

Planning an Online Assessment Course for English Language Teachers in Latin America

La planeación de un curso de evaluación en línea para docentes de inglés de Latinoamérica

Frank Giraldo

Universidad de Caldas, Manizales, Colombia

Xun Yan


University of Illinois at Urbana-Champaign, USA


In this article, we report the results of a study through which we collected English language teachers' needs and wants to design an online language assessment course. Through a mixed-methods approach, we asked 20 teachers from four Latin American countries what they wanted to learn in the course. The teachers wanted a course in which they could address the challenges they faced in assessment; discuss and develop new ways to assess; and learn about authentic, valid, and ethical assessment. Therefore, the findings suggest that the teachers wanted a course that mixed theory, practice, and principles of assessment. Additionally, the course should address emerging topics in English language assessment, namely bilingual assessment and the assessment of learners with special educational needs.

Keywords: assessment literacy, language assessment literacy, language testing, teacher professional development

En este artículo reportamos los hallazgos de una investigación mixta acerca de los deseos y necesidades en evaluación de un grupo de docentes de inglés con el objetivo de diseñar un curso en línea de evaluación de lenguas extranjeras. Tras ser entrevistados, veinte docentes de inglés de cuatro países latinoamericanos informaron que deseaban un curso que les permitiera discutir los retos de la evaluación; estudiar maneras novedosas de evaluar, y aprender sobre la evaluación auténtica, válida y ética. Es decir, un curso que incorpore la teoría, la práctica y los principios de evaluación y que además explore temas emergentes en la evaluación del inglés como lengua extranjera: la evaluación bilingüe y la evaluación de estudiantes con necesidades educativas especiales.

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

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

The official name for this project was *Fostering Language Teachers' Assessment Literacy Through an Online Program: The Latin American Case*. This project was funded by Fulbright Colombia, through a scholarship called "Visiting Colombian researcher."

How to cite this article (APA, 7th ed.): 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>

This article was received on September 8, 2022 and accepted on March 14, 2023.

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

After Davies' (2008) seminal paper, which brought impetus for discussion on language assessment literacy (LAL), there has been a major focus of scholarly work on teachers' LAL, particularly regarding their training (or lack thereof), practices, and beliefs (Crusan et al., 2016; Fulcher, 2012; Sultana, 2019; Vogt & Tsagari, 2014; Yan & Fan, 2021). Researching these areas in teachers' LAL is indeed essential because a clear picture of assessment is a necessary condition for devising appropriate instructional initiatives for teacher education (Baker & Riches, 2017; Fulcher, 2012; Giraldo & Murcia, 2018; Malone, 2017; Scarino, 2013; Yan & Fan, 2021; Yan et al., 2018).

Researchers have also attempted to describe language teachers' LAL profiles, particularly the necessary knowledge and skills required for professional assessment (Kremmel & Harding, 2019; Stabler-Havener, 2018; Taylor, 2013). Additionally, scholars have indicated that drawing the LAL profiles for several stakeholders—including language teachers—is challenging, given the contextual nature of language assessment and pedagogy (Scarino, 2013; Yan, 2021). However, there is abundant research evidence to conclude that language teachers should participate in LAL initiatives (e.g., courses), which may lead to professionalization in language assessment. Such a professional profile might cultivate an assessment culture in which teachers critically consider assessment and its implications for their institutions (Boyd & Donnaruma, 2018; Inbar-Lourie, 2008, 2017; Vogt & Tsagari, 2014).

To propose pedagogically sound initiatives for teachers' LAL, feedback from these stakeholders is a central consideration (Bøhn & Tsagari, 2021). Thus, our purpose in this paper is to present and discuss the findings we gathered from a research study in which we collected data from English language teachers to design an online language assessment course.

Literature Review

Teachers' Language Assessment Literacy

Language testing scholars have discussed LAL within three major components: knowledge, skills, and principles (Davies, 2008; Fulcher, 2012). Knowledge refers to concepts and theories around assessment; skills help stakeholders to design, develop, or evaluate language tests; and principles include considerations such as ethics and fairness in assessment. LAL profiles naturally differ when various stakeholders are involved (e.g., researchers, language testers, school administrators), encouraging scholars to research particular LAL profiles for different people (Kremmel & Harding, 2019; Inbar-Lourie, 2013a; Taylor, 2013).

As mentioned above, teachers' LAL has been a central focus in LAL literature. Descriptions of competency in assessment have been proposed for teachers, from the principles for assessment in general education (American Federation of Teachers et al., 1990) to Fulcher's (2012) LAL model for language teachers and the International Language Testing Association's (2007) guidelines for practice. The existing LAL models and descriptions have further suggested that teachers' LAL profile amalgamates theoretical, technical, and critical considerations (Fulcher, 2012; Giraldo, 2018; Inbar-Lourie, 2013b; Stabler-Havener, 2018). In sum, teachers with solid training in language assessment are better positioned to develop appropriate testing systems for their schools and students and to evaluate existing systems and their impact (Boyd & Donnarumma, 2018; Kremmel et al., 2018).

