

Language Assessment Literacy for Pre-service Teachers: Course Expectations from Different Stakeholders¹

Literacidad en la evaluación de lenguas para futuros profesores: expectativas de un curso desde diferentes partes interesadas

Frank Giraldo and Daniel Murcia^{2*}
Universidad Tecnológica de Pereira, Colombia

Abstract

In the field of applied linguistics, for the past fifteen years, there have been discussions about language assessment literacy (LAL)—the knowledge, skills, and principles related to assessing language ability— (Davies, 2008; Fulcher, 2012). However, the field lacks research on the professional development of language teachers, particularly pre-service language teachers, through training in language assessment. Our paper focuses on the preliminary findings of an action research study whose goal is to identify the impact of a language assessment course for pre-service teachers in a language teaching program in a state university in Colombia. Data collection for the diagnostic stage of the action research cycle used a multiple-choice questionnaire for student needs and wants, an open questionnaire for professors, an interview with an expert, and researchers' journals. Preliminary findings indicate that there is a need to combine theory and practice of language assessment, with an emphasis on current methodologies for language teaching, assessment in bilingual education, and local policies for assessment. The paper highlights recommendations and challenges when designing a language assessment course based on insights from existing literature and includes implications for professional development.

Keywords: Language assessment literacy, language testing, language policies, bilingual education.

¹ Received: January 8th 2018/ Accepted May 8th 2018

² icaros@utp.edu.co; dmurcia@utp.edu.co

Resumen

En los últimos quince años, el campo de la lingüística aplicada ha desarrollado una discusión sobre la literacidad en la evaluación de lenguas (LEL) — los conocimientos, destrezas y principios para evaluar la competencia lingüística — (Davies, 2008; Fulcher, 2012). Sin embargo, el campo carece de investigación en el desarrollo profesional de docentes de lenguas, particularmente profesores de lenguas en formación desde la evaluación de lenguas. Este artículo se enfoca en los hallazgos preliminares de una investigación acción que tiene como objetivo identificar el impacto de un curso en evaluación de lenguas para profesores en formación de un programa para docentes en una universidad pública en Colombia. Para la recolección de datos en la fase de diagnóstico como primer ciclo de la investigación acción se utilizó un cuestionario múltiple para las necesidades y expectativas, un cuestionario abierto para profesores, una entrevista con una experta, y diarios de los investigadores. Los resultados preliminares indican que existe una necesidad de combinar teoría y práctica en evaluación de lenguas, enfatizando en metodologías para la enseñanza de lenguas, la evaluación para la educación bilingüe y políticas locales de evaluación. Se resaltan recomendaciones y retos en el diseño de un curso en evaluación de lenguas basados en tendencias del estado del arte y se discuten implicaciones para el desarrollo profesional docente.

Palabras clave: Literacidad en evaluación de lenguas, evaluación de lenguas, políticas lingüísticas, educación bilingüe.

Resumo

Nos últimos quinze anos, a área da linguística aplicada tem desenvolvido uma discussão sobre a literalidade na avaliação de línguas (LAL) — os conhecimentos, destrezas e princípios para avaliar a competência linguística — (Davies, 2008; Fulcher, 2012). Porém, a área carece de pesquisa no desenvolvimento profissional de docentes de línguas, especialmente professores de línguas em formação desde a avaliação de línguas. Este artigo se enfoca nas descobertas preliminares de uma pesquisa ação que tem como objetivo identificar o impacto de um curso em avaliação de línguas para professores em formação de um programa para docentes em uma universidade pública na Colômbia. Para a coleta de dados na fase de diagnóstico como primeiro ciclo da pesquisa ação se utilizou um questionário múltiplo para as necessidades e expectativas, um questionário aberto para professores, uma entrevista com uma experta, e diários dos pesquisadores. Os resultados preliminares indicam que existe uma necessidade de combinar teoria e prática em avaliação de línguas, enfatizando em metodologias para o ensino de línguas, a avaliação para a educação bilingue e políticas locais de avaliação. Ressaltam-se recomendações e desafios no desenho de um curso em avaliação de línguas, baseados em tendências do estado da arte e discutem-se implicações para o desenvolvimento profissional docente.

