

Evaluating an Online Assessment Course: Teachers' Voices on Their Language Assessment Literacy

Análisis de un curso en línea sobre evaluación: voces de docentes sobre su literacidad en evaluación de lenguas

Frank Giraldo

Universidad de Caldas, Manizales, Colombia

Xun Yan

University of Illinois at Urbana-Champaign, USA

This article reports a mixed-methods case study on English language teachers' evaluation of an online assessment course and their language assessment literacy. Our goal was to examine the teachers' perceptions of the contents, activities, and the impact of the course on their professional development. For data collection, we used a questionnaire and a focus group interview. Results indicate that teachers welcomed all course contents, especially for assessing learners with special needs and integrated-skills assessment. Regarding activities, the teachers favored test analysis and construction while seeing statistics as irrelevant to their contexts. In terms of professional development, teachers reported reflection on prior assessments and improvements for future assessments. In conclusion, the online course contributed to teachers' knowledge and skills in language assessment.

Keywords: assessment literacy, language assessment literacy, language testing, teachers' professional development

Este estudio de casos mixto analizó un curso de evaluación en línea y su impacto en la literacidad evaluativa de docentes de inglés. A través de un cuestionario y una entrevista grupal, se exploraron percepciones sobre contenidos, actividades y su contribución al desarrollo profesional. Los profesores valoraron la evaluación de estudiantes con necesidades especiales y la evaluación integrada, destacando actividades de análisis y diseño de instrumentos, mientras que consideraron irrelevante la estadística para sus contextos. En general, el curso fortaleció sus conocimientos y habilidades en evaluación de idiomas.

Palabras clave: desarrollo profesional docente, evaluación de lenguas extranjeras, literacidad en evaluación, literacidad en evaluación de lenguas.

Frank Giraldo  <https://orcid.org/0000-0001-5221-8245> • Email: frank.giraldo@ucaldas.edu.co

Xun Yan  <https://orcid.org/0000-0002-4544-4938> • Email: xunyan@illinois.edu

This project was sponsored by Fulbright Colombia, thanks to the "Visiting Colombian Scholar/Researcher" scholarship granted to Giraldo.

How to cite this article (APA, 7th ed.): Giraldo, F., & Yan, X. (2025). Evaluating an online assessment course: Teachers' voices on their language assessment literacy. *Profile: Issues in Teachers' Professional Development*, 27(1), 97–114. <https://doi.org/10.15446/profile.v27n1.113304>

This article was received on March 4, 2024 and accepted on October 9, 2024.

This is an Open Access article distributed under the terms of the Creative Commons license Attribution-NonCommercial-NoDerivatives 4.0 International License. Consultation is possible at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Introduction

In language testing, there has been a major interest in researching language assessment literacy (LAL) across stakeholders. As language assessments, large-scale or local, can impact education and society (McNamara & Roever, 2006; Shohamy, 2017), it is in the interest of stakeholders to develop the necessary knowledge, skills, and principles for language assessment. To date, LAL research has focused on describing what LAL is, proposing models and frameworks to conceptualize and problematize the construct (Fulcher, 2012; Xu & Brown, 2016). Additionally, there is a growing focus on describing LAL among stakeholders involved in assessment (Inbar-Lourie, 2013; Taylor, 2013) and how their LAL develops (Baker & Riches, 2018; Yan & Fan, 2021).

Given their engagement in assessment, teachers are a major stakeholder group in LAL research, both from descriptive and pedagogical lenses. Generally, research about teachers' LAL has indicated that they feel underprepared for conducting quality assessments and has suggested hands-on approaches in LAL programs, namely courses or workshops (Fulcher, 2012; Gan & Lam, 2022; Vogt & Tzagari, 2014).

Despite the few LAL courses reported in the literature (for reviews, see Giraldo, 2021 and Gan & Lam, 2022), pedagogical experiences in LAL seem to be gaining momentum internationally. The research on LAL courses has indicated that teachers' assessment experiences contribute to developing their assessment literacy (Arias et al., 2012; Baker & Riches, 2018; Giraldo & Murcia Quintero, 2019). Furthermore, LAL courses with strong hands-on approaches contribute to teachers' theoretical knowledge and design skills (e.g., creating tasks and rubrics for assessing language ability). Most importantly, LAL courses seem most successful when contextualized and based on teachers' feedback, and LAL needs prior to implementation (Bøhn & Tzagari, 2021; Giraldo & Murcia Quintero, 2019).

Although the research on LAL courses has indicated their positive impact on teachers' professional development, much remains to be learned about what fosters LAL in these courses and how this happens. In this paper, we consider that an LAL course is a pedagogical space in which contents, materials, and activities aim to help teachers develop their LAL.

As a background to our study, there is limited research on teachers' LAL development in Latin America. The studies we report were conducted in Brazil, Colombia, and Haiti. We acknowledge, however, that there may be studies we could not find in our literature search. The existing research has remarked on how LAL training helps pre- and in-service teachers to critically reflect on their practices (Arias et al., 2012; Baker & Riches, 2018; Giraldo & Murcia Quintero, 2019; González, 2021; Quevedo-Camargo & Tonelli, 2021; Restrepo Bolivar, 2020). Additionally, like other research, reported LAL training in Latin America suggests that teachers improve assessment knowledge and skills while engaged in test design (Arias et al., 2012; Fernandes & Borges-Almeida, 2021; Giraldo & Murcia Quintero, 2019).

To draw LAL training in this region, Villa-Larenas et al. (2021) surveyed stakeholders' LAL. Of 532 respondents, 340 (66 %) were language teachers. The authors welcome LAL training that includes the connection between assessment, teaching, and learning; the assessment of learners with special educational needs (SEN); and the impact of assessment on society.