Although teachers' LAL is (and probably will be) a matter of discussion, some trends regarding the knowledge, skills, and principles they should have for language assessment are clear. In Table 1, we synthesize significant aspects of each LAL component.

Table 1. Major Components in Teachers' Language Assessment Literacy

| Knowledge | Skills | Principles |
|--|---|--|
| <ul style="list-style-type: none"> • Models of language ability • Bilingualism and multilingualism • Theories and frameworks for assessment • Measurement concepts • Current language teaching pedagogies • Personal beliefs and practices • Local policies | <ul style="list-style-type: none"> • Test analysis and critique • Design of test items and tasks for different language skills • Calculation or interpretation of statistics • Integration among assessment, teaching, and learning | <ul style="list-style-type: none"> • Ethics • Fairness |

Pedagogical Initiatives for LAL

Researchers in language testing have reported on courses for teachers' LAL. These initiatives have shown that teachers have become aware of assessment and its positive impact on their assessment practices, teaching, and students' learning. Notably, teachers in these initiatives have been involved in assessment development tasks which have cultivated their professional development in this area. Below, we describe the major components of these LAL courses regarding contents and pedagogical approaches.

Contents in LAL Courses

LAL courses are primarily based on knowledge and skills, mainly including major measurement concepts, purposes and types of assessment, and theoretical considerations for assessing language skills (Giraldo & Murcia, 2019; Kleinsasser, 2005; Kremmel et al., 2018; Montee et al., 2013; Nier et al., 2009; O'Loughlin, 2006). However, other LAL courses for teachers have used a more specific approach to content selection and use. For example, in Koh et al. (2018), the researchers taught teachers about the principles of task development within Authentic Intellectual Quality; in Walters (2010), teachers in New York learned about test specifications to conduct reversed-engineer specifications to evaluate standards for language learning.

Methodological Approaches in LAL Courses

Regarding how teachers learn about language assessment, there is a clear tendency towards a practical focus, with hands-on workshops being prominent in these courses. In the courses we reviewed, generally, teachers are engaged in test critique, test and task development (including peer feedback), and test and task evaluation (Arias et al., 2012; Giraldo & Murcia, 2019; Koh et al., 2018; Kremmel et al., 2018; Montee et al., 2013).

In our review, we found limited attention to the *principles* side of LAL. Courses addressing concerns for ethics and fairness included discussions of ethics conceptualized as transparency in assessment, that is, informing students of the what and how of assessment (Arias et al., 2012; Levi & Inbar-Lourie, 2019; Restrepo-Bolívar, 2020). Kleinsasser's (2005) study mentions Shohamy's (2001) *The Power of Tests*, but the author does not discuss how principles for assessment were addressed among the participating student teachers.

Thus, LAL initiatives for teachers have mainly focused on the theoretical and practical aspects of testing, but not so much on the critical side of this field, despite a consensus on the need for ethics and fairness in LAL (Fulcher, 2012; Inbar-Lourie, 2008, 2013a).

Research Context

The findings we report in this work come from the diagnostic stage in a project seeking to create an online assessment course for 50 Latin American English language teachers. In the first study of Latin America, Villa-Larenas et al. (2021) surveyed stakeholders' perceptions of and needs in their LAL. The researchers concluded that LAL initiatives should be implemented across this region. Thus, and as Fulcher (2020) encourages, the field of language testing should move from description to action so that we can cater to teachers' needs in language assessment.

Responding to these two calls, we shared information about a free online language assessment course for 50 teachers in five Latin American countries: Brazil, Colombia, Ecuador, Peru, and Venezuela. The initiative was designed as a 10-week, 40-hour course with two weekly workshops of two hours each. To invite participants, we shared this initiative via L-Test (an email list-serve for language testing and assessment), personal contacts and social media, the Latin American Association of Language Testing and Assessment, and the International Research Foundation for English Language Education. We also contacted fellow teacher educators in language teaching programs, expecting they could share the information with English language teachers.

To join the course, the participants had to meet three criteria: (a) be English language teachers at the high school level, (b) work for public/state schools in any of the five countries, and (c) agree to provide information about their assessment practices and learning needs, before course implementation. After two months, 20 teachers' applications were successful and made them course participants (more information in the Results section). In this paper, we report the diagnostic stage based on the feedback from these 20 teachers. This stage was guided by this research question: What do a group of English language teachers in Latin America need and want to learn about language assessment?

Responses to this question would then lead us to interpret how these stakeholders' needs and wants could inform the design and implementation of an appropriate language assessment course to be delivered online.

Method

To draw participants' LAL profiles and needs, we resorted to a mixed-methods approach, which allowed us to have complementary data to plan a context-sensitive course. A concurrent mixed-methods research design (Ivankova & Greer, 2015) provides a complete picture of the research phenomenon from different perspectives (Creswell, 2014; Dörnyei, 2007).

