 2025a.

GIRALDO, F.; YAN, X. Evaluating an Online Assessment Course: Teachers' Voices on Their Language Assessment Literacy. *Profile*: Issues in Teachers' Professional Development, v. 27, n. 1, p. 97-114, 2025b.

GREEN, A. *Exploring Language Assessment and Testing*. New York: Routledge, 2014.

GUBA, E. Criteria for Assessing the Trustworthiness of Naturalistic Inquiries. *Ectj*, v. 29, n. 2, p. 75-91, 1981.

HIDRI, S. (ed.). *Perspectives on Language Assessment Literacy*: Challenges for Improved Student Learning. London: Routledge, 2020.

INBAR-LOURIE, O. 'A Course of Our Own': Reflections Following a Language Assessment Literacy (LAL) Course. *In*: BAKER, B.; TAYLOR, L. (ed.). *Language Assessment Literacy and Competence*: Research and Reflections from the Field. Cambridge: Cambridge University Press & Assessment, 2024. v. 1, p. 231-238.

INBAR-LOURIE, O. Constructing a Language Assessment Knowledge Base: A Focus on Language Assessment Courses. *Language Testing*, v. 25, n. 3, p. 385-402, 2008.

INBAR-LOURIE, O. Language Assessment Literacy: What are the Ingredients? *In*: CBLA SIG SYMPOSIUM PROGRAMME, 4., 2013, Nicosia. *Proceedings* [...]. Nicosia: University of Cyprus, 2013.

KOH, K. *et al*. Developing the Assessment Literacy of Teachers in Chinese Language Classrooms: A Focus on Assessment Task Design. *Language Teaching Research*, v. 22, n. 3, p. 264-288, 2018.

KREMMEL, B. *et al*. Fostering Language Assessment Literacy Through Teacher Involvement in High-Stakes Test Development. *In*: XERRI, D.; VELLA BRIFFA, P. (ed.). *Teacher Involvement in High-Stakes Language Testing*. Cham: Springer, 2018. p. 173-194.

KUMARAVADIVELU, B. *Language Teacher Education for a Global Society*: A Modular Model for Knowing, Analyzing, Recognizing, Doing, and Seeing. New York: Routledge, 2012.

LAM, R. Language Assessment Training in Hong Kong: Implications for Language Assessment Literacy. *Language Testing*, v. 32, n. 2, p. 169-197, 2015.

LEVI, T.; INBAR-LOURIE, O. Assessment Literacy or Language Assessment Literacy: Learning from the Teachers. *Language Assessment Quarterly*, v. 17, n. 2, p. 168-182, 2019.

MERRIAM, S. *Qualitative Research and Case Study Applications in Education*. San Francisco: Jossey-Bass, 1996.

MOHAMMADI, M. Language Teachers' Assessment Literacy in AI-Aided Adaptive Learning Environments. *Journal of Research in Applied Linguistics*, v. 15, n. 2, p. 73-88, 2024.

MONTEE, M. *et al*. LCTL Teachers' Assessment Knowledge and Practices: An Exploratory Study. *Journal of the National Council of Less Commonly Taught Languages*, v. 13, p. 1-31, 2013.

MORAES, I. T. *A promoção do letramento em avaliação para a práxis de formadoras de professores/as de inglês*: implicações para a formação inicial e continuada. 2023. Thesis (Doctorate in Language Studies) – Universidade Estadual de Londrina, Londrina, 2023.

NIER, V.; DONOVAN, A.; MALONE, M. Increasing Assessment Literacy Among LCTL Instructors Through Blended Learning. *Journal of the National Council of Less Commonly Taught Languages*, v. 9, p. 105-136, 2009.

SCARINO, A. Developing Assessment Literacy of Teachers of Languages: A Conceptual and Interpretive Challenge. *Studies in Language Assessment*, v. 6, n. 1, p. 18-40, 2017.

SCARINO, A. Language Assessment Literacy as Self-Awareness: Understanding the Role of Interpretation in Assessment and in Teacher Learning. *Language Testing*, v. 30, n. 3, p. 309-327, 2013.

SHEEHAN, S.; THAI, T. LAL Training in Vietnam: A Teacher-Centred Approach to Resource Development. *In*: GIRALDO, F.; YAN, X. (ed.). *Fostering Language Assessment Literacy*: Materials, Methods, and Emerging Issues. New York: Routledge, 2025. p. 13-30.