Responding to Villa-Larenas et al.'s (2021) findings, in this article, we report on the implementation of an online language assessment (OLA) course for English language teachers in four Latin American countries: Brazil, Colombia, Peru, and Venezuela. We report how these teachers evaluated the OLA course, including its contents, activities, and impact on LAL development. In doing so, we contribute to Fulcher's (2020) call for LAL pedagogies through insights into what seems to work and what does not in LAL courses. We believe

our paper is significant as it responds to three inquiry areas: (a) a contribution to pedagogical approaches to LAL development (Fulcher, 2020; Gan & Lam, 2022); (b) the reported need for LAL training in Latin America, especially classroom-based assessment (Villa-Larenas et al., 2021); and (c) the limited representation of LAL research in this region (Gan & Lam, 2022). We hope to highlight how teachers' voices played a vital role in course implementation and evaluation. Our research question was: What are English language teachers' perceptions of the contents and activities in an OLA course and its impact on their professional development?

Literature Review

Teachers' Language Assessment Literacy

LAL is usually conceptualized as the knowledge, skills, and principles needed to conduct and/or understand various assessment-related activities (Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2013). This broad conceptualization has been used to propose or research LAL across stakeholders (Kremmel & Harding, 2020; Taylor, 2013). However, these three major components differ across stakeholders involved in assessment and represent fluid continua rather than fixed states (Inbar-Lourie, 2013; Pill & Harding, 2013). Stakeholders require different LAL profiles, depending on the contexts where they conduct assessment and the tasks in which they engage.

In the case of language teachers, scholars have emphasized the need to draw and interpret LAL in the teachers' context and their LAL development (Inbar-Lourie, 2017; Scarino, 2013; Yan et al., 2018; Yan & Fan, 2021). As Scarino (2013) argues, understanding teachers' assessment lifeworlds is a key condition to foster their professional development in this area of their practice. More recently, in fact, scholars have made the call that teachers' LAL should not be studied from a deficit perspective, that is, teachers' lack of assessment literacy

(Baker, 2021; Yan, 2021). Rather, language teachers contribute to LAL discussions by bringing their beliefs, experiences, and expertise to the forefront of LAL training (Arias et al., 2012; Baker & Riches, 2018).

Regarding language teachers' beliefs and practices in assessment, research has suggested two interrelated aspects reflecting LAL development. On the one hand, teachers with training in language assessment tend to have positive perceptions of and usually conduct sound assessment practices (Arias et al., 2012; Baker & Riches, 2018; López Mendoza & Bernal Arandia, 2009). Such practices, as reported in the literature, tend to align mostly with theoretical and practical assessment guidelines, thus reflecting LAL's knowledge and skills. On the other hand, teachers with no training may perceive assessment negatively and resort to summative-only, traditional practices (López Mendoza & Bernal Arandia, 2009; Sultana, 2019). Finally, while researchers acknowledge that teachers' beliefs and practices in assessment may be difficult to change, the literature clearly indicates the positive impact of LAL training (Giraldo, 2021; Xu & Brown, 2016).

Against this background, teachers' LAL includes knowledge, skills, principles, as well as their contextual beliefs and practices, to understand and conduct assessment (Davies, 2008; Fulcher, 2012; Scarino, 2013; Taylor, 2013). Knowledge in LAL refers to language use and language ability frameworks, concepts, and a variety of assessment methods. Skills involve the design of assessments and the interpretation of the data they yield. Finally, principles include ethical and fair uses of assessment. Although drawing the construct of LAL for teachers is still in discussion, when designing LAL training, it needs to be contextualized into concrete activities and practices that reflect these three components. To do so, teachers' voices and knowledge from the local context are needed to help them identify assessment purposes, design context-specific assessment tasks, and create a communication network among stakeholders.

Operationalizing Teachers' LAL

For our research and course design, we synthesized authors' ideas on what teachers' LAL might entail (Brindley, 2001; Fulcher, 2012; Giraldo, 2021; Kremmel & Harding, 2020; Scarino, 2013; Taylor, 2013). Given its distinct contextual nature, we note that the components below represent empirical and conceptual discussions and cannot be considered a checklist to describe and/or evaluate teachers' LAL (Csépes, 2021; Xu & Brown, 2016).

- **knowledge:** language ability frameworks, as operationalized in curricula; purposes (assessment of, for, and as learning), methods, and skills to be assessed; central concepts and their use in practice (e.g., validity, reliability, and authenticity)
- **skills:** analyzing and adapting existing assessments; designing and evaluating sound assessments; implementing assessments successfully; using assessments to positively impact teaching and learning
- **principles:** critiquing the uses and misuses of assessments; using assessment ethically and fairly; implementing transparent and democratic assessment practices; analyzing the intended and unintended consequences of assessment
- **assessment context:** knowledge of learners, beliefs, practices, challenges, needs, and skills; institutional assessment policies and practices

Courses to Foster Teachers' LAL

Fulcher (2020) has urged that LAL discussions move from description to action—what he calls “a pedagogy of LAL.” Furthermore, as scholars have stated, LAL training must be based on a sound understanding of teachers' contexts and what they bring to the table (Bøhn & Tzagari, 2021; Scarino, 2013). In our review of published research, we found that successful LAL courses are needs-sensitive (see Arias et al., 2012; Baker & Riches, 2018; Koh et al., 2018; Montee et al., 2013). Thus, teachers' voices are fundamental to LAL course

planning and, eventually, implementation. Additionally, previous research reports that assessment courses for language teachers tend to have these similarities:

- **An explicit connection between theory and practice:** Teachers study language assessment theories and concepts to analyze or design assessments in context (Arias et al., 2012; Giraldo & Murcia Quintero, 2019; Janssen, 2022; Koh et al., 2018; Levi & Inbar-Lourie, 2020; Walters, 2010).
- **A hands-on approach to test analysis and development:** Teachers are engaged in analyzing the usefulness of assessment instruments and then design assessments based on technical considerations (Baker & Riches, 2018; Janssen, 2022; Kremmel et al., 2018; Montee et al., 2013).
- **A growing reflection on assessment:** Although the reported LAL courses are mostly based on knowledge and skills, an outstanding impact on teachers is how they reflect on their prior assessment practices and realize the need to improve the status quo (Arias et al., 2012; Baker & Riches, 2018; Janssen, 2022).

