HOW

Language Assessment Literacy: Insights for Educating English Language Teachers through Assessment[1]

La Literacidad en Evaluación de Lenguas: Percepciones para la Educación de Docentes de Inglés a través de la Evaluación

Frank Giraldo[2]

Abstract

At some point, language teachers need to be engaged in language assessment in their profession. Because language assessment is such a primary task for teachers, the field of language testing is encouraging research around the knowledge, skills, and principles that are foundational for sound assessment. In this paper, I provide a definition of Language Assessment Literacy (LAL), especially when it comes to teachers, by reviewing existing models. I then discuss ongoing issues in this area and end the paper by offering language teacher educators suggestions for fostering LAL among pre- and in-service teachers. In the article, I argue that, if more LAL initiatives take place, we are collectively raising the status and nature of language assessment and its impact on teachers' professional development.

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

Resumen

En algún momento de su profesión, los profesores de idiomas necesitan involucrarse en la evaluación de lenguas. Debido a que la evaluación es un deber central del docente, actualmente el campo

---

[2] He works at the Foreign Languages Department of Universidad de Caldas in Manizales, Colombia. His main research interest is the interface between language assessment literacy (LAL) and teachers' professional development.
frank.giraldo@ucaldas.edu.co
ORCID ID: https://orcid.org/0000-0001-5221-8245

de la evaluación de lenguas extranjeras se encuentra promoviendo la investigación sobre los conocimientos, las destrezas y los principios fundamentales para procesos de evaluación confiables. En este artículo, doy una definición de la Literacidad en Evaluación de Lenguas (LEL), especialmente en lo que concierne a docentes de idiomas, a través de una revisión de modelos de LEL existentes. Luego hago una discusión sobre los temas que siguen en desarrollo en esta área. El artículo lo termino ofreciendo unas recomendaciones para fomentar la LEL entre docentes en formación y aquellos en servicio. En el artículo argumento que, al existir más iniciativas para la formación docente a través de LEL, como formadores de docentes estaremos mejorando el estatus y la naturaleza de la evaluación de lenguas extranjeras y su impacto en el desarrollo profesional docente.

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

Introduction

Language teacher education programs include varied courses in their curricula. A cursory review of these programs in Colombia and elsewhere indicates that they offer courses on linguistics, second language acquisition, general applied linguistics, and methodologies for language teaching. However, language testing and assessment courses may not be as prominent. In their research, López and Bernal (2009) alerted the field of language teacher education in Colombia, explaining that training in language testing seemed limited or missing in these programs; what was more concerning in their study is that language testing was lacking at both the undergraduate and graduate levels. Today, the situation appears to be changing: Another cursory look at teacher education programs in Colombia shows that assessment courses are offered, either as mandatory or optional. The former is especially good news as authors have suggested that compulsory courses be included so that pre-service teachers are better prepared for their in-service practice in assessment (Lam, 2015; Vogt & Tsagari, 2014).

In fact, educating language teachers in and through language assessment should not be taken lightly. All language teachers (and I know this sounds like an absolute) have to assess their students at some point in their careers: Language assessment is inevitable. The opposite, lack of education in language assessment, may have the potential to lead to malpractice and language learners becoming the victims of poor assessment. Thus, our task as teacher educators is (or should be) to offer high quality education in language assessment, because the positive domino effect can be far-reaching. Additionally, as authors in language testing have argued repeatedly, language testing impacts students, institutions, and society at large, so the field needs professionalization (Fulcher, 2012; Giraldo, 2020; Herrera & Macías, 2015; Inbar-Lourie, 2008; 2013; Shohamy, 2001).

The call for education in language assessment has been clear for around 20 years, starting with Brindley's (2001) proposal: Language teachers need to know about what they assess;

79

how they can design reliable assessment instruments; and how they can use assessment for a positive impact on their contexts. A relevant side of this issue is that teachers themselves have reported a lack of training in language testing (Fulcher, 2012; Giraldo, 2019a; Sultana, 2019; Vogt & Tsagari, 2014). What is more, when asked about their needs in language assessment, teachers expect education in all aspects of assessment (Fulcher, 2012).