SPILIOTOPOULOS, V. *et al*. Teacher Professional Learning Communities and Language Assessment Reform in Western Canada. *In*: BAKER, B.; TAYLOR, L. (ed.). *Language Assessment Literacy and Competence*: Case Studies from Around the World. Cambridge: Cambridge University Press & Assessment, 2024. v. 2, p. 8-19.

STABLER-HAVENER, M. Exploring the Effect of Embedding Assessments into a Professional Development Presentation as a Means of Improving Language Teachers' Assessment Competency. *In*: GIRALDO, F.; YAN, X. (ed.). *Fostering Language Assessment Literacy*: Materials, Methods, and Emerging Issues. New York: Routledge, 2025. p. 31-53.

TAYLOR, L. Communicating the Theory, Practice and Principles of Language Testing to Test Stakeholders: Some Reflections. *Language Testing*, v. 30, n. 3, p. 403-412, 2013.

TSAGARI, D.; ARMOSTIS, S. Contextualizing Language Assessment Literacy: A Comparative Study of Teacher Beliefs, Practices, and Training Needs in Norway and Cyprus. *Education Sciences*, v. 15, n. 7, p. 1-29, 2025.

VILLA-LARENAS, S.; BRUNFAUT, T. But who Trains the Language Teacher Educator who Trains the Language Teacher? An Empirical Investigation of Chilean EFL Teacher Educators' Language Assessment Literacy. *Language Testing*, v. 40, n. 3, p. 463-492, 2023.

VOGT, K. Pre-Service Teachers' Language Assessment Literacy Enhancement – Towards Pedagogies for LAL. *In*: GIRALDO, F.; YAN, X. (ed.). *Fostering Language Assessment Literacy*: Materials, Methods, and Emerging Issues. New York: Routledge, 2025. p. 54-76.

WENG, F.; SHEN, B. Language Assessment Literacy of Teachers. *Frontiers in Psychology*, v. 13, e864582, 2022.

YAN, X.; ZHANG, C.; FAN, J. "Assessment Knowledge is Important, but…": How Contextual and Experiential Factors Mediate Assessment Practice and Training Needs of Language Teachers. *System*, v. 74, p. 158-168, 2018.

YIN, R. *Case Study Research and Applications*: Design and Methods. 6. ed. London: SAGE Publications, 2014.

ZHANG, J.; YU, G.; BROWNE, W. Teachers' Language Assessment Literacy: Exploring its Construct and Contextual Factors. *Studies in Educational Evaluation*, v. 87, e101525, 2025.

History

Date of submission: 13/01/2026.
Date of approval: 14/04/2026.

Responsible Editor

Isadora Teixeira Moraes, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais/MG, Brasil. Lattes: http://lattes.cnpq.br/9478476109141850, ORCID: https://orcid.org/0000-0002-5393-286X, email: isadoramoraes@ufmg.br

Authors' Contributions

Giraldo: writing the initial research grant proposal for the Vicerrectoría de Investigaciones y Posgrados (Office of Research and Graduate Studies) at Universidad de Caldas in Manizales, Colombia; preparing research materials; collecting data; analyzing data; preparing draft; validating; writing and revising the article.

Velásquez: preparing research materials; collecting data; analyzing data; preparing draft; validating; writing and revising the article.

Villegas: preparing research materials; collecting data; analyzing data; preparing draft; validating; writing and revising the article.

Declaration of Conflict of Interest

The authors declare that there is no conflict of interest.

Data Availability Statement

Data cannot be made publicly available. When we, the three authors, described the research study to our participants, in the corresponding information and consent letter that all five participants signed, we made it clear that the data they were to provide in the study would only be accessed by the three of us. This was also made clear in the research proposal we wrote for the research office at our university.

Use of AI

The authors declared that they used AI to format all the bibliographical references from APA 7th edition into Brazilian citation standards. AI was not used anywhere else in the article.

Reviews

Review 1

O tema do manuscrito é relevante e se enquadra no escopo da revista. O aporte teórico é preciso, variado e atual. A metodologia é pertinente, detalhadamente descrita e há uma clara preocupação ética com os participantes da pesquisa. Os resultados são apresentados e analisados. O texto está adequadamente organizado e sua leitura é fluída. Ao analisar os impactos de um curso em letramento em avaliação, o manuscrito tem grande potencial de estabelecer procedimentos teórico-práticos para informar futuros cursos e pesquisas na área.