In terms of how training courses operationalize the LAL construct, Giraldo (2021) found that there is a clear allusion to *knowledge* and *skills* and secondary attention to *principles*; in other words, training courses focus on theoretical aspects such as purposes, assessment methods, and concepts such as validity, reliability, and authenticity. However, these courses do not deeply study issues such as fairness and ethics in language teachers' assessment approach. Giraldo also reports that, in terms of assessment skills, the courses focus on the design of rigorous assessment instruments and their relationship with teaching and learning.

The existing studies reporting on teachers' evaluations of LAL courses suggest that both contents and activities contribute to their LAL and, most importantly, help them to improve their assessment

approach. However, when evaluating the outcomes of LAL courses, previous studies tend to utilize the researcher's evaluation of teachers' LAL competencies rather than focusing on how teachers self-assess and perceive their own LAL development (Kremmel et al., 2018; Papadima-Sophocleous, 2022). In the present study, we sought to collect teachers' feedback (i.e., their voices) on their LAL development as influenced by the OLA course.

Method

We used a case-oriented, mixed-methods design (Ivankova & Greer, 2015) to examine the effectiveness of a research-informed LAL course in Latin America. Such design allows for a better understanding of the research focus from complementary perspectives (Creswell & Creswell, 2018). The qualitative data in our study came from a focus group interview, and the quantitative data came from an end-of-course questionnaire in which teachers evaluated the course and their LAL. Specifically, we adopted a convergent parallel design where data collection and analysis of both quantitative and qualitative data co-occur and are analyzed separately. Then, we triangulated the findings from both quantitative and qualitative analyses to provide a cross-validated interpretation of teachers' perceived LAL development during the course.

Context of the Study

Fulbright Colombia fully sponsored the OLA course thanks to a scholarship called Visiting Colombian Researcher, of which the first author was the recipient. This scholarship allows Colombian researchers to visit a university in the United States and conduct research in various areas. The first author's proposal involved designing and implementing the OLA course, which was designed for English language teachers in Latin America (for details on the design stage, see Giraldo & Yan, 2023). We used three selection criteria:

teachers working at state high schools; teaching the English language; and working in Brazil, Colombia, Ecuador, Peru, or Venezuela.

To participate in the OLA course, we sent a call for applicants to a language testing practice emailing list, social and academic networks, the Latin American Association for Language Testing and Assessment, and the International Research Foundation for English Language Education. Fifty teachers could participate in this free online course; 20 applied and participated in the diagnostic stage. From this group, 19 teachers started the course, and 13 finished.

Participants

Thirteen English language teachers participated in the OLA course: Brazil ($n = 3$), Colombia ($n = 4$), Venezuela ($n = 1$), and Peru ($n = 5$). They taught English at the high school level in state-funded institutions. The teachers taught general English courses from elementary up to intermediate proficiency levels. The diagnostic stage of our research (i.e., design) indicated that the course should address teachers' challenges in assessment; include spaces for discussion and interaction with colleagues; target knowledge, skills, and principles in assessment; and discuss current issues in assessment (e.g., bilingual assessment and the assessment of learners with SEN).

Materials, Instruments, and Procedures

Online Language Assessment Course

The project sought to promote English teachers' LAL through a free online course. To plan and deliver the course, we first asked teachers for feedback on topics and skills to be studied. Based on this needs analysis, we designed and taught the course, which formed the basis for the present report. The OLA course lasted ten weeks, from September 6th to November 10th, 2022.

There were 10 modules in the OLA course; for each module, the teachers were expected to read contents from a handbook designed for the course and do a couple of activities about assessment. Below are the 10 modules in the course with some sample contents:

Week 1. Introduction to fundamentals of language assessment: Purposes, methods, and constructs

Week 2. Qualities of language assessment: Validity, practicality, and washback

Week 3. Assessing receptive skills: Real-life reading/listening purposes; multiple-choice and true-false items

Week 4. Assessing productive skills: Construct definition and rubric design

Week 5. Assessing integrated skills: Task-based assessment

Week 6. Bilingual assessment: Translanguaging in assessment

Week 7. Assessing learners with special educational needs (SEN): Accommodations for language assessment

Week 8. Alternative assessment: Self- and peer-assessment; portfolio assessment

Week 9. Ethics and fairness in classroom language assessment: Principles for ethical and fair assessment practices

Week 10. Relating language assessment to language teaching and learning: Positive washback; assessment culture

The OLA course was taught online through Google's Workspace technology: Drive, Docs, Slides, Spreadsheets, and Forms. For small-group discussions, the course utilized Zoom's breakout rooms. There were two synchronous weekly workshops, each lasting two hours. The course included the following activities and tasks: reading contents in the handbook, participating in interactive lectures, participating in small-group discussions about assessment situations, analyzing

and designing assessments, and providing feedback to peers on their assessments.

Questionnaire

The questionnaire, with 41 closed-ended and six open-ended items, was based on OLA course components. The first section asked participants to evaluate the course in general, materials, activities, and the technology used for course delivery; the scale for this section was 1 = *poor* to 5 = *excellent*. In section two, the teachers rated the usefulness of course contents in relation to their assessment practices. In section three, the teachers evaluated the usefulness of course activities in learning about language assessment. The scale for sections three and four was 1 = *not useful at all* to 5 = *extremely useful*. Finally, in section four, the teachers stated their agreement (1 = *strongly disagree* to 5 = *strongly agree*) on how the course impacted their professional development.

Focus Group Interviews

For additional feedback on course contents, activities, and their impact on teachers' professional development, the second author conducted three focus group interviews: two with four teachers and one with five. For each interview, he used a protocol that included the following questions:

1. What do you think about the materials you have used so far?
2. What do you think about the contents you have studied in this course?
3. What do you think about the synchronous workshops and their activities?
4. If any, in what ways has the course impacted your professional development?