Against this background, my purpose with this paper is to problematize the overarching theme in my research and academic career: The interface between Language Assessment Literacy (LAL) and language teachers' (both pre- and in-service) professional development. To do so, I take teacher educators as the main audience for my paper. The paper is divided into three sections: A synthesis of conceptual and research issues in LAL; the present situation with LAL; and recommendations for pre- and in-service teacher education in LAL, with some specific attention to Colombia. Overall, I hope that the ideas in this paper can further contribute to the ongoing research arena in LAL for language teachers.

What is Language Assessment Literacy (LAL)?

Language Assessment Literacy is an area within the wider field of language testing. This construct refers to the different levels of knowledge, skills, and principles that stakeholders (teachers, students, administrators, language testers, teacher educators, and others) have or need to engage in the world of language testing. Davies' (2008) seminal paper suggested that LAL include these components, i.e., knowledge, skills, and principles, an overall conceptualization which is still used in the field (see Giraldo, 2021; and Bohn & Tsagari, 2021).

Authors have also conceptualized LAL through models. For example, Fulcher (2012) places *practice* (knowledge and skills) as foundational in his model; *principles* as elemental for guiding practice; and *contexts* as the broader historical, political, and social impact of language testing. Adding to models, Taylor (2013, p. 410) claims that the LAL profile for teachers should include, on different levels, these aspects: "Knowledge of theory, technical skills, principles and concepts, language pedagogy, sociocultural values, local practices, personal beliefs/attitudes, and scores and decision-making". What Taylor does is to expand the knowledge + skills + principles model and suggests that teachers' contexts of assessment be included in their LAL profile. The inclusion of context, as authors have argued, is fundamental in understanding teachers' LAL (Giraldo, 2020; Hill, 2017; Inbar-Lourie, 2012; Scarino, 2013).

Through a review of the existing conceptual and empirical literature, I offered a descriptor-based definition of LAL (Giraldo, 2018). In this work, I provide specific examples of knowledge, skills, and principles for language assessment that teachers are expected to have. I emphasized then, and should do now, that my conceptual framework did not imply

80

that teachers need to display *all* 66 descriptors; different language assessment contexts require the display of different descriptors. My explication of LAL was guided by these dimensions in language assessment (Giraldo, 2018, p. 187):

- *Knowledge*: Awareness of applied linguistics; theory and concepts; own language assessment context.
- *Skills*: Instructional; design; educational measurement; technological.
- *Principles*: Awareness of and actions towards critical issues in language assessment.

Giraldo (2018) is my most cited, and it has been used to highlight trends in the field (Bøhn & Tsagari, 2021; Butler, Peng, & Lee, 2021), differentiate LAL from its generic term, assessment literacy (Lan & Fan, 2019; Sultana, 2019), and discuss LAL against teachers' professional development and practice (Buendía & Macías, 2019; Rauf & McCallum, 2020). The citations come from researchers both in Colombia and elsewhere. Based on personal communication with fellow teacher educators in Colombia, this paper has also been used in undergraduate and graduate language teaching programs, specifically to revise syllabuses for language assessment courses or as content in them.

More recently, I defined LAL as follows (Giraldo, 2020, p. 190):

In essence, LAL represents the different levels of knowledge, skills, and principles required to engage in language assessment, either from a development perspective (i.e., designing and evaluating language assessments) or from a knowledge perspective, that is, understanding and using scores from assessments to make decisions about people's language ability.

The choice of the words, *different levels*, *engage in*, and *perspective* is deliberate. A group of teachers may need to evaluate whether the introduction of a new test in a school is appropriate. For this task, they will resort to specific arguments (in favor or against) which revolve around either knowledge, skills, or principles in LAL. A student may utilize his/her LAL to argue that a final score on an assignment is fair or unfair (for an example of LAL with young learners, see Butler, Peng, & Lee, 2021). A language tester will need a specific LAL repertoire to develop a reading test for an academic context. Thus, LAL is multidimensional, purpose-driven, and highly context-sensitive (Kremmel & Harding, 2019; Scarino, 2013; Yan, Zhang, & Fan, 2018).

81

When I started my career in language testing and assessment, my view of the matter was operational: I designed tests for different purposes and once the test was read and used, language assessment finished with a score. However, once I delved more deeply into the theoretical and practical implications of language assessment, my view expanded. In other words, to me language assessment is no longer just a test and a score; I fully embrace LAL at large and as defined in this section of the paper: knowledge, skills, and principles. I also

understand language assessment as a key source of feedback for improving learning, its ultimate goal (Bachman & Damböck, 2018).

