

Perceptions and Beliefs of EFL Students Regarding Classroom Assessment at a Colombian University

Percepciones
y creencias de
estudiantes del
programa LLE
sobre los métodos
de evaluación en
una universidad
colombiana

Percepções e
crenças dos alunos
de ILE sobre a
avaliação em
sala de aula em
uma universidade
colombiana

Leonardo Herrera Mosquera* <https://orcid.org/0000-0002-4054-7209>

Ósmar David González Meléndez** <https://orcid.org/0000-0002-6903-3914>



Para citar este artículo

Herrera Mosquera, L. y González Meléndez, O. (2021). Perceptions and Beliefs of EFL Students Regarding Classroom Assessment at a Colombian University. *Folios*, (54). <https://doi.org/10.17227/folios.54-11498>

Artículo recibido
22•03•2020

Artículo aprobado
28•01•2021

* Máster en TESOL. Docente de Planta Universidad Surcolombiana, Neiva, Colombia.

Correo: leonardo.herrera@usco.edu.co

** Licenciado en Lenguas Extranjeras. Docente de tiempo completo Aspaen La Fragua, Neiva, Colombia.

Correo: u20122114120@usco.edu.co

Abstract

The higher education institution, where the present study is conducted, proposes in its school handbook that a formative and democratic learning assessment approach should prevail in all academic activities. In the English as a Foreign Language Teacher Education Program at this university, eighty percent of students' grades come from summative testing, and the teachers are the ones who regularly make the decisions regarding any assessment criteria. That is, the formative and democratic components are not present in the process. To delve deeper into these aspects, we have inquired students on their perceptions and beliefs concerning classroom assessment and suggestions to improve the process. A small-scale qualitative research study was conducted with 49 participants. The data collected through questionnaires, focus-group interviews, and narratives allowed us to organize the findings according to three emerging categories: students' perceptions of assessment as Assessment of Learning (AOL), Assessment for Learning (AFL), and suggestions for improving the overall language evaluation process. The different concepts provided by the participants go from believing that classroom evaluation constitutes a procedure to elicit rote reproduction to that of perceiving it as an opportunity to enhance students' formation. Actions for improving assessment go from particular procedural aspects to those related to a more general approach.

Keywords

assessment for learning (AFL); assessment of learning (AOL); classroom assessment; democratic assessment; formative assessment

Resumen

Las Instituciones de Educación Superior, objeto del presente artículo de investigación, propone que en todas las actividades académicas prevalece un enfoque de evaluación formativo y democrático, tal como se estipula en el manual de convivencia. En el caso del Programa de Licenciatura en Lengua Extranjera de esta universidad, el ochenta por ciento de las notas de los estudiantes proviene de evaluaciones sumativas y los docentes son quienes normalmente toman las decisiones relacionadas con procesos evaluativos; es decir, los componentes formativos y democráticos no parecen estar presentes en el proceso. Para revisar un poco más en profundidad estos aspectos, decidimos preguntarles a los estudiantes sobre sus percepciones y creencias acerca de la evaluación implementada, donde nos ofrecieron sugerencias que permitan cualificar el proceso. De esta forma, se desarrolló un estudio cualitativo a pequeña escala con 49 estudiantes, donde los datos fueron recopilados a través de cuestionarios, entrevistas focales y narrativas, herramientas que permitieron organizar los resultados en tres categorías emergentes: percepciones de los estudiantes sobre evaluación del aprendizaje (AOL, por sus siglas en inglés), sobre evaluación para el aprendizaje (AFL, por sus siglas en inglés) y sugerencias para mejorar el proceso de evaluación. Los diferentes conceptos proporcionados por los participantes van desde la concepción de evaluación como un procedimiento que suscita la reproducción memorística hasta la percepción de este proceso como una oportunidad para fortalecer la formación de los estudiantes. Finalmente, las acciones para mejorar el proceso evaluativo van desde aspectos procedimentales particulares hasta aquellos relacionados con un enfoque de evaluación más general.