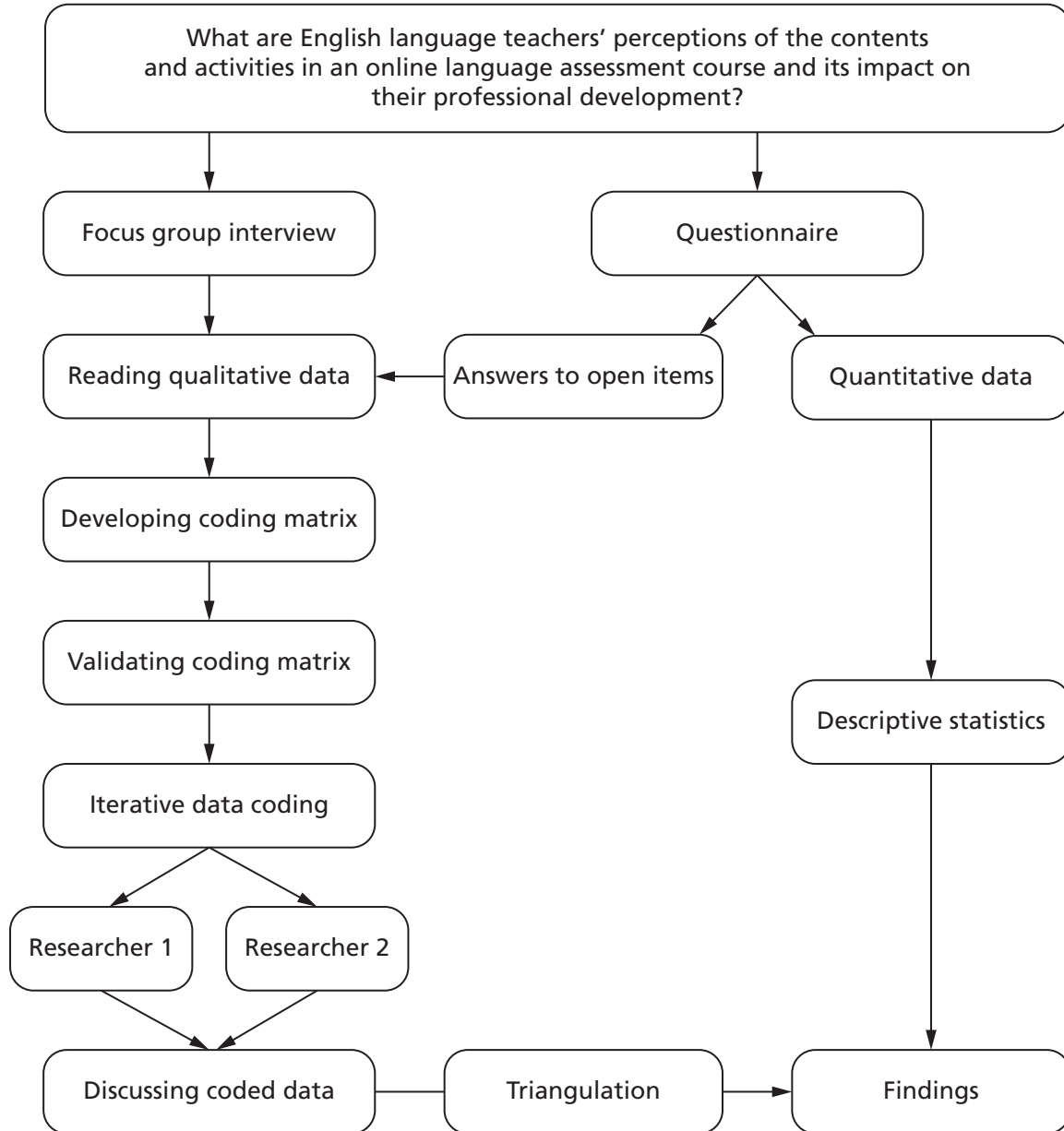
Both data collection instruments were verified through content expert analysis from six teacher educators with experience in language assessment. The experts considered both instruments fit for purpose, although

some changes were required. We modified the closed items in the questionnaire, making sure they addressed the usefulness of course contents; how activities helped teachers to learn about assessment; and course impact on their professional development.

Data Analysis

Figure 1 shows the process we followed to analyze and triangulate the data in this study, to answer our research question.

Figure 1. Data Analysis Procedures



To analyze the qualitative data in the interview and the questionnaire, we used Saldaña's (2016) coding model for qualitative analysis: We started with codes and ended with themes. According to this author, such coding goes from concrete data samples to abstract themes that form commonalities across data samples. After reading over the qualitative data, the first author developed the coding matrix, which was checked by the second author; we then used it independently to code the data, making the adaptations we deemed necessary (e.g., grouping categories further). Then, we shared our analyses to look for commonalities and discrepancies: For 12 categories in the five major themes, Cohen's Kappa was 0.75, which suggests substantial agreement. Where discrepancies occurred, we discussed them to reach an overall agreement for the categories. We conducted this analysis iteratively as we read and refined codes and categories and finally arrived at themes. The Appendix presents our coding scheme. Additionally, we used descriptive statistics (mean, median, and range) to describe the numerical data in the questionnaire. Finally, following our research approach, we grouped data from the interview and the questionnaire to triangulate and arrive at findings that could answer our research question.

Results

In line with our research question, we will present significant findings related to the teachers' perceptions in four areas: the course in general, course contents, course activities, and the impact of the course on their professional development. Following the mixed-methods approach, we will use data from both the focus group interview and the questionnaire to illustrate findings and evaluate the implementation of the OLA course. In all data pieces, we use codes to refer to our participants (e.g., Teacher12).

Teachers' Perceptions of the OLA Course

Data in the focus group interview and the questionnaire clearly indicated that the participating teachers evaluated the course and its components positively. The teachers commented positively on how the course was delivered and emphasized the usefulness of the course for their practice in assessment. For example, Teacher7 comments on how the course has helped them to view assessment differently and align it with professional practice:

Interviewer: So, our first question is, how do you feel about the course so far?