Fortunately, my view is also present in scholarly work in Colombia. A look at papers in Colombian journals suggests that empirical and conceptual discussions have looked at language assessment mostly from a formative lens; the papers indicate that formative assessment contributes to student learning and positive changes in teaching. In the case of LAL literature in Colombia, authors such as López and Bernal (2009) and Herrera and Macías (2015) have rightfully argued how education in assessment is needed for language teachers. This is a call I joined in 2018 (Giraldo, 2018), and I hope that other Colombian academics further discuss and provide empirical information as to what LAL means in Colombia. Some examples of Colombian authors feeding LAL include Jaramillo and Gil's (2019) student-based research on the use of journal writing for learning about language assessment; and Restrepo's (2020) account of how her pre-service teachers heightened their conception of language assessment theory and practice through this same technique. Although studies exist, the field of LAL in Colombia seems nascent. More studies on pre-service and in-service teachers' LAL in the Colombian context should be –and probably will be– welcomed.

Why LAL Should Matter to English Language Teachers and Teacher Educators

With the data teachers collect from assessment instruments and procedures, they make judgements and interpretations of students' language learning. Thus, if language assessment is done poorly –given inadequate levels of LAL–, then the data may fail to really suggest whether students are learning or not. On the contrary, and as suggested in the LAL models above, teachers who exercise their LAL appropriately are in a position to assess professionally, effect positive change, and contribute to a program's success. Appropriate levels of LAL can ultimately lead to benefits for students' language learning and a positive assessment culture (Giraldo, 2020, 2021; Inbar-Lourie, 2008, 2012).

In the empirical research reporting on in-service teacher needs, studies show that these stakeholders did not receive sufficient education in LAL at the pre-service level and –when the studies were conducted– felt underprepared for doing language assessment (Fulcher, 2012; Giraldo, 2019a; Vogt & Tsagari, 2014). In a research study currently underway, four student researchers and I asked 30 teachers whether they had taken any courses on language assessment. The following Table 1 shows the results to this question:

From the sample in the table, at least eighteen teachers were probably assessing empirically and without theoretical and practical foundations. This echoes something I found in a study (Giraldo, 2019a): I asked five teachers in a case study how they learned about

Table 1. Number of Teachers who have (not) Studied Language Assessment

ITEM	Yes	No
During your pre-service teacher education or in-service professional development, have you studied, formally or independently, any language testing and assessment theory and practice?	39% $n = 11.7$	61% $n = 18.3$

language testing; their answers ranged from *by doing it* to *teamwork with peers*, but not officially or explicitly.

In relation to the trends in teacher needs, language assessment programs seem to be catering to such needs. In a review of professional development initiatives for LAL (Giraldo, 2021), I show how teachers improve their LAL drastically when engaged in courses on language assessment. Overall, programs lead to sizeable changes in teachers' perception of assessment, an increased awareness of what is involved in designing assessments, and even a principled approach towards assessing students (Arias, Maturana, & Restrepo, 2012; Baker & Riches, 2017; Boyd & Donnarumma, 2018; Giraldo & Murcia, 2019). In my 2021 paper, I argue that teacher needs are the foundation upon which these programs should be planned and implemented; this is what I call the intersection between LAL and teachers' professional development.

In conclusion, teacher educators tasked with language assessment courses should raise pre- and in-service teachers' awareness of LAL because:

1. Teachers have expressly reported the need for education in LAL.

2. Pre-service teacher education needs to explicitly address LAL in their curricula.

3. Teachers use information from assessment to make crucial decisions about student learning.

4. There is clear evidence that programs in LAL foster positive change.

5. Better levels of LAL may lead to a positive culture where assessment is used to do good.

83

Twenty Years of LAL: Where is the Field Now, especially for Language Teacher Education?

As I commented elsewhere, the first article to discuss LAL, while not explicitly using this acronym, was Brindley's (2001) proposal for professional development for teaching teachers about language testing. Later, Davies (2008) reviewed language testing textbooks to suggest

the three core LAL components: Knowledge, skills, and principles. It was in 2013 that LAL discussions became more prominent in language testing, with a dedicated edition in the specialized journal *Language Testing*.