We first interviewed each teacher individually, asking them about their language assessment practices, challenges, and learning needs. For this, we used a semi-structured interview protocol (see Appendix A), which we administered before the teachers took the online questionnaire so that the interview data would emerge without influence from the language in the questionnaire items (Fulcher, 2012). Before the interviews, each teacher received an email telling them that the interviews were about their practices and challenges in language assessment and aspects they would like to learn about in this area. Each interview took place and was recorded online via Google Meet.

After the interviews, the teachers completed a questionnaire through Google Forms (see Appendix B). The questionnaire, with 52 items, was divided into four sections: background information (six items, including country, age, and studies); prior training in language assessment (six items); assessment activities in which the teachers have been engaged (eight items); and LAL topics for the course (31 items).

In Section 4 of the questionnaire, the teachers were asked to determine to what extent they found various language assessment topics necessary for the course, rating them from 1 (*not important at all*) to 5 (*extremely important*). For this instrument, we decided not to include the technical, generic names of concepts

in language assessment but rather what we considered a teacher-friendly description. For example, the item for authenticity (Item 10) was “Creating test items and tasks that resemble real-life language use.”

To design a questionnaire that could be fit for purpose, we examined and ensured its validity in three ways. First, we wrote items reflecting the central construct under investigation: LAL. To do so, we based the items on the three major components of LAL, as suggested in the literature—knowledge, skills, and principles. For instance, Item 10 was meant to tap into the *knowledge* of authenticity and the *skill* of designing authentic test tasks; Item 29 (Discussing ethics and fairness in classroom language assessment) was meant to collect information on these *principles* for assessment. Secondly, we asked four content experts engaged in language testing and assessment and language teacher education to evaluate the questionnaire. The feedback form had this guiding prompt: “Evaluate whether the instruments are fit for purpose, i.e., that they have the potential to give us information about teacher needs to plan and design a LAL program for English language teachers.” All experts agreed that the questionnaire was appropriate for these purposes. Two experts recommended adding Item 21 (Assessing different age groups: young learners, teenagers) and Item 30 (Conducting fair, ethical, and transparent assessment practices). These additions helped with the method’s construct validity. Finally, we calculated Cronbach’s Alpha to check the internal consistency of the items in Section 4 (LAL contents) of the questionnaire. The value for Cronbach’s Alpha was $\alpha = .87$, suggesting satisfactory internal consistency (George & Mallery, 2003) in measuring what we defined as LAL.

Overall, the content experts suggested that the interview and the questionnaire tap into teachers’ prior training and experiences in assessment. Thus, in the interview, we included one question about experiences and one about assessment challenges or difficulties. As for the questionnaire, we had closed-ended items on prior assessment activities.

Data Analysis

After transcribing the interview answers for each teacher, we used theme coding (Saldaña, 2016) as an iterative process of reading and analysis, allowing us to identify trends in teachers’ practices, challenges, and LAL learning needs across all transcriptions. From this first read-through, we made a list of initial themes; we then used NVivo (QSR International, n.d.) to code the data to refine, discard, or validate our initial themes, especially those amenable to use for planning the online course. The final themes from this data set were *institutional challenges influencing assessment*, *learning about innovative ways to assess professionally*, and *spaces for discussions about assessment*. Table 2 has the complete matrix derived from our theme analysis.

For the questionnaire data, we calculated descriptive statistics for the items in the last section: the topics for the LAL course. The descriptive statistics were mean, median, and range. We used these three measures to understand teachers’ choices clearly. With the results from the questionnaire—specifically, means and medians—we ranked the topics the teachers considered most important for this LAL course. Thus, Table 3 presents the data referring to teachers’ prior language assessment training, while Table 4 describes teachers’ involvement in assessment activities.

Results

This section first characterizes teachers’ challenges and learning needs in language assessment. For this, we include data that illustrate trends across the interviews. Then, we report the results from the questionnaire items asking teachers to rate the importance of LAL topics for the course.

Institutional Challenges Influencing Assessment

When asked about challenges or difficulties that the teachers faced when it came to assessment, most of them indicated, first and foremost, that the number