Teacher7: I think the course has been very helpful because we came from seeing assessing as maybe a tool to get grades, and just maybe give reports to parents, to a tool to make students more aware of learning so it's not just the process, not just the final step of giving students like a test or, well, evaluating them, but just starting from very basic things, like from setting the goals for the course, making them aware of the goals they will achieve, and how their performance will be evaluated, like all the things that evaluation has, so it's been very helpful.

Table 1 shows teachers' general impressions of the course, materials, activities, and the platform through which the course was delivered. The large means and small ranges reiterate the teachers' positive evaluation of these general aspects. The slightly larger range of 2 in Item 3 may reflect teachers' opinions on statistics in the course, to which we will return later.

Teachers' Perceptions of Contents in the OLA Course

All teachers considered the topics in the course to be useful for their assessment practice in their institutions. Particularly, they emphasized how the course contents were organized and sequenced. As they explained, this organization was useful for them in learning about

Table 1. Teachers' Overall Perceptions of the OLA Course

Items	Mean	Median	Range
1. Overall, how would you rate this language assessment course?	4.9	5	1
2. How would you rate the usefulness of the materials you used to learn about language assessment?	4.8	5	1
3. How would you rate the usefulness of the activities for you to learn about language assessment?	4.5	5	2
4. How would you rate the platform used to administer the course (Google's Workspace)?	4.8	5	1

language assessment. Teacher17 commented on the topic sequence and how this was something she expected to have in the course:

Interviewer: What do you think about the contents of the course?

Teacher17: Well, I really like the sequence of the topics because they are related and connected, and all of them are really good for me because that's the thing that I was looking for at the beginning of the course. So, everything is new.

On the other hand, teachers' answers about preferred contents in the courses varied; however, answers in the interview and the questionnaire conclusively showed teachers' interest in two content topics: assessing learners with SEN and integrated-skills assessment. For instance, Teacher19 stated that she had never addressed the topic of assessing learners with SEN but came to value them as she was engaged in an activity that explored this content:

Another topic is assessing learners with special educational needs. I never have idea [sic] how to, or what kind of strategies, or tech, or methods I can use to assess this kind of students, for example. The first author brought two teachers from Colombia who shared with us their experiences, so it led me to value these kinds of students.

In an open item from the questionnaire, Teacher12 states what topic she found most useful and connects

it to her assessment practice: "The most useful course content is assessing integrated skills because the learners learn and practice with their real life, motivating [sic] and the washback is more effective."

Table 2 shows the results of asking teachers about the usefulness of contents in the OLA course. As demonstrated by the large means and narrow ranges, teachers found most topics extremely useful for assessment in their teaching contexts.

Items 16, 17, and 20 have a slightly wider range. Item 16 (*statistics for classroom assessment*) may have gotten the smallest mean because, as teachers explained, this content was not so relevant to their current assessment practices. The questionnaire asked teachers what content they found least useful, to which Teacher 5 answered, "Statistics for classroom assessment because it is not commonly used on a daily basis." In addition, some teachers stated that bilingual assessment was not so useful for the same reason Teacher5 expressed regarding statistics. On the other hand, there is no evidence in the data to describe or explain the range of 2 in Item 20: Ethics and fairness in classroom language assessment.

Teachers' Perceptions of Activities in the OLA Course

Like their perceptions about course contents, all teachers in this study expressed that they found its activities useful, specifically for learning about language assessment. The teachers commented on how the

Table 2. Teachers' Perceptions of Usefulness of Contents in the OLA Course

Please state to what extent you found the following contents useful for your language assessment practices (in your school).			
Items	Mean	Median	Range
9. Qualities of language assessment: validity, reliability, authenticity, practicality, and washback	4.9	5	1
10. How to assess receptive skills: listening and reading	4.9	5	1
11. How to assess productive skills: speaking and writing	4.9	5	1
14. How to create test specifications	4.9	5	1
21. Connecting language assessment to language teaching and learning	4.9	5	1
7. The relation between the three key questions in language assessment: the why, the how, and the what	4.8	5	1
8. Key concepts in language assessment: summative, formative; traditional, alternative; test, assess, evaluate	4.8	5	1
12. How to assess integrated skills	4.8	5	1
13. How to analyze existing assessment instruments	4.8	5	1
15. Task-based assessment	4.8	5	1
19. Alternative assessment	4.8	5	1
18. Assessing learners with special educational needs	4.6	5	1
17. Bilingual assessment	4.5	5	2
20. Ethics and fairness in classroom language assessment	4.4	4	2
16. Statistics for classroom assessment	4.2	4	2

activities in the synchronous workshops of the course fostered interaction among themselves, which in turn led them to learn from each other. The interactive and collaborative nature of the course, as the excerpt below suggests, had a positive impact on teachers' LAL.

Teacher10: The most important part of this group is that we are from different parts of Latin America, and to have this opinion of evaluation, assessment, and how they teach in their countries, it was really awesome for me because I can have all the things that they can do in order to fulfill this topic, this assessment.

For fine-grained feedback to evaluate course activities, the questionnaire asked teachers about their usefulness. Table 3 shows the descriptive statistics for 12

major activities the teachers conducted in the course. All means were large and ranges narrow, except for Item 33 (*Calculating and interpreting basic statistics for classroom language assessment*), with a mean of 4 and a range of 3. This feedback confirms that many of these teachers did not seem to find this activity useful in the course.

Table 3 also shows that all 13 teachers found Activity 29 (*Analyzing assessment instruments*) extremely useful for learning about language assessment. When asked about activities they found most and least useful in the open-ended items of the questionnaire, we saw the same trend we just described. In the open items, the teachers could explain why they had these perceptions about the two activities.

Teacher₃ (most useful activity): The most useful activity for learning is analyzing instruments because it is very important to know why: diagnostic, progress, and achievement and what is the method [*sic*] according to the skills: listening, writing, etc.