After this edition, with LAL scholar Inbar-Lourie (2013a) as the special guest editor, there has been a steady increase in empirical and conceptual information. In synthesis, discussions around LAL have pointed to five areas:

1.   *Drawing the Construct of LAL*

This is an ongoing discussion in the field of language testing. Although LAL's major components appear to be of a consensus, exactly what makes up each component is still in progress. Authors, myself included, have offered specifics for all three LAL components (Giraldo, 2018; Inbar-Lourie, 2013b; Kremmel & Harding, 2019; Malone, 2017; Stabler-Havener, 2018; Taylor, 2013); a look at the available descriptions implies that the construct is far from being completely defined. There may be two reasons for this: On the one hand, each stakeholder group (e.g., teachers, administrators, and even parents) will probably require different LAL profiles (Taylor, 2013; Malone, 2017). On the other hand, as Stabler-Havener (2018) shows, there is thus far no authority –namely a council or board in language testing– that can define, at least, the core knowledge base in LAL. I believe that this is not necessarily a negative development: The lack of a top-down definition may ignite further discussion as to what LAL is. For example, in my case study paper (Giraldo, 2019a), two teachers explained that one of the skills they have is to be caring with students in the assessment process. This skill, as I explain in the paper, is not part of current LAL operationalizations.

2.   *The Stakeholders Involved*

Authors have made it clear that language teachers are a key group in the LAL puzzle (Giraldo, 2021; Inbar-Lourie, 2017; Malone, 2017). However, given the impact of assessment in society, other people need to become conversant with language assessment issues (Pill & Harding, 2013; Malone, 2017). For example, in a recent article, Butler et al. (2021) investigated the LAL of young learners studying English as a foreign language in China. The results in this study showed that even these stakeholders can have sophisticated views of language assessment: The children, aged nine to twelve, reported that they wanted to have more communicative and cognitively demanding tasks for language learning; they were also aware that the construct in their assessments was limited to linguistic aspects rather than communication itself. Butler et al.'s fascinating account attests to the need to hear more voices in LAL, as Inbar-Lourie (2017) welcomes. The next point I address connects this call to the LAL of language teachers.

3. The Need for a Wider Research Agenda in Teachers' LAL

The main argument in my 2020 paper (Giraldo, 2020a) is that we need to expand research methodologies for LAL, both from procedural and conceptual lenses. Fulcher (2012) rightly pointed out the limitations of using a questionnaire to be analyzed statistically. In his study, he noticed that the participating teachers tended to select as many items as were offered to learn about language testing. After echoing this limitation in my paper, I argue that we –language teacher educators– need to see assessment in context (through interviews or observation protocols, for example) so that we can fully grasp teachers' LAL. We should also look at how professional development programs are taught and what impact they have on teachers, once they utilize their improved LAL. Lastly, LAL research should also focus on teacher uses and misuses of language assessments. This last point, I believe, is necessary because the component of *Principles* in LAL is barely touched upon in LAL initiatives (for example Arias et al., 2012). In short, we need to know about dilemmas and problems that teachers face and that can lead to unethical and unfair practices in assessment (Green, Johnson, Kim, & Pope, 2007).

4. The Need for Complementary Literacies in Teachers' LAL

Davies (2008) rightly made it clear that the field of language testing needs to interact with other fields so that it does not fall into isolation. With the unusual developments we have experienced due to COVID-19, technology came to the forefront in teaching, and naturally, assessment. Thus, one complementary literacy that I believe should inform LAL is the technological literacy, more specifically computer and digital literacy (Ng, 2012). Teachers need to use technology efficiently to design, administer, and evaluate assessments online in a way that is fair, transparent, and smooth. Statistical literacy is another field that language teachers need to welcome more. As Brown (2013) explains, teachers may suffer what he calls statistics anxiety. In my own experience teaching pre- and in-service teachers about statistics, the topic is met with apprehension. However, once these stakeholders engage in interpretations rather than the calculations themselves, they start to see the usefulness of statistics, as I argue in Giraldo (2020b).

5. A Call for More Reports of LAL initiatives

85

In Giraldo and Murcia (2019), my colleague and I encourage teacher educators to report on professional development programs for teachers' LAL. The more information we can aggregate from these experiences, the better equipped we are, as teacher educators, to engineer relevant programs for teachers. In my literature review on LAL programs for teachers (Giraldo, forthcoming), I state that, even though the programs I reviewed ($n = 14$) are few, the trends regarding teaching LAL are clear: Design-based courses primarily allow

teachers to develop skills and utilize knowledge consciously; secondarily, LAL courses may ignite awareness of principles such as transparency and fairness. Thus, we should surely benefit from more reports on initiatives that seek to foster language teachers' LAL.

