# Examining the Assessment Practices of Foreign Language Novice Teachers

## Investigación de las prácticas de evaluación de profesores principiantes de lenguas extranjeras

**Gabriel Cote Parra**

Universidad de Pamplona, Pamplona, Colombia


**Alexis A. López**

Educational Testing Service, Princeton, USA

This paper reports a mixed-methods study at a public university in Colombia. It describes the classroom assessment practices and challenges of 75 novice foreign language teachers. To gather the quantitative data, the participants completed an online survey. For the qualitative data, 11 key informants participated in one-on-one online interviews. Findings revealed that novice teachers predominantly used summative assessment in the classroom and aligned their assessment instruments to large-scale tests. Moreover, novice teachers faced many challenges with classroom assessment, including determining how to assess their students, developing assessment instruments, and interpreting and using assessment scores to inform teaching and learning. In conclusion, novice teachers need more knowledge, skills, and support to handle daily assessment-related tasks.

*Keywords:* classroom assessment challenges, classroom assessment practices, novice foreign language teachers

Este artículo presenta un estudio de métodos mixtos, desarrollado en una universidad pública de Colombia, que describe las prácticas y desafíos de evaluación en el aula de 75 profesores principiantes. Para la recolección de información cuantitativa, se utilizó un cuestionario en línea y, para la información cualitativa, se usó una entrevista en línea con once docentes. Los resultados muestran que los participantes utilizaron predominantemente la evaluación sumativa y alinearon sus instrumentos con evaluaciones a gran escala. Además, se evidenció que los participantes enfrentaron muchos desafíos para evaluar a sus estudiantes, desarrollar instrumentos de evaluación, e interpretar y usar puntajes de evaluación para informar la enseñanza y el aprendizaje. En conclusión, los participantes necesitan más conocimientos, habilidades y apoyo para afrontar las tareas diarias relacionadas con la evaluación.

*Palabras clave:* desafíos de la evaluación en el aula, prácticas de evaluación en el aula, profesores principiantes de lenguas extranjeras

Gabriel Cote Parra https://orcid.org/0000-0002-0095-8265 • Email: gcote@unipamplona.edu.co
Alexis A. López https://orcid.org/0000-0002-4616-1091 • Email: alopez@ets.org

## Introduction

In Colombia, five-year undergraduate foreign languages programs provide preservice teachers with educational training on five components: general component (e.g., literacy, research, foreign language learning, information and communications technologies, mathematical reasoning, and citizenship competence), specific and discipline component, pedagogic component, and the didactics of the discipline (Resolución 18583, 2017). Although undergraduate students are provided with standard general education foundations and the specialized knowledge and skills to teach foreign languages, "some teacher education and teacher-training programs in Colombia do not offer extensive training in language assessment" (López & Bernal, 2009, p. 55). In most of these programs, "language assessment preparation from theoretical and practical perspectives appears to be limited to one or two units in the methods courses" (Herrera & Macías, 2015, p. 305). For example, in the foreign languages program where this research took place, assessment represented a single unit in the didactics course. However, two more didactics courses and one assessment course were added due to the latest curricular reform. Hence, future in-service teachers are expected to be better prepared regarding foreign language assessment's theoretical foundations, skills, and principles.

This paper reports one part of a more extensive mixed-method study on language assessment literacy (LAL). It attempts to describe the perceived experiences of novice foreign language teachers' assessment practices and challenges. The participants graduated recently from a foreign language program at a public university in Colombia and reported having two to five years of teaching experience.[1]

---

[1] Several authors have defined the term "novice teacher" differently. For instance, this term has been used interchangeably with others, such as beginning teachers, neophytes, and preservice teachers (Kim & Roth, 2011), newly qualified teachers (Hayes & Chang, 2017), and novice-service teachers (Farrell, 2012). However, for this study, novice teacher refers to a teacher with less than five years of teaching experience after graduation.

This study aimed to examine the current assessment practices of novice foreign language teachers in their first five years of teaching. In addition, this study also examined the challenges novice foreign language teachers face when assessing their students in the language classroom. Thus, this study addresses the following research questions:

1. What characterizes the classroom assessment practices of novice foreign language teachers?
2. What challenges do novice foreign language teachers face when assessing students in the classroom?

## Literature Review

### Classroom Assessment Practices

Classroom assessment is integral to language instruction (Butler, 2022). According to Leung (2005), classroom assessment means assessing student learning within a classroom setting. It is important to note that classroom assessment includes assessments used for either summative or formative purposes (Brookhart, 2004). Teachers are vital in planning what, how, and when to assess and interpret classroom assessment to guide teaching and learning (Cheng et al., 2004; Leung, 2005; López, 2010; Rea-Dickins, 2001). Rea-Dickens (2004) refers to teachers as "agents" of assessment responsible for observing, evaluating, and interpreting their students' progress. Although teachers play a critical role in classroom assessment, students must be actively involved to enhance their understanding of how assessment works (Butler, 2022; Inbar-Lourie, 2008).

Classroom assessment generally includes three basic steps: (a) collecting evidence of student learning, (b) interpreting or judging the evidence, and (c) making use of the evidence (Leung, 2005). To make use of classroom assessment, teachers engage in different practices to determine the purpose of the assessment; plan the assessment; select methods and instruments; implement, score, and report the assessment, and make use of the assessment information (Hill & McNamara,

2011; Leung, 2005; Wolf & López, 2022). However, the classroom assessment practices of language teachers vary greatly (Rea-Dickins, 2001; Yin, 2010).