Table 2. Significant Themes and Codes From Interview Data

| Question focus | Themes in data | Data codes |
|--|---|---|
| Practices in language assessment | A. Purpose: Progress B. Purpose: Administrative C. Methods: Traditional D. Methods: Performance E. Methods: Other F. Skills: Reading/Writing G. Skills: Listening/Speaking H. Skills: All four | A1. Checks on learning/whether students are learning B1. Provides or uploads grades on the platform C1. Administers quizzes/tests D1. Administers interviews/oral presentations E1. Uses exit slips/teaching logs F1.1. Assesses reading F1.2. Assesses writing G1.1. Assesses listening G1.2. Assesses speaking H1. Assesses all skills |
| Challenges/Difficulties in language assessment | I. Institutional J. Involving people | I1.1. Mentions lack of resources I1.2. Mentions lack of time J1.1. Mentions students' attitudes J1.2. Mentions parents' attitudes J1.3. Mentions colleagues' attitudes |
| Language assessment literacy learning needs | K. Conduct better assessment L. Learn other assessment methods N. Fairness in assessment N. Learn with/from others O. Spaces for discussion | K1.1. Wants to do better assessments K1.2. Wants to become a better assessor K1.3. Wants to improve their assessment practices L1.1. Wants to learn about better/other assessment methods L1.2. Wants to learn about new techniques M1.1. Mentions fairness only M1.2. Wants to be fair N1.1. Wants to learn about other teachers' assessment approach N1.2. Wants to learn about other teachers' contexts O1.1. Expects to have discussion spaces O1.2. Recommends spaces for discussion |

of students impedes efficient assessment practices. As the teachers commented, giving personalized feedback becomes challenging due to the high number of students per classroom and grade. In a related manner, the teachers stated that they had limited time to assess their learners, which they see as challenging to conduct appropriate assessments (e.g., administration and checking). As T15 comments:

When I use rubric[s] because, for each student, another difficulty is the time because I have only three times per

week, three pedagogical hours per week. No more. And I have uh, 35 or 38 students, so it's so difficult to do one rubric for each student.

Learning About Innovative Ways to Assess Professionally

One of the main questions in the interview asked teachers what they would like to learn about assessment. A related question asked teachers about expectations they had towards the online course. In their responses to

Table 3. Teachers' Prior Training in Language Assessment

| Items | Yes | No | N/A |
|---|-----|----|-----|
| A complete course in your undergraduate studies | 5 | 15 | |
| A complete course in your graduate studies (MA or PhD) | 3 | 14 | 3 |
| A module in a course in your undergraduate studies | 13 | 7 | |
| A module in a course in your graduate studies (MA or PhD) | 7 | 10 | 3 |
| Attended workshops (about language assessment) after you started to work as an English language teacher | 13 | 7 | |
| Learning about language assessment independently, e.g., reading articles, watching videos | 17 | 3 | |

Table 4. Participation in Assessment-Related Activities

| Items | Yes | No |
|---|-----|----|
| Evaluating tests found in textbooks to be used in your school | 17 | 3 |
| Evaluating the results of regional or national standardized tests in your country | 14 | 6 |
| Participating in regional or national test development projects | 8 | 12 |
| Designing your tests to use in your school | 20 | 0 |
| Administering tests designed by coordinators in your school | 7 | 13 |
| Administering tests designed by organizations external to your school | 12 | 8 |
| Participating in assessment committees at your school | 12 | 8 |
| Participating in assessment committees in your city, region, or country | 5 | 15 |

these two questions, the teachers highlighted an overall need to become better assessors by resorting to new methods they can use with their language learners and help them improve their language ability. The sample below, from T10, shows the need for a more professional approach to language assessment.

I want to innovate for that they can be a good, [*sic*] or a better feedback [*sic*] between them and me, and we will work better. I really want this kind of tool or innovation tools. Not always the same. I want to change.

Spaces for Discussions About Assessment

The interviews showed that teachers found being in a course with English language teachers from other Latin American countries helpful. Thus, the answers showed that they expected to have spaces to share ideas about assessment and, correspondingly, learn from each other's experiences and contexts. T8 comments on the expectation to have these discussion scenarios and the positive consequences they could have: "I don't

know. Probably have more spaces. To talk, to share our experiences, our ideas. . . . And these spaces also help us to improve our English practice, and our English.”

Importance of LAL Topics for the Course

We now turn to the quantitative data drawn from the questionnaire. In Table 5, we present the descriptive statistics about the items in the last part of the questionnaire, with mean and median values ranked from highest to lowest.

The results show that the teachers ($N = 20$) found 23 topics, out of 31, extremely important for the course. *Creating test items and tasks that resemble real-life language use* and *Designing test items that have the potential to collect precise information about your students' language skills* were the items with the highest means ($M = 4.7$, $Mdn = 5$) and a narrow range of 4–5. These items are followed closely by two items with the same mean and median but a slightly wider range of 3–5: *Designing test tasks that have the potential to collect precise information about your students' language skills* and *Establishing a clear purpose for assessing your students' English*. Item 22, *Assessing learners with special educational needs*, has a mean of 4.5 and a median of 5 but a relatively wide range: 2–5. A slightly similar picture is presented by Items 29 and 30: *Discussing ethics and fairness in classroom language assessment* ($M = 4.3$, $Mdn = 4$) and *Conducting fair, ethical, and transparent assessment practices* ($M = 4.4$, $Mdn = 4$); these two items have a wide range of 1–4.