Palabras clave

evaluación para el aprendizaje; evaluación del aprendizaje; evaluación en el aula; evaluación democrática; evaluación formativa

Resumo

A instituição de ensino superior, na qual o presente estudo é desenvolvido, propõe em seu manual escolar que a abordagem formativa e democrática de avaliação da aprendizagem deve prevalecer em todas as atividades acadêmicas. No Programa de Formação de Professores de Inglês como Língua Estrangeira desta universidade, oitenta por cento das notas dos alunos vêm de testes somativos e os professores são os que regularmente tomam as decisões sobre os critérios de avaliação. Ou seja, os componentes formativo e democrático não estão aparentemente presentes no processo. Para aprofundar esses aspectos, questionamos os alunos sobre suas percepções e crenças em relação à avaliação em sala de aula e sugestões para

melhorar o processo. Uma pesquisa qualitativa de pequena escala foi conduzida com 49 participantes. Os dados coletados por meio de questionários, entrevistas de grupos focais e narrativas nos permitiram organizar os resultados segundo três categorias emergentes: as percepções dos alunos sobre a avaliação como Avaliação da aprendizagem, Avaliação para a aprendizagem e sugestões para melhorar processo geral de avaliação da linguagem. Os diferentes conceitos fornecidos pelos participantes vão desde acreditar que a avaliação em sala de aula constitui um procedimento para eliciar a reprodução mecânica até percebê-la como uma oportunidade para potencializar a formação dos alunos. As ações para melhorar a avaliação vão desde aspectos processuais particulares até aqueles relacionados a uma abordagem mais geral.

Palavras chave

avaliação para a aprendizagem; avaliação da aprendizagem; avaliação em sala de aula; avaliação democrática; avaliação formativa

Introduction

As both teachers and students, our experiences have enabled us to witness a traditional educational paradigm in which assessment is regarded as a separate component from teaching. Assessment of learning has been traditionally related to a procedure at the end of a cycle (unit, lesson, term, etc.) and which provides teachers with a passing or failing concept, usually in the form of a numerical value. As Jang (2014) claims, “this disconnect between assessment and instruction happens in some contexts where educational reforms are driven by mass testing and where testing becomes the end in itself” (p. 5).

Research in educational assessment calls for a more comprehensive scope that values aspects other than shreds of evidence on formal written tests. Teachers are being encouraged to consider different alternatives that contribute to collecting information regarding students’ knowledge, abilities, and performance. These alternatives, along with more ‘traditional’ tests typically used in classroom evaluation, will provide the teacher with more elements to make decisions in the teaching–learning process. Thus, *assessment* and *testing* have been regarded as distinct constructs that aim at the same purposes: collecting evidence of students’ knowledge and skills and making decisions based on this information (Hughes, 2003; Suskie, 2018).

Nowadays, different authors support the belief that just a test cannot determine the learning; rather, they point out that there are multiple sources to

obtain evidence of students’ learning (Brown, 2004; Popham, 2008; Shohamy, 1998). These sources can be oral presentations, essays, role-plays, portfolios, and the like. Hence, assessment is considered the ongoing process that helps to improve teaching practice and students’ language learning through the systemic collection and review of information that teachers interpret to attain educational purposes (Iseni, 2011; Jang, 2014). On the other hand, testing focuses on measuring students’ abilities and knowledge in a given domain (Brown, 2004; Hancock, 2006). This domain can be a specific language skill or a particular linguistic content. Testing has been traditionally related to the process of assessing knowledge during a definite learning period, aligned to the course curriculum, and eliciting a score or grade.

Teachers sometimes ignore the impact that testing may have on students’ motivation to learn. Therefore, it is necessary to implement assessment procedures that enhance students’ motivation for learning even if their result on a test was not as expected. As Kandbilder (2009, cited in Iseni, 2011) claims, “Simply by changing the assessment of your subject you can affect the way students engage with the subject content” (p. 61). It means that we, as teachers, have a great responsibility when deciding who passes or fails a test, not only based on a final grade but also through a formative accompaniment that provides meaningful information of their learning process for the learners and us.