Reviewed on: 23/03/2026.
Recommendation: Accept submission.
Reviewer: Estogildo Gledson Batista.

Review 2

This study reports on the design, implementation, and impact of a context-sensitive language assessment course (LAiC) for five in-service English teachers in Colombia. Using a qualitative, multi-method case study design, the authors analyse data from observations, reflective logs, learning artifacts, and interviews. The findings suggest that the course contributed to teachers' language assessment literacy (LAL), particularly in terms of contextualised practice,

theoretical understanding, and professional attitudes. The study also proposes elements of a localised LAL pedagogy grounded in course design and participant engagement.

Theoretical background: The manuscript engages with relevant literature on LAL, particularly emphasising its contextual and sociocultural dimensions. It aligns with key scholars in the field and reinforces the need for context-sensitive LAL pedagogies. The attempt to conceptualise a "local LAL pedagogy" is a valuable contribution.

Methodology: The methodology is generally appropriate and well-structured for the study's aims. The use of a qualitative case study and multiple data sources enhances depth and triangulation. Identified strengths: clear description of instruments and procedures; inclusion of both teacher and instructor perspectives; Effort to ensure validity (e.g., piloting, expert review, inter-rater reliability). Some limitations: the small sample size (n=5) limits broader applicability, and the dual role of researchers as instructors raises potential bias, despite mitigation efforts.

Results and conclusions: The findings are coherent, well-organised, and supported by data excerpts. The identification of three main dimensions (impact, instructional design, and teacher attitudes) provides a clear analytical structure. The conclusions are consistent with the findings and emphasise the importance of contextualisation, the integration of theory and practice, and the role of collaboration and reflection.

Writing and organisation: The manuscript is generally well-organised and readable, with a logical progression across sections. Identified strengths: clear sectioning and structure, and effective use of excerpts to illustrate findings. Areas for improvement: presence of language issues (grammar, phrasing, and occasional awkward constructions), some redundancy, especially across findings and discussion, and proofreading and language editing.

Argumentation: The argument is coherent but could be strengthened analytically. While the study convincingly shows that the course had a positive impact, it tends to emphasise success without sufficiently problematising tensions or contradictions, and underexplore alternative interpretations of the data.

Relevance and originality: The study addresses a highly relevant topic in language assessment and teacher education, particularly in the Global South context. Its originality lies in the tailor-made course design, in the inclusion of instructor perspectives, and in the focus on context-sensitive LAL development.

Contributions to the field: the study offers practical and pedagogical contributions, including insights into designing context-sensitive LAL courses, identification of key features (tasks, materials, interaction), and emphasis on teacher agency and collegiality. These contributions are valuable, especially for teacher educators and program designers.

Reviewed on: 02/04/2026.
Recommendation: Mandatory corrections.
Reviewer: Gladys Quevedo.

Appendix A – Non-Participant Observation Form

Observer's name:
Date of the observation:
Topic of the session:
Number of participants that attended the session:
Session's activities:

Category	Observation Notes (Possible influence on teachers' LAL)	Evidence/Examples (Participants explain or state something related to the category)
Participants' engagement with course content		
Input		
Roles		
Interaction with peers and course instructor		
Reflection on LAL (e.g., Teachers' challenges or barriers to enact LAL)		
Outcomes and takeaways		
Feedback		
Emerging category		

Source: Adapted from Fulcher (2020) and Giraldo (2024).

Observer's Insights to Contribute to LAL Pedagogies

Appendix B – Interview Protocol

Opening question:

How would you evaluate the LAiC? What is your general opinion about it?
What did you like? What did you not like? Why?

Questions:

(1) Are you implementing / Have you implemented anything from the course?

- If yes: What have you implemented?
- If no: Why not?

(2) Did the course contribute to your approach to language assessment? If so, how? If not, why?
- If the teacher asks what approach to assessment means:
    - Your assessment knowledge, beliefs, practices, skills, principles, attitudes, etc.

(3) Are there any changes in your approach to language assessment after having taken the course?

- What about your beliefs regarding language assessment in **your context**?
- What about your practices regarding language assessment in **your context**?

(4) Have you found the course useful or not useful to your assessment context?

- If yes: In what ways have you found it useful?

- If not: Why was the course not so useful to your assessment context?

5) If any, what is something from the course that you would like to implement in your language assessment context?

Questions based on the ideas from the teacher's reflective log:
...