The teachers found seven topics very important for the online assessment course, though their responses varied widely; for instance, Item 19, *Assessing pronunciation* ($M = 3.9$, $Mdn = 4$), has a range of 1–4. Additionally, two items have a mean and median of 4 and a high range of 1–4: Item 24, *Designing alternative assessments (e.g., self- and peer assessment, portfolios)*; and Item 17, *Assessing grammar*. Item 23 (*Designing traditional tests with true-false and multiple-choice*

questions) has lower values ($M = 3.4$, $Mdn = 3$) and a wide range of 1–4. Finally, teachers' responses varied widely for Item 1 (*Learning about the history of language testing*) and Item 14 (*Evaluating large-scale or standardized language tests*), with a wide range of 1–5. Specifically, teachers found Item 1 moderately important ($M = 2.8$, $Mdn = 3$) and Item 14 very important ($M = 3.9$, $Mdn = 4$).

Discussion

Even though the challenges the teachers discussed in the interview are not assessment-related per se, they impact their assessment practices. The difficulties related to limited time and many students have been documented elsewhere in Latin America (Díaz-Larenas et al., 2012; Frodden et al., 2004). Scarino (2013) argues that teachers' lifeworlds need to be considered when discussing teachers' LAL, and this may include problems teachers face in assessment. In the case of this finding, our challenge lay in addressing the high number of students and time limitations as topics in the online assessment course (see Appendix C for how we decided to address these challenges).

As for the learning needs the teachers expressed in the interview, two aspects merit discussion. On the one hand, the teachers expressed their desire to assess well and learn new assessment methods. The teachers did not provide more specific answers for this question, and we believe this is expected because they may lack knowledge of assessment terminology. Two teachers emphatically commented that they wanted to learn how to be fair in assessment, but this trend was not frequent enough in the interview data to discuss it here. We also acknowledge that the lack of granularity in teachers' answers may be an issue in our research: Asking teachers what they want to learn about assessment may not be fruitful. However, we felt compelled to ask the question: “What would you like to learn about assessment?” because this information was included in the email teachers received about the interview (see the

Table 5. Descriptive Statistics for Importance of Language Assessment Literacy Topics in the Course

| | Items | M | Median | Range |
|---------|---|----------|---------------|--------------|
| Item 10 | Creating test items and tasks that resemble real-life language use | 4.7 | 5 | 1 |
| Item 7 | Designing test items that have the potential to collect precise information about your students' language skills | 4.7 | 5 | 1 |
| Item 8 | Designing test tasks that have the potential to collect precise information about your students' language skills | 4.7 | 5 | 2 |
| Item 2 | Establishing a clear purpose for assessing your students' English | 4.7 | 5 | 2 |
| Item 5 | Planning the design of assessment instruments | 4.6 | 5 | 2 |
| Item 9 | Creating test items and tasks that can provide consistent and reliable information about your students' language skills | 4.6 | 5 | 2 |
| Item 16 | Assessing productive skills: speaking and writing | 4.6 | 5 | 2 |
| Item 3 | Evaluating whether an assessment instrument is meeting, or not, its purpose | 4.5 | 5 | 2 |
| Item 4 | Evaluating whether a test has the potential to collect information about a determined set of language skills | 4.5 | 5 | 2 |
| Item 22 | Assessing learners with special educational needs | 4.5 | 5 | 3 |
| Item 20 | Assessing integrated skills | 4.5 | 5 | 1 |
| Item 15 | Assessing receptive skills: listening and reading | 4.4 | 5 | 1 |
| Item 30 | Conducting fair, ethical, and transparent assessment practices | 4.4 | 4 | 3 |
| Item 6 | Describing clearly the particular skills you want to assess | 4.4 | 4 | 2 |
| Item 13 | Evaluating the positive or negative influence that assessment can have on teaching and learning | 4.4 | 4 | 2 |
| Item 26 | Relating language assessment to language teaching and learning | 4.4 | 4 | 2 |
| Item 31 | Analyzing misuses of language assessment | 4.3 | 4 | 2 |
| Item 18 | Assessing vocabulary | 4.3 | 4 | 2 |
| Item 11 | Evaluating available resources (e.g., technology) for test development | 4.3 | 4 | 1 |
| Item 12 | Using available resources efficiently during test development | 4.3 | 4 | 2 |
| Item 29 | Discussing ethics and fairness in classroom language assessment | 4.3 | 4 | 3 |
| Item 21 | Assessing different age groups (young learners, teenagers) | 4.2 | 4 | 2 |
| Item 25 | Administering assessments successfully | 4.2 | 4 | 2 |
| Item 24 | Designing alternative assessment methods (e.g., self- and peer assessment, portfolios) | 4 | 4 | 3 |
| Item 28 | Assessing students in a bilingual mode | 4 | 4 | 3 |
| Item 17 | Assessing grammar | 4 | 4 | 3 |

| | Items | <i>M</i> | <i>Median</i> | <i>Range</i> |
|---------|---|-----------------|----------------------|---------------------|
| Item 27 | Interpreting statistical information and scores from language assessment | 4 | 4 | 2 |
| Item 19 | Assessing pronunciation | 3.9 | 4 | 3 |
| Item 14 | Evaluating large-scale or standardized language tests | 3.9 | 4 | 4 |
| Item 23 | Designing traditional tests with true-false and multiple-choice questions | 3.4 | 3 | 3 |
| Item 1 | Learning about the history of language testing | 2.8 | 3 | 4 |

discussion for questionnaire results following and the Limitations section for more on this matter).