Among some of the factors that influence the classroom assessment practices of language teachers, we find teachers' beliefs, professional training and development, teaching experiences, and LAL (Cheng et al., 2004; Yin, 2010). Several scholars have highlighted the importance of professional development to help teachers develop their LAL and acquire the skills needed to assess and enhance students' learning (e.g., Giraldo, 2018; Inbar-Lourie, 2008; Mohammadkhah et al., 2022; Tajeddin et al., 2022). Other factors that influence the classroom assessment practices of language teachers include the school's curriculum, culture, and climate (Rogers et al., 2020). In addition, teachers' workload, time constraints, and available resources also impact their classroom assessment practices (Wolf & López, 2022).

Although many scholars agree on the importance of classroom assessment, not many studies have examined the classroom assessment practices of foreign language teachers, particularly novice teachers (Cheng et al., 2004; Hill & McNamara, 2011; Wang, 2017). Thus, it is critical to examine novice teachers' classroom assessment practices to effectively determine the best way to support them in their assessment duties.

## Language Assessment in Colombia

There has been an increased interest in language assessment in Colombia in recent years, as indicated by the number of studies published in research journals. These studies have focused on different aspects of language assessment. For example, many have examined LAL, including describing the implications of LAL for language teachers (Giraldo, 2018), evaluating research constructs and methodologies on LAL (Herrera & Macías, 2015), investigating LAL initiatives and professional development (Giraldo, 2021), and examining the training preservice and in-service teachers receive in different assessment-related areas (Giraldo & Murcia, 2018).

Other studies have focused on issues related to large-scale tests, such as examining the validity of a Colombian national English language proficiency exam (López & Janssen, 2010) and the washback or impact of large-scale language tests on teaching and learning (Muñoz & Álvarez, 2010). Moreover, a few research-based papers describe the assessment of specific skills and competencies such as students' speaking skills (Duque-Aguilar, 2021), self-evaluation of the grammatical range and grammar accuracy (Caicedo Pereira et al., 2018), teachers' and students' perceptions on language assessment (Herrera Mosquera & Zambrano Castillo, 2019), and understanding effective ways to assess speaking skills (Duque-Aguilar, 2021).

Finally, a few studies have focused on classroom assessment in English language teaching. For example, exploring how language teachers use formative assessment to inform teaching and learning (López, 2010), investigating the influence of peer assessment (Gomez Sará, 2016), examining the challenges of designing and administrating tests (Ramírez, 2020), and understanding the effects of classroom (face-to-face) assessment practices (Lopera Medina, 2015). More recently, Hernández-Ocampo's (2022) review of five studies on assessment highlights that teachers observed "fair and democratic assessment practices" and learners' autonomy is fostered through "peer- and self-assessment practices." However, according to the author, "more teacher education in language assessment" is needed (p. 231).

Although the abovementioned studies describe language assessment in Colombia, they are not explicitly centered on novice teachers. Therefore, more studies are needed to examine language teachers' classroom assessment practices, particularly the assessment practices of novice teachers. These studies can inform preservice and in-service training programs to support language teachers using classroom assessment more effectively.

## Method

We adopted a sequential explanatory design (Tashakkori & Teddlie, 2003) that allowed us to collect and analyze quantitative and qualitative data. Although this design prioritizes the quantitative data, we uniformly integrated both while analyzing and interpreting the data.

### Participants

The target sample for this study were novice English and French foreign language teachers who had recently graduated from a foreign language program at a public university in Colombia and had two to five years of teaching experience. The sampling plan for this study included teachers who graduated from this program between 2013 and 2018. After inviting 190 novice foreign language teachers to respond to an online survey, 75 agreed to participate and completed an online survey (a response rate of 39.47%). These teachers varied in terms of years of teaching experience and the type of institution they worked in (private school, public school, language institute; see Figure 1). The average age of the participants was 28; 52% of the teachers were women, and 48% were men.

**Figure 1.** Survey Participants' Background Information ($N$ = 75)
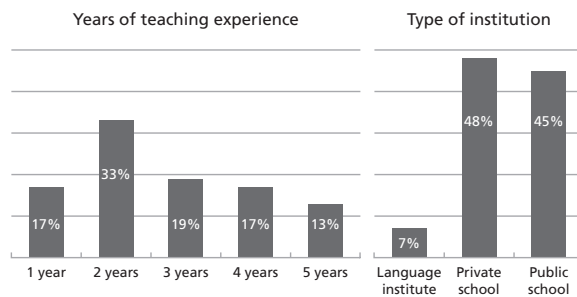


Of those 75 teachers, we selected 11 to participate in one-on-one online interviews following a maximal variation sampling (Creswell, 2002). These 11 teachers displayed different dimensions of four characteristics we thought critical in examining novice teachers' assessment practices and challenges: (a) year of graduation (between 2013 and 2018); (b) different workplaces (private and public sector); (c) type of institution (school and language institute); and (d) location (rural and urban schools). Table 1 provides additional information about the novice teachers who participated in the interviews.

**Table 1.** Interviewees' Additional Background Information

| Participant | Degree completion | Years of teaching experience | Type of school | Location |
|---|---|---|---|---|
| Teacher 1 | 2017 | 2 | Language institute | Urban |
| Teacher 2 | 2017 | 2 | Private primary school | Semi-urban |
| Teacher 3 | 2018 | 1 | Private high school | Semi-urban |
| Teacher 4 | 2016 | 3 | Public high school | Semi-urban |
| Teacher 5 | 2016 | 3 | Public high school | Semi-urban |
| Teacher 6 | 2016 | 3 | Public university | Semi-urban |
| Teacher 7 | 2014 | 5 | Public high school | Rural |
| Teacher 8 | 2017 | 2 | Language institute/Private school | Urban |
| Teacher 9 | 2017 | 2 | Language institute | Urban |
| Teacher 10 | 2014 | 5 | Public high school | Urban |
| Teacher 11 | 2015 | 4 | Language institute | Urban |

### Procedures

We employed an online survey to gather information about how teachers used assessment and their challenges in assessing their students in the classroom (see Appendix A). The survey was based on the instrument used by Cheng et al. (2004). However, we made some modifications to reflect the Colombian context. We piloted the online survey with ten last-year students in the foreign language program to check the questions' clarity of direction and relevance. The survey consisted of 29 items organized into three sections: background information (8 items), current assessment practices (8 items), and challenges in using classroom assessment (13 items). On average, it took participants 20 minutes to complete the online survey.