Palavras chave: Literalidade em avaliação de línguas, avaliação de línguas, políticas linguísticas, educação bilingue.

Introduction

Educational policies in Colombia include broad guidelines for the assessment of/for learning (*Decreto 1290* from 2009 by the Ministerio de Educación Nacional, MEN). Teachers in schools are expected to assess student learning, and that includes English language teachers. In the case of language learning, the *Currículo Sugerido* (Suggested Curriculum) (MEN, 2016) defines summative and formative assessment as two approaches to language assessment. However, there are no documents to guide this type of assessment, which is unique due to language as a central construct (Inbar-Lourie, 2008). Notwithstanding the importance of assessment for language teachers (Davison & Leung, 2009), the lack of published guidelines derived from empirical evidence exercises a great impact on language teacher education programs. Training for pre-service and in-service teachers, however, is needed and encouraged (Herrera & Macías, 2015; López & Bernal, 2009) in these programs.

Language teachers make decisions based on assessment data, whether this endeavor involves reporting achievement or improving learning (Brown & Abeywickrama, 2010). Since data on language ability is used for decisions, they directly influence learning, teaching, and schools. Against this background, there is a need to improve the language assessment literacy of language teachers, a crucial stakeholder group in language assessment (Giraldo, 2018). In general, language assessment literacy (henceforth LAL) refers to the knowledge, skills, and principles for contextualizing, planning, developing, executing, evaluating, and interpreting language assessments, whether these are devised in the classroom or out of it (Fulcher, 2012).

To foster LAL among teachers, Brindley (2001) proposes that professional development programs include at least three modules: one around the why of assessment (purposes), one around the how (methods for language assessment), and one about the what (the meaning of language ability). While this proposal is indeed welcomed in the language testing field (see Inbar-Lourie, 2012, for example), research seems to be in beginning stages in Colombia. In the past five years, there has been an emergence of LAL research targeting in-service language teachers (see Kremmel, Eberharter, Holzknicht, & Konrad, 2017; Yan, Fang, & Zhang, 2017). However, there is scarce research on LAL development of pre-service language teachers. Specifically in Colombia, Restrepo and Jaramillo's (2017) preliminary findings on pre-service language teachers' LAL has led them to suggest that language testing courses do have a positive impact on participants: these courses

help them become aware of the meaning and purpose of language assessment –the how, why and what in Brindley’s (2001) proposal. As of the time we wrote this article, Restrepo and Jaramillo’s study was the only available research in Colombia targeting the LAL of pre-service language teachers.

Because of the need to further conduct research on LAL (Inbar-Lourie, 2013; Taylor, 2013), we are currently investigating the impact that a language assessment course can have on pre-service language teachers at a state university in Colombia. The language assessment course under scrutiny had its first cohort during the second semester of 2017, and was added to the language teaching program thanks to a recent curriculum modification (see details in *Methodology* below).

Specifically, we report the preliminary findings of the action research study underlying our approach to professional development in LAL. The results in this paper come entirely from the diagnostic stage of the action research cycle, whereby we collected the needs and wants from multiple stakeholders on what they would expect from a language assessment course. Therefore, the diagnostic stage in our study was a needs assessment exercise in curriculum development and sought to listen to stakeholder voices to foster LAL (Inbar-Lourie, 2017b). The overarching goal of our study is to characterize the impact of a language assessment course on pre-service foreign language teachers, while the goal of the diagnostic stage was to characterize the needs and wants for a language assessment course for pre-service teachers.

Literature Review

Assessment Literacy and Language Assessment Literacy

In education, assessment literacy refers to the skills and knowledge for the practice of assessment by stakeholders such as teachers and school administrators. This rather general definition by Stiggins (1995) has broadened in scope and more recently includes issues such as design and evaluation of assessments and knowledge of how students learn; such knowledge, as Brookhart (2011) contends, should inform assessment practices. Language Assessment Literacy (LAL), on the other hand, shares these generic definitions in education; however, it has been stressed that what is specific in LAL is the central construct for assessment: language (Davies, 2008; Inbar-Lourie, 2012).

In general terms, therefore, LAL refers to the knowledge, skills, and principles that stakeholders possess for the exercise of *language* assessment (Davies, 2008; Fulcher, 2012, our emphasis). Such

exercise includes language assessments for specific contexts, purposes, populations, and decisions; as authors argue, language assessment as a process is planned, monitored, and evaluated.