The elements of summative and formative assessment processes insofar described were fundamental for developing the present study. Students from the EFL Teacher Education Program were inquired about their perceptions, beliefs, and suggestions regarding their learning assessment processes conducted in their courses in general and English as a Foreign Language (EFL) courses in particular. Thus, the research stated these questions:

- What are the students' perceptions and beliefs concerning their efl learning assessment process?
- What suggestions do they make to improve the efl learning assessment process?

The perceptions and beliefs were analyzed according to constructs offered within the literature of assessment and testing.

Theoretical Framework

Whenever we talk about assessing the learning of a language, whether as a second or foreign language, the terms *test* or *testing* seem to come to our minds immediately. These terms come along with other words such as multiple-choice, fill-in-the-blank, translation, matching, dictation, and the like. This top-of-mind reaction is owing perhaps to the long tradition of large-scale testing, objective testing, or selected-response type of tests, which still prevail in our language classrooms today.

Nevertheless, with the advent of the massive use of educational technologies and the emergence of new pedagogical paradigms, the field of learning assessment has moved towards other approaches that somehow take a certain distance from the paradigm of psychometric testing. Hence, we talk today about *assessment* as a more general construct for all that occurs in collecting evidence of students' learning. And with assessment, as a counterpart of testing, have emerged perspectives such as formative, alternative, dynamic, embedded, and classroom-based assessments, some of which can even be used interchangeably. Likewise, the school has emphasized on evaluative procedures such

as feedback, self and peer evaluations, portfolios, journals, rubrics, etc.

Thus, in the present theoretical framework, we will first describe conceptual differences between testing and assessment as two major constructs in the history of language evaluation. Secondly, we will delve into the more modern paradigms of assessment for learning versus assessment of learning; we will then review some important aspects of classroom-based assessment, milieu on which the present study participants will be providing insights, in terms of perceptions, beliefs, and suggestions. Finally, we will present some research studies related to learning assessment conducted in Colombia and other countries.

What is testing?

The paradigm of testing has traditionally dominated the certification of students' knowledge and abilities. When we talk about testing, we may be referring more to measurement, periodic tests, multiple-choice tests, and paper-and-pencil type tests. These qualities of measurement have commanded the tradition of learning assessment and are still prevalent in today's classrooms. Standardized and norm-referenced tests have been used to meet the demands, not necessarily of teachers and students, but actors external to the classroom.

Tests serve different purposes beyond regular classroom testing. First, tests can be used to measure students' general language competence, known as proficiency tests. Secondly, they can be used to assign students to a determined course or language level at a school program, known as placement tests. Finally, if teachers are concerned with the efficacy of their courses, lessons, or units, they can conduct achievement tests to report how much students have learned. Generally, these sorts of tests are administered at the end of a unit or course (Brown, 2004; Harmer, 2007). Other test types commonly described in the literature, which may resemble the ones just mentioned, are diagnostic tests, progress tests, and admission tests (Carr, 2011).

Testing does not imply a random selection of tasks or items for the sake of testing. It means great preparation from both the test developers (whether large-scale or classroom testing) and the test takers. In the case of classroom testing, it is not something that the teacher does without connecting to the criteria and objectives of the unit or course taught. Teachers need to think of purposes, instruments, feedback, and other relevant factors inherent to testing. As Fulcher and Davidson (2007) claim, the making of a test implies “a particular skill and training in test design” (p. 28).

In response to the tradition of testing, the term ‘assessment’ has been suggested as a more comprehensive construct for a type of evaluation that targets performance rather than just receptive skills or grammar and vocabulary knowledge.

What is Assessment?