Eleven of the 75 teachers also participated in a one-on-one online semi-structured interview. Each interview lasted approximately 60 minutes. All the interviews were audio recorded and later transcribed verbatim. We used an interview protocol to discuss the respondents' instructional context, current classroom assessment practices, and challenges when assessing their students in the classroom (see Appendix B).

### Data Analysis

Descriptive statistics (e.g., frequencies, percentages) were calculated on SPSS for each item in the three sections of the online survey. We coded each interview transcript manually and independently. We first used structural coding (Saldaña, 2009) to identify categories that indicated the assessment practices of novice foreign language teachers, how they developed their classroom assessment practices, and the challenges they faced. Then, we used descriptive coding for each structural code we identified (Saldaña, 2009) to help establish recurring themes. Descriptive coding refers to assigning labels to data to inventory critical issues within each structural code. After completing the coding, we met to compare our ideas and codes to determine if we had arrived at similar data interpretations. All the disagreements in

the coding were resolved through discussion to reach a consensus, as suggested by Strauss and Corbin (1990).

## Findings

The findings are organized into two themes related to the research questions. First, we discuss the main characteristics of novice teachers' classroom assessment practices. Then, we present the main challenges they faced when assessing their students in the classroom.

### Novice Teachers' Classroom Assessment Practices

The survey and interviews revealed three main features that characterized novice teachers' assessment practices. First, all participating teachers mainly used classroom assessment for summative purposes, meaning that classroom assessment was used primarily for accountability (i.e., to assign grades or to determine who passed or failed a course). This classroom assessment practice is evident in the following comments. Teacher 3 explained how she typically assessed her students in the classroom: "It would be through the bimonthly exam that is done at the end, in which 25% is assigned to each skill: listening, reading, speaking, and reading competence." Although teachers assess multiple skills in their tests, each skill is considered separately.

Similarly, Teacher 4 discussed what is usually included in his classroom assessments: "At the end of the period, a summative test is also done. It includes all the themes that were covered in the period." Furthermore, Teacher 8 talked about how teachers in her school used summative assessment: "We use end-of-module exams at the [institution]. They are the most important exams because they decide if the students fail or pass the module."

The second main feature is that all the teachers used traditional paper-and-pencil tests in the classroom, generally based on multiple-choice, true and false, and open-ended questions. Finally, most assessment practices were teacher-centered; that is, the assessment is done exclusively by the teacher with little to no active

student participation. This usually serves to monitor learning rather than promote learning. According to the teachers, they either designed their assessments, adapted existing assessments, or used assessments provided by their institutions.

Conversely, this study also found that novice teachers sometimes used other classroom assessment practices (see Figure 2). A few teachers (23%) used assessment for formative purposes to inform teaching and learning, and many teachers (63%) used assessment for diagnostic purposes at the beginning of the instructional process. For instance, Teacher 2 commented about using other types of assessments in his classroom: "We also use a diagnostic assessment at the beginning of the school year." In the interviews, a few teachers reported using alternative assessments such as projects, role-plays, and games to assess their students. To highlight using alternative assessment, Teacher 4 explained the following: "We also use project-based assessment, such as theater. For the past two years, our project has been a municipal theater meeting for various schools, both public and private." Although these alternative assessment practices allowed novice teachers to assess their learners differently using meaningful interactive tasks, they were used sparingly.

**Figure 2.** Other Classroom Assessment Practices (*N* = 75)
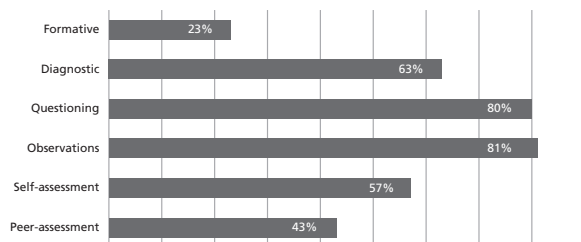


Moreover, most teachers reported using informal assessments in the classroom. For example, approximately 80% of the teachers who completed the survey indicated that they asked students many questions during instruction and observed them while they engaged

in classroom activities and discussions. Asking and monitoring students' questions helps teachers determine how much their students have developed their language skills. The survey also revealed that many teachers incorporated student-centered assessments in their classrooms, such as self-assessment (57%) and peer assessment (43%), which are used sporadically.

We also found that the schools' policies and assessment systems determined the classroom assessment practices of novice teachers. For example, Teacher 1, who works at a language institute, explained, "When I started teaching, I realized it doesn't matter how much assessment training you have. Your assessment system has to be governed by some guidelines or approaches that the institute already has established." Similarly, Teacher 10 (public school teacher) stated: "The school already had a series of criteria for assessment. So, they asked me to assess my students according to those criteria."