Given these ongoing issues, it may not be surprising to see further developments in LAL research. For example, as I predict in Giraldo (2021), there may be more reports on LAL, as in my last point above. The main implication of the probable future in LAL dynamics is that we, language teacher educators, need to keep abreast of these developments. This can be done through consulting both specialized language testing journals and more general language teaching journals, e.g., *HOW Journal* or *TESOL Quarterly*. In addition to this major recommendation, in the next section I provide other suggestions for language teacher educators to address LAL in their contexts.

LAL Recommendations for Language Teacher Education

The recommendations below come from the conceptual and empirical research I have conducted in LAL, which is fully described in Giraldo (2018), Giraldo and Murcia (2018; 2019), Giraldo (2019a), Giraldo (2020a), and Giraldo (2021). Furthermore, these recommendations come from my experience as a language teacher educator teaching language assessment to pre- and in-service teachers in the Colombian context. Naturally, the recommendations have my personal biases and, consequently, should be analyzed with specific contexts in mind. What may work in one teacher education setting may not work in another.

At a Conceptual Level

As I express in the ongoing issues above, LAL is still in refinement. Models and frameworks will probably be discussed, empirically validated (for example Kremmel & Harding, 2019), and refined. Thus, language teacher educators in Colombia should track the progress of the construct. For example, the issue of principles in LAL is still unresolved, specifically regarding classroom assessment. Research on this LAL component, and how teachers experience it in their assessment lifeworlds, might provide ideas for LAL training with pre- and in-service teachers. In a related manner, when information in LAL is nascent or lacking, we can resort to the general field of educational assessment; to illustrate, Rassoli, Zandi, and DeLuca (2019) provide an extensive treaty of fairness in classroom assessment. Thus, ideas from papers like this one can –and indeed should– inform LAL. Finally, a recommendation mostly applicable to the Colombian context is to include the *Decreto 1290* (decree 1290) which states the general features of assessment for elementary and high schools in Colombia. In Giraldo and Murcia (2018) we learned that pre-service teachers needed to have this content so language assessment can be studied with the Colombian context of assessment in mind.

At a Methodological Level

The professional development programs for language teachers' LAL that I review in Giraldo (2021; forthcoming) have these methodological features in common:

- Successful programs start from teacher needs. The reports by Baker and Riches (2017) and Giraldo and Murcia (2018) provide evidence that by asking language teachers about their needs, wants, or expectations, appropriate LAL programs can be devised.

- There is an explicit, strong design component: Teachers criticize the design of assessment instruments and then create their own by following rigorous design guidelines.

- Theory is best approached through design, not abstractly. Teacher educators in LAL programs embed concepts and theories of language assessment through the critique and design of instruments. For example, how can a robust rubric enhance the validity of interpretations from a speaking assessment? (Arias et al., 2012)

- The programs with more positively impactful results tend to address LAL at large: Knowledge, skills, and principles. However, as I highlight in the present paper and Giraldo (2020a), the issue of principles in LAL still needs to be informed by further empirical research.

Since more LAL programs may surface in the literature, there will most likely be other ideas that can inform local LAL initiatives. Altogether, the methodological suggestions above can be used to plan, teach, or evaluate language assessment programs for pre- and in-service teachers in Colombia and elsewhere. Whereas the knowledge base for defining the construct of LAL is ongoing (Inbar-Lourie, 2013; Stabler-Havener, 2018), methodological trends for teacher education seem warranted. Specifically in Colombia, I suggest that course discussions on language assessment issues include the decree 1290. In a course I taught for pre-service teachers, reported in Giraldo and Murcia (2019), I could see that they could make connections between LAL and this decree and further contextualize and even criticize language assessment in the Colombian context.

At a Procedural Level

In this last part of the paper, I would like to share five teaching techniques that I have used when teaching pre- and in-service teachers in Colombia. The techniques have helped me to help these stakeholders raise awareness of LAL at large. For a more elaborate treatment of these techniques and other principles, please refer to my book on the LAL of pre-service teachers (Giraldo, in press).