While there is discussion that LAL involves several stakeholders (Pill & Harding, 2013; Taylor, 2013), clearly language teachers remain a central group of stakeholders that need to have adequate levels of LAL (Giraldo, 2018; Inbar-Lourie, 2017b). Consequently, in addition to the core components outlined and explored by Davies and Fulcher, Scarino (2013) argues that language teachers' worldviews are indeed part of their LAL and helps them shape this set of skills.

In a review of LAL for language teachers, Giraldo (2018) proposes a core list of knowledge, skills, and principles that this group is expected to have, according to discussions in the field. Knowledge specifics include theories and methodologies for language teaching and learning, theoretical issues such as validity and authenticity, and personal contextual issues such as institutional guidelines for assessment. Besides, the author brings attention to instructional skills such as collecting classroom data on language development; design skills for closed- and open-ended instruments; statistical skills such as calculation of descriptive statistics; and technological skills such as the use of statistical software. Finally, specific principles include ethics, fairness, democracy, and transparency.

As can be observed, the construct of LAL has embraced a wide variety of issues in language education, and it is now a central discussion in language assessment. In fact, LAL has gained specific entries in language testing literature (see for example, Shohamy, May & Or, 2017).

Related research

Research studies tapping into language teachers' LAL have indicated that, in general, teachers want training in a wide variety of topics. The findings in Fulcher (2012) and Vogt and Tsagari (2014) state that in-service language teachers require training in issues such as test design and skills for test evaluation.

An interesting trend in the research conducted with teachers is that this group prioritizes the practical aspect of language assessment, while not entirely disregarding theory. Fulcher (2012) and Kremmel et al. (2017) have suggested that language teachers want to learn about the construction of items for multiple-choice questions, for instance. However, as their results show, teachers do not seem to be interested in

topics such as ethics and fairness in language assessment, which have been topics of heated debate in the field (Kunnan, 2003).

Other studies have helped language teachers to increase their LAL rather than diagnose their needs. For example, the study by Walters (2010) empowered teachers to become critical towards standards-based testing by means of developing test specifications. In the study by Arias, Maturana, and Restrepo (2012), Colombian in-service teachers improved their assessment practices and made them more valid, ethical, fair and democratic.

In conclusion, LAL is an expanding construct, welcoming what Inbar-Lourie (2017b) calls the genesis of assessment *literacies*. Thus, more research is being encouraged and welcomed to help characterize LAL. As a response to the call for contributing to the LAL discussion, we present a research study that brought together the perceptions of what a language assessment course should include, viewed from the lenses of pre-service teachers and language teacher educators.

Methodology

Context and participants

This action research study was conducted in the B.Ed. in bilingualism with an emphasis on English language teaching from a state university in Colombia. This program modified its curriculum based on the requirements established by the MEN in decrees 02041 and 18583 published during the years of 2016-2017. Because of this modification, a course titled *Seminar in Language Assessment* was added to the program's curriculum for the 8th semester. The seminar meets four hours a week, and the second semester of 2017 was the first time it was offered to students. LAL contents for the course derived from the diagnostic stage of this study.

The participants for the diagnostic stage of the research were professors ($n= 5$) who completed an open questionnaire, and one language teacher education expert; the expert participated in a semi-structured interview. These participants have been full-time professors in the aforementioned program for more than six years, were active members of the curricular transition, and participated in an event in 2017 on the concepts of bilingualism and national language policies.

Another group of participants in the diagnostic stage were pre-service teachers of the program ($n= 30$) whose age ranged from 17 to 25 years old. These participants were enrolled in the *Professional*

Development Course, a subject in the seventh semester, which explores concepts related to national language policies and bilingualism in language education. The students were selected as they were going to be part of the *Seminar in Language Assessment* for the upcoming semester. They expressed their expectations regarding the possible contents for a course in language assessment through a multiple-choice questionnaire, which was administered during the first semester of 2017.

As a response to the diagnostic stage, we concluded that for the first month of the course, the pre-service teachers needed to have an overall review of approaches and principles in language assessment; the following two months, students would deal with the design of language assessments; finally, the last month of the semester should be focused on issues such as bilingual assessment and general policies for assessment in Colombia.