According to Popham (2008), the term assessment involves “testing students in the midst of an ongoing instructional sequence and then using the test results to improve instruction” (p. 3). For Barkley and Howell (2016), assessment corresponds to “appraisal of the knowledge, skills, attitudes, and beliefs that students have acquired, most often as the result of learning in their courses” (p. 24). Brookhart and Nitko (2019) go beyond the realm of the teacher and the student and state that assessment is “a process for obtaining information for making decisions about students; curricula, programs, and schools; and educational policies” (p. 2).

Thus, assessment has been regarded as an evaluation process that transcends the discrete-point approach and targets other impacts and aspects of students’ development. It may give us the feeling that the term ‘assessment’ broadens the perspective of measurement that has been traditionally attributed to ‘testing.’

Within the category of assessment, different conceptual dichotomies have emerged, such as quantitative versus qualitative assessment, formal versus informal, traditional versus alternative, and

the widely discussed summative versus formative assessment. Some of these distinctions have been suggested due to the nature and characteristics of their evaluative procedures. For instance, most authors refer to summative assessment as the process of measuring students’ learning, usually at the end of a period of teaching; while formative assessment has been regarded as a permanent process intended to improve both students’ learning and teachers’ instruction (Brown, 2004; Carr, 2011; Moss & Brookhart, 2009). Summative assessment usually corresponds to formal assessment, whereas informal assessments can play an important function within the formative assessment. Likewise, summative can be associated with quantitative assessment while formative with more qualitative assessment. The alternative assessment constitutes the promise of choice instead of more traditional assessments such as selected-response type tests.

Recently, a new perspective toward assessment has emerged. According to Shohamy (1998), assessment serves diverse purposes, and educators have developed various strategies to take advantage of every feature for evaluation. This ‘multiplism’ perspective is intended to orientate the selection of different options on the different assessment phases. Multiplism offers various possibilities when a teacher assesses learners, and these options vary depending on the purpose of the assessment. Shohamy, as one of the leading advocates of multiplism, states that “Assessment is shaped by its specific context, its purposes, the type of knowledge it addresses, the procedure it selects, by the different criteria for determining success, by different interpretation and different ways of reporting results” (p. 258).

The evolution of various teaching methodologies and approaches throughout the second half of the twentieth century has generated the distinctions above between testing and assessment and between the different types of tests and assessments.. We will briefly describe some of those stages in the following section.

Assessment of Learning (AOL) and Assessment for Learning (AFL)

The first uses of the AFL concept appear to occur in the early 90s, but according to William (2010), the term AFL in contraposition to AOL seemed to have been coined by Gipps and Stobart a few years later.

Even though AOL has been associated with summative evaluations and assessment for learning (AFL) with formative evaluations, some authors find discrepancies in these connections. William (2010) cites Black et ál. (2004) to differentiate between AFL and formative assessment:

Assessment for learning is any assessment for which the priority in its design and practice is to serve the purpose of promoting students' learning. It thus differs from assessment designed primarily to serve the purposes of accountability, or of ranking, or of certifying competence. An assessment activity can help learning if it provides information that teachers and their students can use as feedback in assessing themselves and one another and in modifying the teaching and learning activities in which they are engaged. Such assessment becomes "formative assessment" when the evidence is actually used to adapt the teaching work to meet learning needs. (p. 10)

However, the distinctions between formative and summative assessments keep a remarkable resemblance with those between AFL and AOL, mainly in terms of their function; understanding the distinction between AFL and AOL "is pivotal to realizing gains in student achievement" (Chappuis and Stiggins, 2020, p. 26).

According to Chappuis and Stiggins (2020), these constructs bear clear distinctions regarding the reasons for assessing, the audience, the place in time, the primary users, and their typical uses. AFL focuses on student growth and improvement, whereas AOL seeks to document achievement for purposes of reporting or accountability. AFL takes place during learning so that both the teacher can adjust their instructional practices if needed and students can modify or continue with the learning strategies they have been using. Instead, AOL takes

place after learning has occurred. In some contexts, especially during final exams, there is not even time to provide students with feedback regarding their performance on particular tasks or their general learning progression. In AOL, the audience and the primary users of the information emerging from the assessment consist more of external agents instead of the students themselves, as in AFL. Likewise, grades, certification, graduation, and promotion issues are primary concerns of AOL (p. 26).