1. *Use assessment instruments designed by real teachers*. With the instruments, engage your students in analyzing them from various perspectives. For example, you can ask your students to analyze and speculate on the relative presence of qualities: Validity, reliability, authenticity, practicality, and washback. Alternatively, you can have your students analyze the instrument's items (e.g., a multiple-choice test) or criteria (e.g., a rubric) against design guidelines. My paper (Giraldo, 2019b) includes one example of how this can be done. In it, I explicate why context for language assessment is a key criterion for test critique.

2. *Emphasize the importance of test specifications*. In my experience, when teachers write up a document of specifications for an assessment instrument that they want to use, theoretical discussions such as purposes or constructs and design skills converge. Specifications are also key opportunities to have your students review what they have learned up to a certain point in the course. The added benefit is that they should end up being more careful when it comes to drafting and finalizing an assessment.

3. *Have students conduct small-scale projects*. With your students, design an interview protocol to be used with a teacher. The protocol can include these three core questions: 'Why do you assess?', 'How do you assess?', and 'What do you assess?' Then, your students can design an assessment instrument for this specific teacher. The project can culminate in a discussion involving you, your students, and the interviewed teacher. This project should remark upon the highly contextual nature of language assessment.

4. *Instill alternative assessment procedures in the course*. For instance, after your students design a true-false listening test, you can give them a checklist so they can give feedback to each other's instrument. After the checklist is used, have them talk to each other about how they designed the instrument and provide comments to make each other's instruments better.

5. *Evaluate the course from different angles*. In Giraldo and Murcia (2019), we evaluated a language assessment course for pre-service teachers through class observations, interviews, analysis of students' portfolios, and content analysis of their instruments (see Giraldo, in press). The data we collected from these methods contributed to our understanding of what worked and what needed to be improved in the course in question.

Inbar-Lourie (2008, 2012) argues that an assessment culture is needed in language education, one in which assessment for learning is paramount and in which contextual considerations, rather than external forces, inform language learning. This same notion can be

instilled in language assessment courses: Teacher educators in both Colombia and elsewhere need to invite pre- and in-service teachers to understand language testing and assessment as an approach to monitor and improve language education, nothing else. The opposite –that assessment is used to control and to scare or only to produce a number– should be discouraged. In essence, language teacher educators and language teachers themselves should cultivate their LAL so an assessment culture can flourish. The conceptual, methodological, and procedural suggestions in this paper may aid in doing so.

Conclusions

At the start of this paper, I argue that language assessment is an inevitable part of the teaching profession. The call, then, is to provide language teachers with spaces for them to problematize language assessment and, in doing so, cultivate their Language Assessment Literacy. In this paper, I provided a definition of LAL and overviewed areas which are still under discussion. Based on these, I offered and explained recommendations to plan and teach LAL courses. Echoing other authors, I encourage the inclusion of LAL programs starting at the pre-service level. In the meantime, local LAL initiatives should capitalize on in-service teachers' needs and voices as springboards for professional development, especially for those teachers who have received little to no training in language assessment. These efforts should lead to a heightened awareness of what LAL can bring to the teaching table, and, in the end, help teachers to do the inevitable task professionally.

References

Arias, C. I., Maturana, L., & Restrepo, M. I. (2012). Evaluación de los aprendizajes en lenguas extranjeras: hacia prácticas justas y democráticas [Evaluation in foreign language learning: Towards fair and democratic practices]. Lenguaje, 40(1), 99–126. https://doi.org/10.25100/lenguaje.v40i1.4945

Bachman, L., & Damböck, B. (2018). *Language assessment for classroom teachers*. Oxford University Press.

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. http://dx.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. Vel-

laBriffa (Eds.), Teacher involvement in high-stakes language testing (pp. 105–126). Springer. https://doi.org/10.1007/978-3-319-77177-9_7

Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, C. MacNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 126–136). Cambridge University Press.

Buendía, X. P., & Macías, D. F. (2019). The professional development of English language teachers in Colombia: A review of the literature. *Colombian Applied Linguistics Journal, 21*(1), pp. 89-102.

Butler, Y. G., Peng, X., & Lee, J. (2021). Young learners' voices: Towards a learner-centered approach to understanding language assessment literacy. *Language Testing.* https://doi.org/10.1177/0265532221992274

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25*(3), 327-347. https://doi.org/10.1177/0265532208090156.

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly, 9*(2), 113-132. https://doi.org/10.1080/15434303.2011.642041.

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. (2019a). Language assessment practices and beliefs: Implications for language assessment literacy. *HOW, 26*(1), 35-61. https://doi.org/10.19183/how.26.1.481.