Teacher₉ (least useful activity): Calculating and interpreting basic statistics because my country's curriculum is based on the formative assessment [*sic*] and most of the time, we use symbols and letters to give a learning process result.

Another aspect that became apparent in the questionnaire was the perceptions teachers had about discussions. As mentioned above, teachers found interaction with others useful to learn about each other's teaching and assessment contexts. This is confirmed in the large mean and low range in Item 26 in Table 3. However, Item 27 had a smaller mean and slightly

wider range, which suggests teachers had differing views on small-group discussions. A closer look at the data in the open items of the questionnaire showed that four teachers found small-group discussions to be the least useful. As Teacher₁₁ states, "Working with colleagues in small groups did not promote a lot of interaction, in my opinion. Teachers were quiet most of the time and only one or two engaged." Based on these data, it seems like the teachers in the OAC were more willing to participate in whole-group discussions than small-group ones.

Teachers' Perceptions of the Impact of the OLA Course on Their Professional Development

All the teachers commented that the course positively impacted their professional development regarding their LAL. The teachers explored how

Table 3. Teachers' Perceptions of Usefulness of OLA Course Activities to Learn Language Assessment

Items	Mean	Median	Range
29. Analyzing assessment instruments	5	5	0
24. Interacting with the handbook: reading its contents	4.8	5	1
28. Listening to and interacting in lectures given by the course instructor	4.8	5	1
32. Sessions to provide and receive feedback on assessments you created in the course	4.8	5	1
26. Participating in whole-group discussions	4.7	5	1
34. Discussing a bilingual assessment mode in your English classes	4.7	5	1
35. Discussing unethical and/or unfair practices in language assessment	4.6	5	2
25. Doing activities in the handbook	4.5	5	2
27. Participating in small-group discussions	4.5	5	2
30. Planning the design of assessment instruments with other teachers	4.5	5	2
31. Designing assessment instruments for language skills with other teachers	4.5	5	2
33. Calculating and interpreting basic statistics for classroom language assessment	4	4	3

the course taught them about language assessment, impacted their assessment practice, and raised their awareness of what language assessment implies. Teacher18 commented on how she can better assess communicative competence and what kind of assessments she can use to do so.

Interviewer: How do you feel about the course so far?

Teacher18: I improve [*sic*] my knowledge related to how to assess the communicative competence of my students. It means that construct for [*sic*] language skills and what kind of instrument we can use to know the learning progress of our students. Another thing that I improve to recognize or identify [*sic*] in order to differentiate, for example, how can we evaluate the task of students, and what kind of method we can use.

Although the positive impact of a course is expected in such an educational scenario, we want to emphasize two areas that were apparent across the data. On the one hand, the teachers commented that the course ignited serious reflection about their past language assessment practices. In particular, the teachers reflected on mistakes they made in assessment and now need to improve. Teacher20 commented on how her professionalism in assessment evolved thanks to the OLA course:

I realized that I was applying some methods, some strategies, some techniques, and some instruments, also, but it was not in a professional way. It was empiric or just part of my teaching practice. Now that I'm going to be more professional with my teaching practices also.

On the other hand, the teachers explained how, thanks to the course and their reflections, they felt the need to improve assessment practices to better serve their students. The answers in the interview suggest that teachers see professional assessment practices as fair. For instance, Teacher15 stated how she was not fair in assessment but hopes to change her practices so they are fair in future assessment experiences:

It somehow makes me a better teacher. I think it will help me be more fair to my students. I wasn't quite fair before with some of the things I was asking from them. They were not very clear, maybe, or something like that. So, that is something that I will improve from now on, I hope.

The data in Table 4 reinforces the finding that the OLA course positively impacted teachers' professional development. The data suggest that the teachers strongly agreed with the positive outcomes of the course, as shown in Item 38. Further, the impact of the course reflected LAL's three overall components: knowledge (Item 40), skills (Items 41 and 42), and principles (Item 43). Item 44 shows how all 13 teachers strongly agreed that the course led them to reflect on areas for improvement in their LAL, which is evident and aligns with the interview sample above.

Suggestions for Improving the OLA Course

Notwithstanding the overall positive evaluation that teachers had of the course, in the data, we noticed an area that the course could have improved. Teacher15 and Teacher20 commented that it was challenging for them to move from one technological tool to another (e.g., from Google Meet to Zoom). As Teacher20 comments, this was exacerbated by their slow internet connection: "I would like to have the activities in the same platform, to be honest, because we just [used] Google Meet, and we had to go different links. It was more difficult for me because of my Internet connection." Thus, we believe the use of technological apps, especially for conferencing, must be considered carefully when offering online assessment courses.

Discussion

As with previous research, our study has indicated language teachers' favorable opinions of language assessment courses, especially regarding the connection between assessment and teaching (Kremmel et al., 2018;

Table 4. Teachers' Perceptions of How the OLA Course Impacted Their Professional Development

Please decide to what extent you agree or disagree with the following statements. (1 = strongly disagree to 5 = strongly agree)			
Items	Mean	Median	Range
44. While being in this course, I became aware of aspects I need to improve in language assessment.	5	5	0
38. Overall, I feel this course impacted my professional development positively.	4.9	5	1
39. The course fostered my knowledge, skills, and principles in language assessment.	4.9	5	1
41. Thanks to the course, I am better at analyzing assessment instruments professionally.	4.9	5	1
42. Thanks to the course, I am better at designing assessment instruments professionally.	4.9	5	1
43. Thanks to the course, I am more aware of unethical or unfair practices in language assessment and their impact.	4.9	5	1
47. Thanks to this course, I feel now I am more prepared to assess my students well.	4.9	5	1
40. The course made me more knowledgeable of general theories and concepts in language assessment.	4.8	5	1
45. Because of this course, I feel better prepared to discuss language assessment issues in my institution.	4.8	5	1
46. After this course, I think I can better connect language assessment to language teaching and learning.	4.8	5	1

Montee et al., 2013). The general positive perception of the teachers in our study may be attributed to the localized nature of the course, which depended heavily on their feedback for planning and implementation, as various scholars have encouraged in LAL training (Bøhn & Tsagari, 2021; Scarino, 2013; Yan & Fan, 2021). Thus, an in-depth needs analysis designed for the local context could precede successful assessment courses for teachers.