Data Collection and Analysis

The approach in the diagnostic stage of our study was anti-positivistic (Burrell & Morgan, 1979) as it sought to collect data from stakeholders' views. Our methodology followed mostly the tenets of qualitative research because we collected information from an open-question survey, an interview, and researchers' journals. The survey for professors asked them to express what knowledge, skills, and principles for language assessment the students in the course should develop. The interview with the expert had open questions (e.g. What knowledge of language assessment do you think the students of this program should learn about?) and probes (e.g. In your opinion and experience, what areas of language assessment have you noticed students in the program have difficulties with?). Journal entries were guided by key phrases such as "Trends We have Identified in the Data" and "Deciding upon Topics for the Language Assessment Course".

Additionally, our method had a quantitative approach to data collection, through a questionnaire adopted and adapted from Fulcher (2012). Results from this instrument led us to calculate percentages and ranks on language assessment topics to be included in the course. The questionnaire had 28 topics the pre-service teachers could choose from (e.g. History of Language Testing, Writing Assessment Tasks and Items), and an open question: What other topics about language assessment should the course include?

For the qualitative side of the research, we used grounded theory for the data from three instruments: The answers to the open-question

survey, the transcription from the interview, and five comprehensive entries from our researchers' journals. We followed three levels of iterative data analysis in grounded theory (Glasser & Strauss, 1967; Dillon, 2012). *Open coding* –an initial list of trends in the data– occurred as each one of us looked through answers in all instruments and compiled a list of emerging codings; both open lists were then merged to create one open list with fifteen categories; this list was used in the next analysis level. For *axial coding* (which examines and groups trends across open codings), each one of us looked over the merged list and made a second, more specific list of axial codings, using data from all instruments to confirm trends. We then compared both axial lists and calculated how much agreement there was between the two researchers. Inter-Rater agreement for axial coding was 88%; we discussed disagreements and came to a consensus on different codings. This agreement led us to the last data analysis level: *selective codings*, whereby axial codings were further grouped to arrive at major categories supported by research data. Finally, we discussed *selective codings* (i.e. we talked about what they should be labeled, according to data from all instruments) and brought everything down to five major categories that emerged from the diagnostic stage of our study, as follows:

- A. Overall awareness of language assessment
- B. Theoretical foundations in language assessment
- C. Praxis in language assessment, including design, score interpretation, and test critique.
- D. Bilingualism and language policies on learning, teaching, and assessment in Colombia
- E. Washback and impact from language assessment

Results and discussion

The purpose of the diagnostic stage in this action research study was to characterize the knowledge, skills, and principles for language assessment that a group of stakeholders would expect to have in a language assessment course. The data below shed light on the LAL expectations of pre-service teachers and professors from a language teaching program in Colombia. We first present the results from the questionnaire administered to students, then answers from the survey provided by professors, sample data from the interview with the language teaching expert, and journal entries from both researchers. The results below are divided into three major expectations: the first one focuses on designing assessments, the second one on general

language education issues vis-à-vis language assessment, and the third one on other pertinent issues for language assessment.

Overall Emphasis on the Practice of Language Assessment

Table 1 ranks the three top topics chosen by the pre-service teachers in our study. The two topics with the highest rank are related to the design of instruments for assessing language skills, and the third topic refers to the connection between language assessment and specific methodologies for language teaching.

The data stress the need to have praxis in language assessment within a coherent theoretical context for language education (e.g. CLIL). This may indicate that pre-service teachers have a lack of training in the design of instruments for assessing language skills. Besides, this group of stakeholders may consider the practical side of assessment as a priority, rather than a focus on theory they may already be familiar with, however superficially. To illustrate, the instrument uses the word “Design”, which may trigger among the students the idea of practical –not theoretical– language assessment. Lastly, the program the students are enrolled in has approached them through CLIL teaching, hence the evident expectation to learn about CLIL assessments.