Similarly, the teachers commented that they did not have complete autonomy to create their assessments because they needed to align them to large-scale language proficiency national (e.g., SABER 11 test[2]) or international tests (e.g., Cambridge tests). Aligning their classroom assessment to large-scale tests implied having the same format (test structure), using the same item types (types of questions), and assessing similar language skills. This also limited the teachers' autonomy to teach and assess their learners. Teachers 5 and 7 described how the SABER 11 test considerably shaped their assessment practices. The school where Teacher 7 worked had been ranked number one among rural high schools the previous year and number three in

---

2   The SABER 11 test—the national standardized test in Colombia—serves the twofold purpose of exit requirements for high school students and entry requirements for higher education. The test consists of two sections of 131 and 137 questions, respectively, which include the core subjects of critical reading, math and quantitative reasoning, sciences, social studies and citizenship competencies, and English. The English section only assesses reading and language use. It does not include writing or speaking. In addition, it ranks students from A- to B+ levels of proficiency (according to the Common European Framework of Reference).

the SABER 11 test in Arauca.[3] As a result, the board of directors required teachers to focus on preparing students to take this test. Teacher 7 explained how the SABER 11 test impacts her classroom assessment practices:

> Part of the vision of our institution is to show the government how much the students are learning based on the results of the SABER tests, especially in ninth and eleventh grade. So, it started as a guideline for the institution, where most of the assessments are done in this way so that the student becomes familiar, from sixth grade until he reaches eleventh, with the SABER tests.

As a result, Teacher 7 designed SABER-test-item types[4] in an attempt "to continue showing these good results." Although this participant acknowledged that she must always be mindful of the SABER 11 test while teaching and assessing her students, she felt these high-stakes tests constrained her teaching autonomy. Her assessment practices involved familiarizing learners with the test structure and item types.

Likewise, Teacher 5 was required to focus her class activities on the language skills measured on the SABER 11 test. She explained that her school "aimed at improving the test scores of its students on the national test. Then, my goal was to improve their reading comprehension skills." She reported having adapted assessment to this teaching reality. She used the textbook activities to assess her students' reading, grammar, and vocabulary skills and helped them better prepare for the national test.

Similarly, Teacher 2's assessment practices were influenced by international high-stakes tests. Teacher 2 worked at a private school where learners must take the Cambridge Flyers exam to determine their English language proficiency levels. Thus, teachers at this school are compelled to prepare their students to perform well on that exam. One way to accomplish this is by aligning classroom assessments with external tests. As Teacher 2 explained:

> I used the same types of questions as in the Cambridge exam, the exam they take, but I used content from the textbook we used in class. So, I always designed the vocabulary with a Flyers-type structure, one or two grammar points, and a reading text like the ones in the Flyers exam.

Another classroom assessment practice reported by novice teachers was the modification of existing assessments from textbooks and the Internet or the ones provided by their institutions. Typical changes involved shortening the length of the assessments, adding other types of questions, and adapting the assessment to meet the institutions' parameters.

Regarding shortening the assessment length, Teacher 9 explained that he did this by reducing the number of questions to adjust to the limited time allocated to complete the assessment (i.e., one class period). Another change involved adding different questions (e.g., matching or unscrambling). For example, Teacher 9 required his students to support one or two answers with their points of view. Other changes were made because of the pace of the course and the students' progress. For instance, Teacher 6, who taught English to forestry, oil, electronics, and industrial engineering students at a public university, modified the coursebook tests depending on the importance he gave to specific skills. These two novice teachers needed to make critical decisions to adapt existing language assessments so they could be aligned with their instruction. In other words, they did so because they sensed a mismatch between the assessment and their instruction.

A further change involved adapting the assessment to meet the institutions' parameters. Teacher 8 stated that her school's curriculum is aligned with a Cambridge curriculum. Consequently, she needed to adapt what she taught in French to the Cambridge learning objectives.

---

3   One of the 32 departments in Colombia.
4   Mostly related to multiple-choice items.

She said the following about how she modified her assessments:

> I had to adapt the assessments because the school worked with a Cambridge curriculum, so everything was focused on Cambridge content. So, I had to seek some Cambridge tests and then make some changes to fit the courses I was teaching.

The teachers explained that they have to fully comply with the school assessment system to meet the specified assessment standards of the school. In doing so, novice teachers needed to design the assessments according to the guidelines provided by their institutions. Generally speaking, how these three participants changed the original assessments did not compromise the essence of what the predesigned tests were supposed to measure. As the participants expressed, these changes were introduced to keep a more balanced assessment of their students' language skills.

## Challenges Assessing Students in the Classroom

The participants also reported facing many challenges when assessing their students in the classroom. To discuss these challenges, we present them based on the following five categories: (a) conceptualizing assessments, (b) designing assessments, (c) implementing assessments, (d) scoring assessments, and (e) interpreting and using assessments.

### Challenges Conceptualizing Classroom Assessments

Survey results revealed that most novice teachers had challenges determining what, when, and how to assess their students in the classroom (see Figure 3). According to the interviews, novice teachers' main challenges in conceptualizing assessments were determining which skills to assess, which assessment strategy to use (e.g., type of question or task), and aligning assessment and instruction. For example, Teacher 2 discussed

his difficulties in determining what to include in the assessments he designed:

> We have a textbook, and I feel that sometimes there is a lot of content. I try to select a few questions from each content, but I don't know if this is appropriate. So, I need to learn how to select the content for the assessment.