On the other hand, in the interviews, the teachers clearly stated that they wanted to learn about other teachers' perspectives of assessment through interactive spaces in the online course. LAL scholars have suggested that LAL involves social learning, which can happen through learning communities of practice to help teachers move forward in LAL (Baker, 2021; Scarino, 2013; Yan, 2021). In the case of this group of teachers, learning from others may be helpful as they all have been involved in language assessment activities of some kind (see Table 4). Thus, this finding should have a clear implication for planning and teaching the online assessment course—including interactive and collaborative activities. Therefore, asking teachers about their course expectations, as in the present study, yielded valuable data for course planning.

Regarding the questionnaire results, these teachers found 23 topics out of 31 to be extremely important for learning in the LAL course, further corroborating their need for LAL: knowledge, skills, and principles. That teachers need rounded LAL has been found in other studies (Fulcher, 2012; Giraldo & Murcia, 2018; Vogt & Tzagari, 2014). Remarkably, the questionnaire results provide feedback to plan a generic assessment course: One that includes test design as a significant task (see the six top items in Table 5); general assessment concepts and principles, such as authenticity and validity (Items 7, 8, and 10); and attention to the assessment of language skills (Items 15, 16, and 20). As we show in our literature

review, test design, and core concepts are prominent in language testing courses for teachers.

Notwithstanding the generic course nature we outline above, this group of teachers signaled areas in language assessment that have not been documented prominently in other studies regarding English teachers' LAL needs: assessing learners with special educational needs (SEN); ethics and fairness; and bilingual assessment, with somewhat lesser importance. In other studies, principles are ranked low by teachers (Fulcher, 2012; Vogt & Tzagari, 2014), whereas the assessment of SEN learners and bilingual assessment have not appeared as topics for English teachers' LAL development.

Limitations

The data for this study was used to plan a language assessment course for 20 English language teachers in four Latin American countries. The data cannot represent nor be considered trends in Latin America or any given country in this region. The data, however, were beneficial to plan a course that could cater to the LAL needs and expectations of the teachers involved.

As we stated earlier, asking teachers directly what they want to learn about assessment may not be fully informative, as they provide general answers or state they want to learn everything the course offers. Aware of this issue, we also used the online questionnaire, which confirmed the need for a balanced—although contextual—LAL profile. Thus, we believe more research should be done regarding the use of interviews to collect data for planning a LAL course for English

language teachers. In our case, asking teachers about their expectations—rather than what they want to learn—helped collect their needs, but this technique may require further scrutiny.

Finally, as we commented in our Method section, we phrased the items in the last part of the questionnaire to avoid using technical names such as validity and reliability. Therefore, we must acknowledge that these descriptions simplify heated debates in language testing and educational assessment in general. However, we needed to design items that would be useful to collect teachers' LAL learning needs; this is something that, we argue, the questionnaire did reasonably well.

Conclusions and Recommendations

The purpose of this research report was to characterize the learning needs of a group of English language teachers in Latin America in the area of LAL. The data we collected through the interview and questionnaire led us to suggest that this group of stakeholders needed a course that balanced the generic of LAL and the specific, that is, challenges at their institutions. The answers taught us that the course should be driven by practice, including interactive discussions, address traditional topics such as validity and authenticity, critical issues such as ethics and fairness, and address emerging issues in English language testing (e.g., SEN learners and bilingual assessment). Appendix C has the LAL topics we decided to include in the course, with some commentary reflecting what we found.

Based on our LAL needs analysis exercise, we recommend that researchers, especially those engaged in teaching teachers about language assessment, utilize a mixed-methods approach to data collection. This methodology may provide information to substantiate an LAL course and unveil the stakeholders' particularities. Additionally, we recommend that when and if appropriate, the items to survey teachers' LAL should be phrased as practices rather than merely topics, from *test specifications* to *planning assessments carefully*. This change in

LAL language may be more friendly to teachers who may not have the terminology but have LAL learning needs.