Table 1. Three Top Topics from Questionnaire for Pre-Service Teachers

Rank	Topics	Respondents	%
1	Design of listening and reading assessments	28	93.3%
2	Design of speaking and writing assessments	27	90%
	Language assessments and methodologies (CLIL; task-based instruction)*	27	90%

Based on the quantitative results of Table 1, the conclusion that the pre-service teachers want a course that deals with the practical aspects of language testing mirrors what scholars have found in different parts of the world. Studies investigating in-service language teachers’ LAL (Kremmel et al., 2017, Yan et al., 2017) have shown that teachers want to develop skills for item-writing, which is consistent with the results we are presenting. What is interesting about the overall results of these studies is that both populations seem to have similar interests for their professional development in language assessment.

Table 2, on the other hand, lists the bottom three topics ranked by the pre-service teachers. Based on these data, the topics with fewer respondents are not altogether connected to the design of language assessments; in other words, topics such as ethics and test administration generally occur once an assessment has been designed and used; topic ranked 13 (*test specifications*), however, deals with a theoretical-practical aspect of language testing, which is writing instructions on how to write a test item or task. In this case, there may either be a contradiction in students' responses, or lack of knowledge of what test specifications are. The data in Table 2 suggest that the pre-service teachers would expect little attention to these matters, or it may be the case that they are not aware of what these issues imply in language assessment. The data also seem to reiterate what we present in Table 1: these pre-service teachers appeared inclined towards the design of language assessments. It may also be the case that Ethical Considerations and Test Administration are topics the pre-service teachers could be taking for granted; according to these stakeholders' responses, these two topics should not be prioritized in a language assessment course.

Table 2. Three Bottom Topics from Questionnaire for Pre-Service Teachers

Rank	Topics	Respondents	%
13	Writing test specifications/blueprints	10	33.3%
14	Language test administration	6	20%
	Ethical considerations in language assessment	6	20%

Overall, the results above show a lack of interest in specific language assessment issues, a trend which has been observed in other studies. For example, in the study by Fulcher (2012), participating in-service teachers showed little interest in test administration. Similarly, in-service teachers in Kremmel et al. (2017) and Yan et al. (2017) showed low interest in ethics and fairness in language assessment. However, in Fulcher's study, there was a high level of interest for training in item specifications, which differs from our study. In conclusion, it appears that pre- and in-service language teachers do not seem to want training in administrative and ethical issues for language assessment.

Lastly, Table 3 displays the list of subsequent highly ranked topics which link language assessment with theoretical issues. All topics have the same number of respondents (21 pre-service teachers) and include the selection of contextually sensitive type of language assessment, exploration of validity and how to validate language assessments, and how language assessment operates within bilingual contexts. The percentages in this table (70%) reveal that the pre-service teachers in our study seem to have some previous conceptions of theoretical aspects of language assessment and, most interestingly, that they perceive these topics as fundamental in their professional development. In the case of the first topic (assessments for own purposes), we believe this item directly addresses a need: the participants might want to consider assessments for their future practice, so selecting these assessments might be a relevant task for which LAL is needed. As for the second topic, we infer that students have studied validity before and, given their background, attach importance to it. The words validity and validation are crucial in language assessment, and these students may be aware of this fact. Finally, the program the participants are studying has the word *bilingualism* in its name, which should be a reason why assessment in bilingual contexts is a topic that ranks high.

Table 3. Sample of Topics Related to Theoretical Issues in Language Assessment

Rank	Topics	Respondents	%
5	Selecting assessments for your own purposes.	21	70%
	Validity and validation of language assessments.	21	70%
	Language assessment in bilingual contexts**	21	70%

The tendency in table 3 depicts the need of pre-service teachers for in-depth exploration of the selected topics for a language assessment course. These theoretical issues align with the knowledge dimension in the core list of LAL by Giraldo (2018). However, as the author and others emphasize (Davies, 2008; Fulcher, 2012), the three components of LAL should not be seen separately but integrated. All three tables above show that, at least, attention to knowledge and skills must be present in the language assessment course for these pre-service teachers.