**Figure 3.** Novice Teachers' Challenges Conceptualizing Classroom Assessments (*N* = 75)
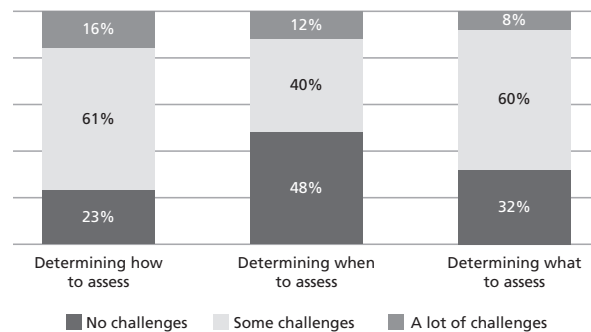


### Challenges Designing Classroom Assessments

Some novice teachers are tasked with creating their assessments or adapting existing assessments to use in their classrooms. Teachers reported having difficulties fulfilling these two tasks (see Figure 4). In terms of creating their classroom assessments, selecting the content of the assessment was very challenging for novice teachers. They expressed that it was daunting to rank the level of importance among the contents taught in class. Teachers also pointed out that their students' low proficiency level was an additional problem when designing the assessments. Teachers struggled to find appropriate texts or adapt these texts and questions to the level of their students. Teacher 8 talked about the difficulties of finding level-appropriate assessment materials for her students:

> My assessment practices were a bit hectic because I had to adapt them depending on my students' proficiency levels and the group's needs. It was initially challenging because I had just graduated, and putting all my knowledge into practice was not easy.

Other challenges include writing multiple-choice questions and determining how many questions to have in the assessment.

**Figure 4.** Novice Teachers' Challenges Designing Classroom Assessments (*N* = 75)
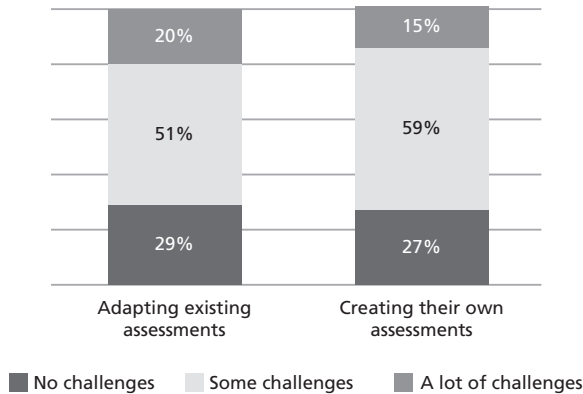


Regarding the challenges novice teachers face when modifying existing assessments, they reported it was problematic to identify what and how to adapt them. For example, Teacher 7 adapted the assessment in the best way possible. However, her coordinator advised her to ask more specific questions, adjust the difficulty level of the assessment to the students' level of proficiency, improve the clarity of instructions, and select the type of audio recordings used more carefully. In addition to the test construction flaws this participant experienced, she was advised to use a better page layout to facilitate the learners' completion of the assessment.

### Challenges Implementing Classroom Assessments

According to the survey, novice teachers also have challenges implementing classroom assessments. We found that 63% of the teachers had some challenges administering their assessments in the classroom. Their main challenge is determining how much time students need to complete an assessment. Other challenges include having little time to assess their students because their classes are relatively short (e.g., less than 50 minutes) and assessing large classrooms (e.g., more than 30 students). Teacher 10 highlighted some of these difficulties: "Today I had three class blocks, but they are classes of 45 students. It is very complicated to develop an optimal assessment process with such large groups; however, that is our reality."

### Challenges Scoring Classroom Assessments

The survey also indicated that novice teachers had difficulty scoring classroom assessments (see Figure 5). Over 70% of participants reported problems establishing or adapting existing scoring criteria (e.g., rubrics) to score their assessments. Likewise, 60% of participants indicated difficulty assigning scores. All the interview participants stated that they faced challenges establishing clear scoring criteria, setting scores for each question on the assessment, applying scoring criteria consistently, and needing more time to score the assessments. Teacher 3 exemplifies these challenges when discussing what areas she needs to improve:

> I would like to learn some strategies to make assessments easier to understand and easier to score. In the beginning, it took me a long time to score all the assessments. You know that there is little time to score them in school.

**Figure 5.** Novice Teachers' Challenges Scoring Classroom Assessments (*N* = 75)
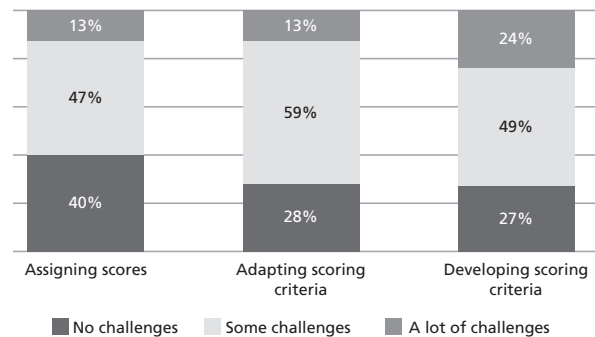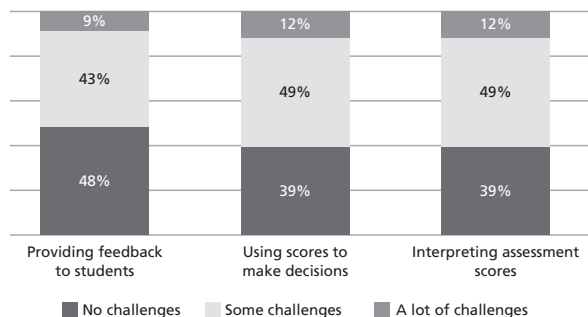
### Challenges Interpreting and Using Classroom Assessments

The survey revealed that 61% of the novice teachers had difficulties making sense of the assessment scores and what the scores mean (see Figure 6). The interview participants reiterated that their main challenge is understanding how the assessment scores relate to learning. Furthermore, 61% of the participants also reported difficulties using scores to make instructional decisions. The interviewed participants indicated that it is challenging to use the assessment scores to understand what they need to do next and to use the assessment scores to improve teaching and learning. Finally, 50% of the participants indicated difficulties providing valuable and relevant feedback to their students (e.g., how they performed and what they needed to improve). The novice teachers said they would like to provide more descriptive feedback to their students (i.e., more than just a grade). Teacher 5 summarized the challenges of interpreting and using classroom assessment scores when reflecting on what areas she needs to improve:

> I really need to work on using formative assessment, providing feedback, and understanding how significant these assessments can be for our students. So, I need to improve how I can use assessment to help my students and focus more on that.