Other research has shown that teachers generally have a positive perception of contents in LAL training (Baker & Riches, 2018; Kremmel et al., 2018), with our study emphasizing the extent to which OLA course contents were useful for the teachers' assessment context. Further, the findings in our study corroborate those

by Villa-Larenas et al. (2021) regarding teachers' LAL in Latin America: Assessing SEN learners seems to be an assessment need and/or interest among language teachers. We argue that this topic, then, may merit further attention in LAL research and pedagogy, especially as it does not seem prominent in current discussions around teachers' LAL.

Regarding activities to promote LAL, other studies have reported teachers' positive perceptions of the activities in the assessment courses or workshops in which they participate (Baker & Riches, 2018; Kremmel et al., 2018; Papadima-Sophocleous, 2022). Our findings provide empirical evidence for the usefulness of test analysis as a core activity or task in language assessment training (Arias et al., 2012; Baker & Riches,

2018; Giraldo, 2021; Koh et al., 2018). On the other hand, the teachers in our study did not find calculating and interpreting statistics useful perhaps because these were not common activities in their assessment lifeworlds. Considering these results, our study contributes to LAL pedagogies for teachers as we gathered fine-grained data and feedback (see Table 3) on the relationship between course activities and teachers' development of LAL.

Finally, previous studies have shown how LAL training contributes to teachers' conception of what assessment involves (Arias et al., 2012; González, 2021; Montee et al., 2013). Such findings may not be surprising and, in fact, expected. However, in our study, teachers reflected on their LAL, identified past mistakes in their practice, and reported willingness to change their assessment approach. This finding further supports the idea of teachers gradually forming their "assessor" identity as they learn about assessment during the course (Pastore, 2023; Xu & Brown, 2016). The configuration of LAL's theoretical, technical, and critical aspects (see Items 41 to 43 in Table 4) fostered teachers' consideration of how they approach(ed) assessment in their contexts.

Limitations

Given that our research is a case study in a specific geographic and educational context, our findings cannot be directly extrapolated to wider settings. However, they contribute to the growing interest in problematizing pedagogies for LAL at large. Thus, we hope the findings are analyzed in relation to LAL courses elsewhere, especially those delivered online. Another limitation of this study is that, even though we addressed the ethics and fairness angle of LAL in the questionnaire (see Item 43 in Table 4), we did not delve into this topic in the focus group interview; thus, we do not have converging empirical evidence to advance discussions regarding ethics and fairness in teachers' LAL.

Conclusions

This study reports on language teachers' evaluation of an online assessment course regarding three components: course contents, activities, and impact on participants' professional development. Overall, thanks to course organization and delivery, the teachers found all contents relevant for their assessment practice—especially integrated assessment and assessing learners with SEN. Statistics, conversely, were not so relevant for these teachers. In terms of activities to foster LAL, teachers favored those that involved test analysis and design while evaluating the calculation of statistics as less useful for their teaching contexts. Finally, the impact on teachers' professional development was evident in two areas: Teachers reflected on prior assessment experiences and reported that they would like to change them; on the other hand, the teachers evaluated their LAL, stating that improvements will lead to better practices to serve their students' language learning needs.

References

- Arias, C. I., Maturana, L. M., & Restrepo, M. I. (2012). Evaluación de los aprendizajes en lenguas extranjeras: hacia prácticas justas y democráticas. *Lenguaje*, 40(1), 99–126. <https://doi.org/10.25100/lenguaje.v40i1.4945>
- Baker, B. (2021, July 1–3). *Unpacking the concept of language assessment literacy for all key stakeholders* [Conference session]. New Directions in English Language Assessment, Bogotá, Colombia (Online).
- Baker, B. A., & Riches, C. (2018). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557–581. <https://doi.org/10.1177/0265532217716732>
- Bøhn, H., & Tsagari, D. (2021). Teacher educators' conceptions of language assessment literacy in Norway. *Journal of Language Teaching and Research*, 12(2), 222–233. <https://doi.org/10.17507/jltr.1202.02>
- Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, E. Grove, K. Hill,

- N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 126–136). Cambridge University Press.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Csépes, I. (2021). The evolving concept of (language) assessment literacy: Implications for teacher education. *Central European Journal of Educational Research*, 3(1), 120–130. <https://doi.org/10.37441/CEJER/2021/3/1/9360>
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- Fernandes, M., & Borges-Almeida, V. (2021). *Letramento em avaliação na formação inicial de professores de línguas*. Editora CRV. <https://doi.org/10.24824/978652511072.1>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Fulcher, G. (2020). Operationalizing language assessment literacy. In D. Tsagari (Ed.), *Language assessment literacy: From theory to practice* (pp. 8–28). Cambridge Scholars.
- Gan, L., & Lam, R. (2022). A review on language assessment literacy: Trends, foci and contributions. *Language Assessment Quarterly*, 19(5), 503–525. <https://doi.org/10.1080/15434303.2022.2128802>
- Giraldo, F. (2021). Language assessment literacy and teachers' professional development: A review of the literature. *Profile: Issues in Teachers' Professional Development*, 23(2), 265–279. <https://doi.org/10.15446/profile.v23n2.90533>
- Giraldo, F., & Murcia Quintero, D. (2019). Language assessment literacy and the professional development of pre-service language teachers. *Colombian Applied Linguistics Journal*, 21(2), 243–259. <https://doi.org/10.14483/22487085.14514>
- Giraldo, F., & Yan, X. (2023). Planning an online assessment course for English language teachers in Latin America. *Profile: Issues in Teachers' Professional Development*, 25(2), 147–164. <https://doi.org/10.15446/profile.v25n2.104703>
- González, E. F. (2021). The impact of assessment training on EFL writing classroom assessment: Voices of Mexican university teachers. *Profile: Issues in Teachers' Professional Development*, 23(1), 107–124. <https://doi.org/10.15446/profile.v23n1.85019>
- Inbar-Lourie, O. (2013). Guest Editorial to the special issue on language assessment literacy. *Language Testing*, 30(3), 301–307. <https://doi.org/10.1177/0265532213480126>
- Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, I. G. Or, & S. May (Eds.), *Language testing and assessment: Encyclopedia of language and education* (3rd ed., pp. 257–270). Springer. https://doi.org/10.1007/978-3-319-02261-1_19
- Ivankova, N. V., & Greer, J. L. (2015). Mixed methods research and analysis. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource* (2nd ed., pp. 63–81). Bloomsbury Publishing.
- Janssen, G. (2022). Local placement test retrofit and building language assessment literacy with teacher stakeholders: A case study from Colombia. *Language Testing*, 39(3), 377–400. <https://doi.org/10.1177/02655322221076153>
- Koh, K., Burke, L., Luke, A., Gong, W., & Tan, C. (2018). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, 22(3), 264–288. <https://doi.org/10.1177/1362168816684366>
- Kremmel, B., Eberharter, K., Holzknicht, F., & Konrad, E. (2018). Fostering language assessment literacy through teacher involvement in high-stakes test development. In D. Xerri & P. Vella Briffa (Eds.), *Teacher involvement in high-stakes language testing* (pp. 173–194). Springer. https://doi.org/10.1007/978-3-319-77177-9_10
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17(1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>
- Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers.