References

- American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students*. <https://buros.org/standards-teacher-competence-educational-assessment-students>
- 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>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Baker, B. (2021, July 2). *Unpacking the concept of language assessment literacy for all key stakeholders* [Conference session]. New Directions in English Language Assessment Conference, Bogotá, Colombia. <https://bit.ly/3F9dOTT>
- Baker, B. A., & Riches, C. (2017). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557–581. <https://doi.org/10.1177/0265532217716732>
- 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>
- Boyd, E., & Donnarumma, D. (2018). Assessment literacy for teachers: A pilot study investigating the challenges, benefits and impact of assessment literacy training. In D. Xerri & P. Vella Briffa (Eds.), *Teacher involvement in high-stakes language testing* (pp. 105–126). Springer. https://doi.org/10.1007/978-3-319-77177-9_7
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE.
- Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing*, 28, 43–56. <https://doi.org/10.1016/j.asw.2016.03.001>

- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- Díaz-Larenas, C., Alarcón-Hernández, P., & Ortiz-Navarrete, M. (2012). El profesor de inglés: sus creencias sobre la evaluación de la lengua inglesa en los niveles primario, secundario y terciario. *Íkala, Revista de Lenguaje y Cultura*, 17(1), 15–26. <https://doi.org/10.17533/udea.ikala.10039>
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Frodden, M., Restrepo, M., & Maturana, L. (2004). Analysis of assessment instruments used in foreign language teaching. *Íkala, Revista de Lenguaje y Cultura*, 9(1), 171–201. <https://doi.org/10.17533/udea.ikala.3146>
- 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 assessment literacy. In D. Tsagari (Ed.), *Language assessment literacy: From theory to practice* (pp. 8–28). Cambridge Scholars.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Allyn & Bacon.
- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179–195. <https://doi.org/10.15446/profile.v20n1.62089>
- Giraldo, F., & Murcia, D. (2018). Language assessment literacy for pre-service teachers: Course expectations from different stakeholders. *GIST – Education and Learning Research Journal*, (16), 56–77. <https://doi.org/10.26817/16925777.425>
- Giraldo, F., & Murcia, 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>
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402. <https://doi.org/10.1177/0265532208090158>
- Inbar-Lourie, O. (2013a). 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. (2013b, November). *Language assessment literacy: What are the ingredients?* [Paper presentation]. 4th CBLA SIG Symposium, University of Cyprus, Cyprus.
- 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
- International Language Testing Association. (2007). *ILTA guidelines for practice in English*. <https://bit.ly/3F17GNp>
- Ivankova, N., & Greer, J. (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.
- Kleinsasser, R. C. (2005). Transforming a postgraduate level assessment course: A second language teacher educator's narrative. *Prospect*, 20(3), 77–102.
- Koh, K., Burke, L. E. C., 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. (2019). 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. (2019). 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>

- Malone, M. E. (2017). Training in language assessment. In E. Shohamy, I. G. Or, & S. May (Eds.), *Language testing and assessment: Encyclopedia of language and education* (3rd ed., pp. 225–239). Springer. https://doi.org/10.1007/978-3-319-02261-1_16
- 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.
- Nier, V. C., Donovan, A. E., & Malone, M. E. (2009). Increasing assessment literacy among LCTL instructors through blended learning. *Journal of the National Council of Less Commonly Taught Languages*, 7, 94–118.
- O'Loughlin, K. (2006). Learning about second language assessment: Insights from a postgraduate student on-line subject forum. *University of Sydney Papers in TESOL*, 1, 71–85.
- QSR International. (n.d.). Nvivo® (Version 10) [Computer software]. <https://bit.ly/3Xweiug>
- Restrepo-Bolívar, 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*. 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>
- Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Working Papers in Applied Linguistics & TESOL*, 18(1), 1–22.
- 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* [Paper presentation]. New Directions in English Language Assessment Conference, Bogotá, Colombia.
- 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>
- Yan, X. (2021, June 14–17). *Towards a conceptual model of LAL development across stakeholder groups and contexts* [Paper presentation]. 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 holds an MA in English Didactics from Universidad de Caldas (Colombia) and an MA in TESOL from the University of Illinois at Urbana-Champaign (USA). His research interests include the interface between language assessment literacy and teachers' professional development.

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

Appendix A: Interview Protocol

Procedures

- Start by greeting the teachers kindly and then reminding them of the purpose of this data collection. Finally, thank them for their interest in the course.
- Start with the ice-breaker, which need not be recorded.
- Start recording once you are ready to ask Question 2 (see Questions and Probes below).
- Ask probes as needed, especially when teachers seem to be stuck or needing help.
- Rephrase questions so that teachers can understand them better.
- After the interview, tell the teachers about the online questionnaire they will take.
- End the interview by thanking the teachers again and welcoming them to the course.

Questions and Probes

1. Ice-breaker: Tell me a little about your teaching context.
2. Tell me about how you assess your students.
Probes: What is your purpose? Why do you assess? What do you do with the information you collect? What skills do you assess?
3. In your context, what challenges (if any) do you face regarding language assessment?
Probe: What difficulties do you face? How do these challenges influence your assessment?
4. What would you like to learn about language assessment in the course?
Probes: What topics? What skills would you like to develop in this area? Anything else?
5. What expectations do you have about this course?

Appendix B: Questionnaire for the Diagnostic Stage

Section 1: Background Information

What is your full name?

Country where you work:

Colombia

Venezuela

Ecuador

Peru

Brazil

Name of the institution where you work:

What is your email account? Please type the one you use most frequently.