**Figure 6.** Novice Teachers' Challenges Scoring Class-room Assessments (*N* = 75)



| Providing feedback to students | Using scores to make decisions | Interpreting assessment scores |
|---|---|---|
| 9% | 12% | 12% |
| 43% | 49% | 49% |
| 48% | 39% | 39% |

■ No challenges ▫ Some challenges ▪ A lot of challenges

### Discussion

We examined the classroom assessment practices of novice foreign language teachers who had graduated from a foreign language program at a Colombian public university. Regarding what characterizes the assessment practices, we learned that most assessments are teacher-centered and used for summative purposes. For example, all teachers mainly used end-of-unit assessments, final exams, midterms, and quizzes to assess their students in the classroom. However, some novice teachers reported other classroom assessment practices, such as using assessments for formative and diagnostic purposes. They also indicated using informal assessment practices such as asking students questions during instruction or observing their students as they engage in classroom activities. Moreover, a few novice teachers also incorporated active participation of students in their classroom assessment process through self-assessment and peer assessment.

The focus on using summative assessment for accountability purposes in the classroom might indicate that novice teachers need additional training, as López and Bernal (2009) suggested. It is imperative to support novice language teachers using formative assessment to guide teaching and learning (Hill & McNamara, 2011; Leung, 2005; López, 2010; Wolf & López, 2022). Nonetheless, we found evidence that some novice teachers attempted to use other types of assessment in the classroom, albeit sparingly.

Novice teachers' classroom assessment practices also highlight how national and international tests shaped how they assessed their students' learning. These practices also indicate a solid washback effect on instruction. Previous research concerning the washback of large-scale assessments has yielded positive and sometimes unintended negative impacts (e.g., Green, 2007; Menken, 2017; Wall & Horák, 2011). This study did not examine whether the washback of national and international English language proficiency exams was positive or negative. However, novice teachers indicated

they felt forced to narrow the curriculum by focusing only on the language skills measured in the exams and spending instructional time to familiarize students with the format and item types in these large-scale exams.

Many factors influence language teachers' classroom assessment practices, including teaching beliefs, professional development opportunities, experience, and LAL (Cheng et al., 2004; Rea-Dickins, 2001; Yin, 2010). However, for the novice teachers who participated in this study, the main factor influencing their classroom assessment practices was their schools' policies and assessment systems, as Rogers et al. (2020) found. It is essential to emphasize that the primary goal of classroom assessment is to gather evidence of student learning to make informed instructional decisions to enhance the teaching and learning process (Leung, 2005; López, 2010). Therefore, institutions must ensure that their assessment systems and policies are aligned with their curriculum and instruction (Menken, 2017).

Finally, we also investigated novice foreign language teachers' challenges when assessing their students in the classroom. This study found that novice teachers' main challenge is making sense of classroom assessment scores and using this information to inform their decision-making process to support teaching and learning. However, novice teachers also reported needing help designing or adapting sound classroom assessment tasks. These difficulties include determining appropriate assessment approaches (e.g., question types, assessment activity), the assessment content (e.g., language skills), and selecting linguistically appropriate texts for their students. Moreover, novice teachers need some help establishing clear scoring criteria and using the scoring criteria consistently.

Several contextual factors impacted novice teachers' ability to assess their students in the classroom effectively. For example, we found that novice teachers' workload (e.g., number of classes they teach, number of students), time constraints, and available resources made classroom assessment more challenging, similar to what Wolf

and López (2022) reported. Equally problematic is selecting assessment materials (e.g., reading and listening passages) that are appropriate for students to ensure that the assessment is more responsive to students' characteristics and needs (e.g., their foreign language proficiency levels).

## Limitations of the Study

This study had several limitations. For example, the sample only included novice teachers who had all graduated from the same foreign language program. Furthermore, we only gathered reported data, did not observe the novice teachers assessing their students in the classroom, and did not collect sample classroom assessment materials from the novice teachers. Finally, we used a convenience sample for the online survey that only included novice teachers willing to participate. To support the findings from the online survey, we interviewed a subsample of the survey participants using a maximum variation sampling approach to examine a diverse range of cases. Despite the limitations, the study's findings yielded evidence about novice foreign language teachers' classroom assessment practices and challenges when assessing their students. This study is expected to contribute to a better understanding of novice teachers' classroom assessment needs and challenges and inform how to provide better support so they can use assessment to enhance teaching and learning.

## Implications for Practice and Future Research

The overall findings of this study suggest that novice teachers need more knowledge, skills, and support to handle their daily assessment-related tasks. One thing that needs to happen is for novice foreign language teachers to shift the focus of classroom assessment from summative to formative (Earl, 2003; Mohamad Hanefar et al., 2022); that is, for novice teachers to focus more on using classroom assessment to inform their instructional practices and students' learning processes. Nevertheless,

to make this shift, novice foreign language teachers need much support (Karataş & Karaman, 2013). First, educational institutions must create optimal learning environments to foster language learning and facilitate classroom assessment; for example, smaller classrooms, more extended class periods, appropriate resources, and less workload for the teachers to ensure they have ideal conditions to implement and use classroom assessments successfully. Second, educational institutions must review their assessment policies and systems to provide an adequate alignment between assessment and instruction in terms of content (language skills) and the language proficiency level of the students.