In a nutshell, teachers should plan educational assessment—language assessment in this case—and use the information collected from it as a resource to promote further learning and development. In other words, all assessments conducted in the classroom should be approached as assessment for learning, even if the evaluative procedures and instruments seem to be more aligned to summative assessment. Both AFL and AOL have a place in Classroom-based Assessment (CBA), another important construct in language learning assessment.

Classroom-based Assessment (CBA)

CBA constitutes an assessment paradigm that, in alignment with other constructs such as formative assessment or assessment for learning, has emerged from socio-cultural learning theories, in somehow contraposition to the large-scale psychometric testing. CBA responds to the need to look at assessments that are more internal to the classroom instead of traditional standardized language exams that are distant from its realities.

Assessment theorists have used different labels to describe the assessment that takes place in the classroom and intends to support learning. They have used classroom assessment, teacher assessment, performance, alternative assessment, dynamic assessment, among others. Nevertheless, as Turner (2017) points out, "several of these terms when related to classroom activity are used simultaneously with and considered to be an extension of the more general term formative evaluation" (p. 67).

In the process of resorting to multiple sources to evidence students' learning progression, it some-

times becomes difficult to tell when an action or activity taking place in the classroom corresponds to a simple instructional activity or an evaluative procedure. Many teaching tasks can easily turn into assessment tasks depending on how teachers assume them. Sometimes a grade or score may indicate if the task is indeed an assessment task. This circumstance allows for two types of assessment modes, known as formal and informal assessment (Brown, 2004) or implicit and explicit assessment (Bachman and Dambock, 2017).

According to Brown (2004), informal assessments take several forms and do not necessarily imply recording the result of a performance. On the other hand, formal assessment constitutes more systematic and planned evaluative exercises or procedures (p. 6). For Bachman and Dambock (2017), one of the differences between the implicit and explicit modes lies in that students are unaware that an assessment is taking place on the former. In contrast, on the latter, both the teacher and the students know that the activity corresponds to an assessment (p. 16).

Whether formal or informal, explicit or implicit mode, one of the main considerations to be made in CBA is that teachers use the information they collect about students' learning, through any of the modes, to inform both the learning and the teaching process. Both the decisions and consequences deriving from the assessment must be formative and conducive to improvements for teachers and students.

Studies related to perceptions and beliefs about assessment

There is a vast quantity of research studies devoted to assessment in education. Most research studies focus on showing the effectiveness of formative assessment and the role teachers and students play in this process. A research study conducted by Falchivoke (2004, cited in Mussawy, 2009) showed that significant participation in peer assessment yielded better results in the students' learning process. In the same line of peer assessment, Iseni (2011) conducted a study about noting and reviewing students' errors.

She concluded that peer assessment "can help them become more analytical, and in this way, they learn more and gain an understanding of their progress, they learn how to be critical and analytical, and they engage themselves in metacognition" (p. 70). However, she emphasized the importance of making the assessment process continuous to obtain productive results for learners.

Mussawy's study also determined faculty members' and pre-service teachers' perceptions of classroom assessment. The results showed that most faculty members do not emphasize the formative part of the evaluation but rather focus their attention on the performance measured at the end of the course. In addition, Mussawy (2009) concluded that "students and teachers had a recognition of various forms and purposes of classroom assessment. Besides, both students and the faculty members viewed assessment as activities for educational purposes, not just assigning score or grade..." (p. 59–60).