Close Connection between Theory in Language Education and Practice in Language Assessment

The data below comes from qualitative research instruments, specifically answers to the open question survey, and insights from the interview and researchers' journals. Overall, the data samples below confirm the stakeholders' need to have a course that combines theory and practice in language assessment. Particularly, the samples highlight practical aspects such as assessment methods and theoretical issues like knowledge and purposes for rubrics. In the survey, professor#4 states his wants: "Variety of methods for assessing second language proficiency (beyond testing). General policies for assessing learning in the Colombian Education System and its articulation with theoretical principles." The language education expert comments on the mix between theory and practice:

They (students) should develop the competence for doing that, so meaning that their theoretical knowledge: how is it, or what is it and what are their purposes and um, what are their criteria for developing them but also the practical part where they can develop rubrics

Further, the reflection below comes from one of our researcher's journals and further highlights the need for the theory-practice connection:

Entry 2, Topic 2, (17/07/17): It is crucial then, the capacity to link notions like testing and rubric design, marking, adoption and adaptation of materials, and language teaching approaches which are connected from theoretical perspectives to evaluation and assessment and which have been superficially covered throughout the semesters of the student-teachers' academic program to their current practice in institutions or schools.

Our first interpretation from these results is that stakeholders are not conceiving language assessment as a grade or test but rather as a universe of closely connected practical and theoretical issues. The range of topics to be included in the course may be signaling this fact. Furthermore, as it has been shown in other studies, when teachers have the chance to choose topics to improve or learn about, they tend to require as many topics as possible; this trend is also evident in the data we are presenting. Specifically, the topics these stakeholders want to include not only language assessment as an act of doing or using tests, but as an integration of contextual matters in language education; see results such as variety of assessment methods, assessment policies in Colombia, and methodologies for language teaching.

Similar attention to theory and practice of language assessment has in fact been highlighted by language assessment scholars (Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2013). They argue that language teachers need to critically relate knowledge, skills, and principles for their language assessment practice. Specifically, in the study by Vogt and Tzagari (2014), the in-service teachers reported that they needed training across the spectrum of language assessment. Taken together, the findings in our study and those of others' studies seem to point to a comprehensive need for furthering professional development through training in language assessment for language teachers.

Other Issues for Training in Language Assessment

For this last section of our findings, the data come from journal entries, the survey for professors, the open question in the questionnaire for students, and the interview with the expert. Additionally, we interconnect data in this section with data from tables in the previous two sections.

We identified particularities that the language assessment course should address. According to participants in our study, the course should include information about methodologies for language teaching (e.g. CLIL -Content and Language Integrated Learning) and their relationship with language assessment. As one of the entries in the journal shows (Entry 1, Topic 1, 20/04/17): "Most of the students are currently involved in projects that use CLIL as the instructional approach and they still require much guidance on the conceptions of testing when content is embedded in the educational practices of language teaching." This trend is reiterated in Table 1, specifically the three top themes from the questionnaire for students.

The stakeholders expect information about general policies for assessment in the Colombian context. Professor#1 expects "Regulations or agreements from the national authorities regarding testing and evaluation." Professor#4 echoes: "General policies for assessing learning in the Colombian Education System and its articulation with theoretical principles." Finally, Student#20 expects the following: "National policies for language assessment."

Another issue the course should target, according to the expert, is the impact of language assessment on learning and teaching. The sample data below shows her view on this matter.

[Impact on teaching] And the thing is that evaluation from my point of view should also include how am I doing as a teacher. What is the effect of my course. Is it being successful?

[Impact on learning] assessment is not something instrumental, assessment is something that also has uh effects and consequences on the person who's assessed.

Lastly, assessment of bilingualism is a topic that should be part of the language assessment course reported in the present study. As Professor#5 suggests: "I would include evaluation in bilingualism: translanguaging and evaluation of L1 and L2." Table 3 above reiterates this theoretical issue in language assessment.

The combination of the four issues above seem to suggest that stakeholders have a determination for contextualizing professional development through language assessment. As may be suggested by the data, each of these issues could have some degree of relation to the curricular transition the program experienced, especially when it comes to bilingualism and language policies in Colombia. For instance, the samples above address CLIL as an approach which has been explored in the *Curriculum Design Course* given its close relation to bilingual education and language policies. There is also evidence to connect national language policies in Colombia (a topic students address in their practicum courses) and language assessment practices. Lastly, the language teaching expert expresses her views on the impact of language assessment for pre-service teachers; she seems to be highlighting a comprehensive purpose for assessment, rather than seeing it only as a grading tool.