Moreover, it is imperative to continue providing professional development opportunities for novice teachers to increase their knowledge and skills (Heritage & Wylie, 2020). Professional development is crucial in the early stages of the teaching career (Rogers et al., 2020). Some topics of particular interest include adapting texts, designing tasks, scoring, interpreting, and using assessments. The foreign language program could provide these training opportunities to support novice teachers at the beginning of their teaching journey or by the institutions based on their unique needs and challenges. However, novice teachers must be proactive and look for other opportunities to continue enhancing their knowledge and skills by accessing information in printed materials or the Internet and attending conferences, webinars, or other sharing-knowledge sessions. In addition to all these training opportunities, teachers need to continue self-monitoring and self-regulating their classroom assessment practices.

We also suggest creating communities of practice where novice teachers can interact and collaborate with their colleagues at school or elsewhere (e.g., other graduates from the foreign language program). In communities of practice, participants unite in a continued effort to enhance their knowledge and skills (Wenger, 1998). These communities of practice can help novice teachers discuss common interests regarding classroom assessment and share their assessment concerns or challenges to deepen their skills, knowledge, and expertise in classroom assessment. We also encourage language coordinators or experienced teachers to mentor novice teachers. This mentorship includes working closely with novice teachers, reviewing their assessment instruments, and providing feedback to improve them.

Finally, this study also has some implications for future research. Large-scale exams strongly influenced novice teachers' instructional and assessment practices. Thus, we suggest future research to examine the impact of these large-scale exams on teaching and learning. Moreover, we found that novice teachers used different types of assessment (e.g., formal, informal, and student-centered). Future research should also investigate how novice teachers use all these assessments, their benefits, and their challenges.

## References

Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, *106*(3), 429–458. https://doi.org/10.1111/j.1467-9620.2004.00346.x

Butler, Y. G. (2022). Exploring the idea of "assessment with learners" in discussing Language assessment literacy. *Language Teaching Research Quarterly*, *29*, 92–101. https://doi.org/10.32038/ltrq.2022.29.06

Caicedo Pereira, M. J., Lozano Bermúdez, J. A., & Vanegas Medina, L. A. (2018). Improving L2 oral accuracy and grammatical range through self-assessment of video speech drafts. *Profile: Issues in Teachers' Professional Development*, *20*(2), 127–142. https://doi.org/10.15446/profile.v20n2.61724

Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, *21*(3), 360–389. https://doi.org/10.1191/0265532204lt288oa

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating qualitative and quantitative research* (1st ed.). Pearson Education.

Duque-Aguilar, J. F. (2021). Teachers' assessment approaches regarding EFL students' speaking skills. *Profile: Issues in Teachers' Professional Development*, *23*(1), 161–177. https://doi.org/10.15446/profile.v23n1.85964

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning* (1st ed.). Corwin Press.

Farrell, T. S. C. (2012). Novice-service language teacher development: Bridging the gap between preservice and in-service education and development. *TESOL Quarterly*, *46*(3), 435–449. https://doi.org/10.1002/tesq.36

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. (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., & 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

Gómez Sará, M. M. (2016). The influence of peer assessment and the use of corpus for the development of speaking skills in in-service teachers. *HOW Journal*, *23*(1), 103–128. https://doi.org/10.19183/how.23.1.142

Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education: Principles, Policy and Practice*, *14*(1), 75–97. https://doi.org/10.1080/09695940701272880

Hayes, D., & Chang, K. (2017). South Korean novice English language teachers' experience of induction into teaching. *English Teaching*, *72*(1), 49–71. https://doi.org/10.15858/engtea.72.1.201703.49

Heritage, M., & Wylie, E. C. (2020). *Formative assessment in the disciplines: Framing a continuum of professional learning*. Harvard Education Press.

Hernández-Ocampo, S. P. (2022). Colombian scholars' discussion about language assessment: A review of five journals. *Profile: Issues in Teachers' Professional Development*, *24*(2), 231–245. https://doi.org/10.15446/profile.v24n2.94468

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 Linguistic Journal*, *17*(2), 302–312. https://doi.org/10.14483/udistrital.jour.calj.2015.2.a09

Herrera Mosquera, L., & Zambrano Castillo, L. C. (2019). Assessment of English learning in a language teacher education program. *GIST - Education and Learning Research Journal*, (19), 193–214. https://doi.org/10.26817/16925777.709

Hill, K., & McNamara, T. (2011). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, *29*(3), 395–420. https://doi.org/10.1177/0265532211428317

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

Karataş, P., & Karaman, C. (2013). Challenges faced by novice language teachers: Support, identity, and pedagogy in the initial years of teaching. *The International Journal of Research in Teacher Education*, *4*(3), 10–23. https://open.metu.edu.tr/handle/11511/75962

Kim, K.-A., & Roth, G. L. (2011). Novice teachers and their acquisition of work-related information. *Current Issues in Education*, *14*(1).

Leung, C. (2005). Classroom teacher assessment of second language development: Construct as practice. In E. Hinkel (Ed.), *Handbook of research in second language learning and teaching* (pp. 869–888). Lawrence Erlbaum Associates.

Lopera Medina, S. A. (2015). Effects of classroom assessment practices in a foreign language reading course. *GIST - Education and Learning Research Journal*, (11), 9–25. https://doi.org/10.26817/16925777.377

López, A. A. (2010). La evaluación formativa en la enseñanza y aprendizaje del inglés. *Voces y Silencios: Revista Latinoamericana de Educación*, *1*(2), 111–124. https://doi.org/10.18175/vys1.2.2010.01

López, A. 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.