- Language Assessment Quarterly*, 17(2), 168–182. <https://doi.org/10.1080/15434303.2019.1692347>
- López Mendoza, A. A., & Bernal Arandia, R. (2009). Language testing in Colombia: A call for more teacher education and teacher training in language assessment. *Profile: Issues in Teachers' Professional Development*, 11(2), 55–70.
- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Blackwell.
- Montee, M., Bach, A., Donovan, A., & Thompson, L. (2013). LCTL teachers' assessment knowledge and practices: An exploratory study. *Journal of the National Council of Less Commonly Taught Languages*, 13, 1–31.
- Papadima-Sophocleous, S. (2022). Teacher education for language assessment and testing: Postgraduate program evaluation from its students' perspective. *Colombian Applied Linguistics Journal*, 24(1), 67–88. <https://doi.org/10.14483/22487085.17373>
- Pastore, S. (2023). Teacher assessment literacy: a systematic review. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1217167>
- 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>
- Quevedo-Camargo, G., & Tonelli, J. R. A. (2021). Contribuições de um curso EAD para o letramento em avaliação de professores de línguas adicionais para crianças. *Revista do GEL*, 18(1), 230–248. <https://doi.org/10.21165/gel.v18i1.3094>
- Restrepo Bolivar, E. M. (2020). Monitoring preservice teachers' language assessment literacy development through journal writing. *Malaysian Journal of ELT Research*, 17(1), 38–52.
- Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). SAGE Publications.
- 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>
- Shohamy, E. (2017). Critical language testing. In E. Shohamy, I. Or, & S. May (Eds), *Language testing and assessment: Encyclopedia of language and education* (3rd ed., pp. 441–454) Springer. https://doi.org/10.1007/978-3-319-02261-1_26
- Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 2–14. <https://doi.org/10.1186/s40468-019-0077-8>
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412. <https://doi.org/10.1177/0265532213480338>
- Villa-Larenas, S., Quevedo, G., & González, F. (2021, July 1–3). *Survey on language assessment literacy in Latin America: Preliminary results* [Conference session]. New Directions in English Language Assessment, Bogotá, Colombia (Online).
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402. <https://doi.org/10.1080/15434303.2014.960046>
- Walters, F. S. (2010). Cultivating assessment literacy: Standards evaluation through language-test specification reverse engineering. *Language Assessment Quarterly*, 7(4), 317–342. <https://doi.org/10.1080/15434303.2010.516042>
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Yan, X. (2021, June 14–17). *Towards a conceptual model of LAL development across stakeholder groups and contexts* [Conference session]. Language Testing Research Colloquium (Online).
- Yan, X., & Fan, J. (2021). “Am I qualified to be a language tester?”: Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing*, 38(2), 219–246. <https://doi.org/10.1177/0265532220929924>
- Yan, X., Zhang, C., & Fan, J. J. (2018). “Assessment knowledge is important, but ...”: How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158–168. <https://doi.org/10.1016/j.system.2018.03.003>

About the Authors

Frank Giraldo is a professor at the Foreign Languages Department of Universidad de Caldas (Colombia). He holds an MA in English Didactics from this University and an MA in TESOL from the University of Illinois at Urbana-Champaign (USA). His research interest is the interface between language assessment literacy and teachers' professional development.

Xun Yan is an associate professor of Linguistics, Second Language Acquisition, Teacher Education, and Educational Psychology at the University of Illinois at Urbana-Champaign (USA). He holds a PhD in Second Language Studies from Purdue University (USA). His research interests include speaking and writing assessment, psycholinguistic approaches to language testing, and language assessment literacy.

Appendix: Theme Coding for Focus Group Interview and Answers from Open Questionnaire Items

Categories for Theme 1: Perceptions of contents

1. Appropriate organization and sequencing
2. Variety in preferred topics
3. Integrated skills assessment and assessing learners with SEN as the most useful contents
4. Statistics as least useful content

Categories for Theme 2: Perceptions of materials

1. Fostering learning about language assessment
2. Appropriateness of handbook: design, sequencing, and content itself

Categories for Theme 3: Perceptions of activities

1. Workshops promote interaction, which leads to cooperative learning
2. Positive impact of interacting with others
3. Calculating stats as the least useful activity

Categories for Theme 4: Perceptions of how the course impacted teachers' professional development

1. Overall positive impact on an LAL profile
2. Reflection on the improvement of assessment practices
3. Relationship between learning about assessment and being better for students

Categories for Theme 5: Course suggestions

1. More test analysis
2. One platform for course delivery