Data displayed above strongly suggest the need to merge theory, practice, and general assessment issues in language assessment. Therefore, it is not surprising that the participants in our study want to have varied topics in the course. This is indeed a trend in the literature on LAL, whereby in-service teachers have expressed their wish to learn about diverse topics proposed in questionnaires (for an example, see Vogt & Tsagari, 2014). An interesting emergent finding from the present study is that participants expect to learn about language assessment within a broader context for language education, which includes approaches such as task-based language teaching, CLIL, and bilingual education.

Our findings point to what may be an overlooked area in the research on LAL. Participants in our study believe it is important to learn about language assessment alongside Colombian policies for assessment in education. Discussions about LAL have not explicitly

highlighted how general education policies may or may not have an impact on language assessment, but Davison and Lynch's (2002) idea of test mandate (a norm that officially establishes the need for a test) may come close to such discussions.

However, there are no discussions as to how documents like the *Decreto 1290* can illuminate and/or constrain language assessment, particularly in the case of Colombian language teachers, assessment in bilingual education, and contemporary methodologies for language teaching. Specifically, existing English language policies like the *Suggested Curriculum* have not explicitly linked the *Decreto 1290* to overall paradigms in language assessment; for instance, formative and summative assessment.

This lack of discussion serves as an argument for Inbar-Lourie's (2017b) call to listen to different stakeholders who can collectively understand "assessment targets, tools, procedures, analysis and intended but also unintended consequences" (p.267) around language assessment practices. Thus, we believe that a research avenue worth pursuing should include studies on how general assessment policies—along with general approaches to language teaching—coexist with language assessment issues.

Conclusions and recommendations

Language assessment literacy is an expanding area of debate in applied linguistics. The topics and research presented at the 39th Language Testing Research Colloquium (International Language Testing Association) held in Bogota in July 2017 attest to this fact. In the colloquium's closing plenary, Inbar-Lourie (2017) argued that, instead of LAL, the field of language testing should embrace the genesis and development of *LALs*. That different stakeholders—as shown in our study—want to see national policies for assessment in a solely language assessment course supports Inbar-Lourie's invitation. More importantly, our study reiterates language assessment as a key dimension of language teacher education, best encapsulated in this excerpt from the interview with the expert:

They look at assessment as something that is final with a purpose of just giving students a grade. And that is one of the areas that I, I think that we, that the course should emphasize on: On the purposes of assessment, on the usefulness of assessment, other than just giving students a grade.

The findings in our study show that the participants would like to have a course that deals primarily with practical issues in language assessment (e.g. item writing), even though not avoiding issues such as content-based assessment, purposes of assessment, Colombian assessment policies, and bilingualism. In fact, the data lead us to infer that these topics should be critically combined in the language assessment course to be administered. Therefore, the course should contribute to pre-service teachers' overall awareness of what language assessment means, its implications, and its overall impact on language learning and teaching. In turn, such heightened LAL should contribute to their overall professional development.

Regarding the issue of bilingualism, we suspect that the contents stakeholders want in the course under scrutiny were influenced by contextual factors. Recently, the program where the course belongs went through a reform, and its name now explicitly includes the word bilingualism, which is now reflected on the expectation to have assessment of this construct in the course.

Finally, we wish to highlight two recommendations and a related limitation in our study. First, none of the participants mentioned ethics and fairness as principles that a language assessment course should include, despite the allusion of such themes in LAL literature (Davies, 2008; Kunnan, 2003). We wonder why this was the case and warn that inclusion of these two topics, if not identified in a needs analysis, will merely be a judgement call. Second, Fulcher's (2012) questionnaire does not specify what classroom assessment (a topic in his questionnaire) involves, i.e. subcategories of classroom assessment. Because of this lack of specificity, we realized that there was no information on topics such as portfolio assessment, arguably a pertinent issue for language teachers' LAL. Thus, we feel this was a limitation –and indeed an artifact–of our research, and recommend that researchers wanting to use ready-made questionnaires have a critical stance towards these instruments to adapt them, even if they come from experts.

Overall and in tandem with Inbar-Lourie's (2017b) call to further research in LAL, we especially encourage language teacher education programs to share information on the development and impact of LAL courses, let alone when LAL is expanding and welcoming.