Giraldo, F. (2019b). Designing language assessments in context: Theoretical, technical, and institutional considerations. *HOW, 26*(2), 123-143. https://doi.org/10.19183/how.26.2.512

Giraldo, F. (2020a). A post-positivist and interpretive approach to researching teachers' language assessment literacy. *Profile: Issues in Teachers' Professional Development, 22*(1), 189–200. https://doi.org/10.15446/profile.v22n1.78188

Giraldo, F. (2020b). Statistics for classroom language assessment: Using numbers meaningfully. *HOW, 27*(2), 135-155. https://doi.org/10.19183/how.27.2.541

Giraldo, F. (2021). A reflection on initiatives for teachers' professional development through language assessment literacy. *Profile: Issues in Teachers' Professional Development, 23*(1), 197–213. https://doi.org/10.15446/profile.v23n1.83094

Giraldo, F. (forthcoming). Language assessment literacy and teachers' professional development: A review of the literature. *Profile: Issues in Teachers' Professional Development*, *23*(2).

Giraldo, F. (in press). *Language assessment literacy and the professional development of pre-service foreign language teachers*. Universidad de Caldas.

Giraldo, F., & Murcia, D. (2018). Language assessment literacy for pre-service teachers: Course expectations from different stakeholders. *GiST Education and Research Learning Journal*, (16), 56–77. https://doi.org/10.26817/16925777.425

90

Giraldo, F., & Murcia, D. (2019). Language assessment literacy and the professional development of pre-service language teachers. *Colombian Applied Linguistics Journal, 21*(2). https://doi.org/10.14483/22487085.14514

Green, S., Johnson, R., Kim, D., & Pope, N. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education 2*(7), 999–1011. https://doi.org/10.1016/j.tate.2006.04.042

Herrera, L., & Macías, D. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language. *Colombian Applied Linguistics Journal, 17*(2), 302-312. https://doi.org/10.14483/udistrital.jour.calj.2015.2.a09.

Hill, K. (2017). Understanding classroom-based assessment practices: A precondition for teacher assessment literacy. *Papers in Language Testing and Assessment, 6*(1), 1-17.

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. (2012). Language assessment literacy. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-9). John Wiley & Sons. https://doi.org/10.1002/9781405198431.wbeal0605.

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 presented at the 4th CBLA SIG Symposium Programme, University of Cyprus.

Jaramillo-Delgado, J., & Gil-Bedoya, A. M. (2019). Pre-service English language teachers' use of reflective journals in an assessment and testing course. *Funlam Journal of Students' Research*, (4), 210–218.

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

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197. https://doi.org/10.1177/0265532214554321.

Lan, C., & Fan, S. (2019). Developing classroom-based language assessment literacy for in-service EFL teachers: The gaps. *Studies in Educational Evaluation (61)*, 112-122. https://doi.org/10.1016/j.stueduc.2019.03.003

López, A., & Bernal, 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.

Malone, M. (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. 225–240). Springer. https://doi.org/10.1007/978-3-319-02261-1_16

91

Moraes, I. T., & Furtoso, V. A. B. (2020, October 29-31). *Developing language assessment literacy in pre-service teachers: Analysis of two courses carried out in Brazil.* [Paper presentation]. East Coast Organization of Language Testers. Online.

Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education, 59*, 1065–1078. https://doi.org/10.1016/j.compedu.2012.04.016

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

Rasooli, A., Zandi, H., & DeLuca, C. (2019). Conceptualising fairness in classroom assessment: Exploring the value of organizational justice theory. *Assessment in Education: Principles, Policy & Practice, 26*(5), 584-611. https://doi.org/10.1080/0969594X.2019.1593105

Rauf, M., & McCallum, L. (2020) Language assessment literacy: Task analysis in Saudi universities. In L. McCallum & C. Coombe (Eds.), *The Assessment of L2 Written English across the MENA Region.* Palgrave Macmillan. https://doi.org/10.1007/978-3-030-53254-3_2

Restrepo-Bolívar, E. M. (2020). Monitoring preservice teachers' language assessment literacy development through journal writing. *Malaysian Journal of ELT Research, 17*(1), 38–52.

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. (2001). *The power of tests: A critical perspective on the uses of language tests.* Pearson Education.

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.

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.

Yan, X., Zhang, C., & Fan, 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