What is your level of education?

Bachelor of Arts (undergraduate studies)

Specialization (one year)

Master's degree

PhD

Choose your age range:

20–25 years old

26–30 years old

31–35 years old

36–40 years old

41–45 years old

46–50 years old

51–55 years old

56–60 years old

61 and older

Section 2: Prior Training in Language Assessment

Please, choose whether you have had any training in language assessment:

- A complete course in your undergraduate studies
Yes__ No__
- A complete course in your graduate studies (MA or PhD)
Yes__ No__ Do not have graduate studies__
- A module in a course in your undergraduate studies
- A module in a course in your graduate studies (MA or PhD)
- Attended workshops (about language assessment) after you started to work as an English language teacher
- Learning independently, e.g., by reading articles, watching videos
- Other? Please specify:

Section 3: Assessment Activities

From the list below, choose the assessment activities that, as an in-service teacher, you have been involved in:

- Evaluating tests found in textbooks to be used in your school.
- Evaluating the results of your country's regional or national standardized tests.
- Participating in regional or national test development projects.
- Designing your tests to use in your school.
- Administering tests designed by coordinators in your school.
- Administering tests designed by organizations external to your school.
- Participating in assessment committees at your school.
- Participating in assessment committees in your city, region, or country.
- Other? Please specify.

Section 4: Content Selection for the Course

Please state to what extent you think the topics below are important to include in the assessment course you will take:

1. Learning about the history of language testing.
Not important at all__ Slightly important__ Moderately important__ Very important__ Extremely important__
2. Establishing a clear purpose for assessing your students' English.
3. Evaluating whether an assessment instrument is meeting or not its purpose.
4. Evaluating whether a test has the potential to collect information about a determined set of language skills.
5. Planning the design of assessment instruments.
6. Describing clearly the particular skills you want to assess.
7. Designing test items that have the potential to collect precise information about your students' language skills.
8. Designing test tasks that have the potential to collect precise information about your students' language skills.
9. Creating test items and tasks that provide consistent and reliable information about your students' skills.
10. Creating test items and tasks that resemble real-life language use.
11. Evaluating available resources (e.g., technology) for test development.
12. Using available resources efficiently during test development.
13. Evaluating the positive or negative influence that assessment can have on teaching and learning.
14. Evaluating large-scale or standardized language tests.
15. Assessing receptive skills: listening and reading.
16. Assessing productive skills: speaking and writing.
17. Assessing grammar.
18. Assessing vocabulary.
19. Assessing pronunciation.

20. Assessing integrated skills.
21. Assessing different age groups (young learners, teenagers).
22. Assessing learners with special educational needs.
23. Designing traditional tests with true-false and multiple-choice questions.
24. Designing alternative assessments (e.g., self- and peer assessment, portfolios).
25. Administering assessments successfully.
26. Connecting language assessment to language teaching and learning.
27. Interpreting statistical information and scores from language assessment.
28. Assessing students in a bilingual mode.
29. Discussing ethics and fairness in classroom language assessment.
30. Conducting fair, ethical, and transparent assessment practices.
31. Analyzing misuses of language assessment.

Appendix C: Outline of Topics for the LAL Course – With Commentary

- Teach participants these topics in a short handbook; complement the handbook by addressing other topics during the synchronous sessions.
- Most of the topics in the table below come from the questionnaire results.
- The challenges teachers mentioned in the interview must be addressed in synchronous course sessions, especially during group discussions.

| Week | Topic |
|------|---|
| 1 | <i>Introduction to Fundamentals of Language Assessment</i> Commentary: Refer to types of assessment and three key questions: why, how, and what |
| 2 | <i>Qualities of Language Assessment</i> Commentary: Resort to Bachman and Palmer's (1996) usefulness framework |
| 3 | <i>Assessing Receptive Skills</i> Commentary: Address test specs and design here; include traditional test formats; address stats here, too. |
| 4 | <i>Assessing Productive Skills</i> Commentary: Remark on task-based design for authenticity; address stats here, too. |
| 5 | <i>Assessing Integrated Skills</i> Commentary: None |
| 6 | <i>Bilingual Assessment</i> Commentary: Emphasize the need for exploratory translanguaging in ELT and assessment. |
| 7 | <i>Assessing Learners with Special Educational Needs and Disabilities</i> Commentary: Resort to a discussion-based format, given the wide range for this item. |
| 8 | <i>Alternative Assessment</i> Commentary: Connect this topic to challenges teachers expressed; for instance, they can use peer assessment when the number of students is high. |
| 9 | <i>Ethics & Fairness in Classroom Language Assessment</i> Commentary: Since this item has a wide range, involve teachers in self-reflection of unethical or unfair practices. |
| 10 | <i>Relating Language Assessment to Language Teaching and Learning</i> Commentary: After teachers have refined their LAL, elicit how this relationship can occur in the English language classroom. |