References

- Arias, C., Maturana, L., & Restrepo, M. (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.
- Brindley, G. (2001). Language assessment and professional development. In C. Elder, A.
- Brown, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 126-136). Cambridge: Cambridge University Press.
- Brookhart, S. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3-12.
- Brown, H. D. & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practice*. New York: Pearson Education.
- Burrell, G. & Morgan, G. (1979). *Sociological paradigms and organisational analysis. Elements of the sociology of corporate life*. England: Ashgate.
- Davidson, F., & Lynch, B. K. (2001). *Testcraft: A teacher's guide to writing and using language test specifications*. New Haven (US): Yale University Press.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, (43)3, 393-415.
- Dillon, D. R. (2012). Grounded theory and qualitative research. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-10). Oxford, UK: John Wiley & Sons.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132.
- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179-195.
- Glasser, B. J. & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New Jersey: Aldine Transaction.

- Herrera, L., & Macías, D. F. (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.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402.
- Inbar-Lourie, O. (2012). *Language assessment literacy*. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-9). Oxford, UK: John Wiley & Sons.
- Inbar-Lourie, O. (2012). Language assessment literacy. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-9). Oxford, UK: John Wiley & Sons.
- Inbar-Lourie, O. (2013). Guest Editorial to the special issue on language assessment literacy. *Language Testing*, 30(3) 301-307.
- Inbar-Lourie, O. (2017, July). *Language assessment literacies and the language testing community: A mid-life identity crisis?* Paper presented at the language testing research colloquium, Bogotá, Colombia.
- Inbar-Lourie, O. (2017b). Language assessment literacy. In E. Shohamy, S. May, & I. Or (Eds.), *Language Testing and Assessment (third edition), Encyclopedia of Language and Education* (pp. 257-268). Cham (Switzerland): Springer.
- Kremmel, B., Eberharter, K., Holzknrecht, F., & Konrad, E. (2017, July). *Enhancing language assessment literacy through teacher involvement in high-stakes test development*. Paper presented at the language testing research colloquium, Bogotá, Colombia.
- Kunnan, A. (2003). Test fairness. In M. Milanovic & C. Weir (Eds.), *Selected papers from the European Year of Languages Conference, Barcelona* (pp. 27-48). Cambridge: Cambridge University Press.
- 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.
- Ministerio de Educación Nacional de Colombia (2009). *Evaluación del aprendizaje y promoción de los estudiantes en los niveles de educación básica y media*. Bogotá (Col.): Ministerio de Educación Nacional.

- Ministerio de Educación Nacional de Colombia (2016). *Diseñando una propuesta de currículo sugerido de inglés para Colombia*. Bogotá (Col.): Ministerio de Educación Nacional.
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381-402.
- Restrepo, E. & Jaramillo, D. (2017, July). *Preservice teachers' language assessment literacy development*. Paper presented at the language testing research colloquium, Bogotá, Colombia.
- 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.
- Shohamy, E., May, S., & Or, I. (2017). *Language testing and assessment: Third edition. Encyclopedia of Language and Education*. Cham (Switzerland): Springer.
- Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238-245.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402.
- Walters, F. (2010). Cultivating assessment literacy: Standards evaluation through language-test specification reverse engineering. *Language Assessment Quarterly*, 7(4), 317-342.
- Yan, X., Fan, J., & Zhang, C. (2017, July). *Understanding language assessment literacy profiles of different stakeholder groups in China: The importance of contextual and experiential factors*. Paper presented at the language testing research colloquium, Bogotá, Colombia.

Authors

***Frank Giraldo** holds a B.Ed. in English Language Teaching from Universidad Tecnológica de Pereira (UTP), Colombia; an MA in English Didactics from Universidad de Caldas, Colombia; and an MA in Teaching English as a Second Language, from University of Illinois at Urbana-Champaign, the United States. He works as an academic consultant for the Instituto de Lenguas Extranjeras (ILEX) at UTP and as a professor at the Licenciatura en Bilingüismo con Énfasis en Inglés (UTP).

Daniel Murcia holds a B.Ed. in English Language Teaching and an MA in linguistics from Universidad Tecnológica de Pereira (UTP), Colombia. Daniel belongs to the academic staff at the Licenciatura en Bilingüismo con Énfasis en Inglés in UTP, Risaralda. He works as a professor and coordinates the intercultural area of the program.