López, A. A., & Janssen, G. (2010). Validation study of Colombia's ECAES English exam. *Revista Lenguaje*, *38*(2), 423–448. https://doi.org/10.25100/lenguaje.v38i2.4916

Menken, K. (2017). High-stakes tests as de facto language education policies. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment* (3rd ed., pp. 385–396). Springer. https://doi.org/10.1007/978-3-319-02261-1_25

Mohamad Hanefar, S. B., Anny, N. Z., & Rahman, S. (2022). Enhancing teaching and learning in higher education through formative assessment: Teachers' perceptions. *International Journal of Assessment Tools in Education*, *9*(1), 61–79. https://doi.org/10.21449/ijate.946517

Mohammadkhah, E., Kiany, G. R., Tajeddin, Z., & ShayesteFar, P. (2022). EFL teachers' assessment literacy: A contextualized measure of assessment theories and skills. *Language Teaching Research Quarterly*, *29*, 102–119. https://doi.org/10.32038/ltrq.2022.29.07

Muñoz, A. P., & Álvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing*, *27*(1), 33–49. https://doi.org/10.1177/0265532209347148

Ramirez, A. (2020). Challenges in the design and implementation of an English placement test for a Colombian public university. *GIST - Education and Learning Research Journal*, (21), 191–208. https://doi.org/10.26817/16925777.825

Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, *18*(4), 429–462. https://doi.org/10.1177/026553220101800407

Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, *21*(3), 249–258. https://doi.org/10.1191/0265532204lt283ed

Resolución 18583, Diario Oficial No. 50.357 (2017). https://bit.ly/3psrVOw

Rogers, A. P., Reagan, E. M., & Ward, C. (2020). Preservice teacher performance assessment and novice teacher assessment literacy. *Teaching Education*, *33*(2), 175–193. https://doi.org/10.1080/10476210.2020.1840544

Saldaña, J. (2009). *The coding manual for qualitative researchers* (1st ed.). Sage.

Strauss, A., & Corbin, L. (1990). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Sage.

Tajeddin, Z., Khatib, M., & Mahdavi, M. (2022). Critical language assessment literacy of EFL teachers: Scale construction and validation. *Language Testing*, *39*(4), 649–678. https://doi.org/10.1177/02655322211057040

Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social & behavioral research* (1st ed.). Sage.

Wall, D., & Horák, T. (2011). *The impact of changes in the TOEFL® exam on teaching in a sample of countries in Europe: Phase 3, The role of the coursebook. Phase 4, Describing change* (Research report No. 11-41, TOEFL iBT-17). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02277.x

Wang, X. (2017). A Chinese EFL teacher's classroom assessment practices. *Language Assessment Quarterly*, *14*(4), 312–327. https://doi.org/10.1080/15434303.2017.1393819

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press. https://doi.org/10.1017/CBO9780511803932

Wolf, M. K., & López, A. A. (2022). Developing a technology-based classroom assessment of academic reading skills for English language learners and teachers: Validity evidence for formative use. *Languages*, *7*(2), 71. https://doi.org/10.3390/languages7020071

Yin, M. (2010). Understanding classroom language assessment through teacher thinking research. *Language Assessment Quarterly*, *7*(2), 175–194. https://doi.org/10.1080/15434300903447736

## About the Authors

**Gabriel Cote Parra** is a BA in Foreign Languages professor at the Universidad de Pamplona, Colombia. He holds a PhD in Education (University of Nebraska, Lincoln, USA). His research focuses on learning and teaching foreign languages, which is one of the lines of investigation of the Research Group in Foreign Languages (GRILEX).

**Alexis A. López** holds a PhD in Education (University of Illinois at Urbana-Champaign, USA) and is a senior research scientist at Educational Testing Service. His research focuses on developing learning-oriented, formative, and multilingual assessments for English learners and understanding the classroom assessment practices of ESL/EFL teachers.

# Appendix A: The Online Survey

**Table A1.** Classroom Assessment Practices

| Assessment method | Never | Occasionally | Frequently |
|---|---|---|---|
| Summative assessment | | | |
| Formative assessment | | | |
| Diagnostic assessment | | | |
| Questioning | | | |
| Observations | | | |
| Self-assessment | | | |
| Peer-assessment | | | |
| Other | | | |

**Table A2.** Challenges Implementing Classroom Assessment

| Challenge | No challenges | Some challenges | A lot of challenges |
|---|---|---|---|
| Determining how to assess | | | |
| Determining what to assess | | | |
| Determining when to assess | | | |
| Creating their assessments | | | |
| Adapting existing assessments | | | |
| Administering the assessments | | | |
| Setting scoring criteria | | | |
| Adapting existing scoring criteria | | | |
| Assigning scores | | | |
| Interpreting assessment scores | | | |
| Using assessment information | | | |
| Providing feedback to students | | | |
| Other | | | |

# Appendix B: Interview Protocol

1. What role does classroom assessment play in instruction or your teaching practices?

2. How do you assess your students in the classroom? What types of assessments or instruments do you use?

3. Who designs the assessments you use in your classroom?

4. How often do you assess your students?

5. How do you score your assessments? How do you set your scoring criteria?

6. How do you interpret the assessment scores?

7. What kinds of decisions do you make based on the evaluations you make?

8. How do you provide feedback to your students?

9. What kind of difficulties do you have or have you had when assessing your students?

10. Why do you think you have had these difficulties in adequately assessing your students? To what factors do you attribute these difficulties?