Herrera and Zambrano (2019) inquired 160 students and five English teachers from a language teacher education Program at a Colombian university about their views on learning assessment, in general, and the assessment practices administered in the different language courses, in particular. They found that even though both teachers and students favor the implementation of formative evaluative practices and that institutional documents state that a formative assessment approach should be followed, there are shreds of evidence of summative evaluative procedures in the Program. The study concluded that the implementation of professional development opportunities on educational assessment, in addition to the incorporation of alternative assessment procedures, could contribute positively with the consolidation of "principles, purposes, functions, forms, means, techniques and evaluation instruments, on which teachers and students, as the main actors of the evaluation process, must coincide, thus avoiding surprises, disagreements, and injustices" (p. 210).

Muñoz, Palacio, and Escobar (2012) found in their study, led at a Colombian language center, a

discrepancy between how teachers perceive assessment and the evaluative practices in the classrooms. Teachers regard assessment as a relevant contributor to the teaching and learning process, even more when assessment is assumed from a formative perspective. Nonetheless, summative assessment orientations are still evident in the teachers' comments. The authors recommend professional development opportunities so that teachers can reflect on both their assessment beliefs and practices.

Palacio et ál. (2016) conducted a study to improve testing in a language program at a public school in Colombia. They found that traditional tests harm their program since the format usually followed in tests (fill-in-the-blanks, multiple-choice, and matching) did not show a positive washback on students' learning. Likewise, the nature of these testing instruments does not allow teachers to assess students in a holistic way, which transcends the mere testing of students' subject knowledge. Also, the authors suggested revising the effectiveness of test tasks and items to improve the quality of course testing continuously.

López and Bernal (2009) found that trained teachers tend to view assessment as something integral during the instruction and as a tool that enhances learning. Teachers with no training (pre-service or in-service teachers) expressed negative opinions about assessment; they think that assessment serves only to give an evaluative grade and not a strategy for improving learning. This leaves much to be desired, as future teachers may continue with the same testing instruments without considering new techniques that may help to improve both their teaching practice and their students' learning.

Some studies that have focused on classroom assessment and testing (Areiza, 2013; López and Bernal, 2009; Palacio et al., 2016) have concluded that teachers rely heavily on summative assessment without considering other alternatives to assess students. Moreover, public universities do not offer courses or professional development seminars through their teaching programs to tackle assessment issues.

Thus, the present study aimed at identifying students' perceptions and beliefs regarding their language learning assessment process to contribute to the body of knowledge in the area, both in the Colombia context and abroad.

Methodology

Research Design

Even though this small-scale research study used numerical elements in data analysis, it is predominantly qualitative and descriptive. The study aimed to describe an educational phenomenon (classroom assessment) that has not been quite researched in the targeted language teacher education program. This study sought to identify students' perceptions and beliefs regarding the assessment of their EFL learning process. The study's authors collected the data within the academic setting through semi-structured questionnaires, focus-group interviews, and narratives. All of these instruments were designed by them and revised by peer researchers.

The questionnaires were designed with open-ending questions so that students could provide more descriptive responses. The focus-group interviews and narratives complemented the information collected through the questionnaires, allowing us to respond to the research questions better.

Since we talked about perceptions, beliefs, and suggestions, it was imperative to implement a descriptive qualitative approach. In this regard, Creswell (2012) defines qualitative research as:

A means for exploring and understanding the meaning individuals or groups ascribe to a social or human problem. The process of research involves emerging questions and procedures, data typically collected in the participant's setting, data analysis inductively building from particulars to general themes, and the researcher making interpretations of the meaning of the data. (p. 22)

About the quantitative elements, some of the information will be presented in terms of frequencies and percentages. Although the study

is predominantly descriptive and qualitative, the combination of quantitative and qualitative elements used in the study offered an invaluable contribution to its analysis.

Participants

All of the students from an upper-intermediate English course for both academic terms 2016A and 2016B participated in this study. This is a convenience sample because the professor assigned to these courses is the co-author of this study. The number of participants was 49, and their ages ranged between 18 and 24 years old. Twenty-nine participants were female students, and twenty were male students. Students were given the option to withdraw from participating in any research stage, especially in completing the questionnaires. However, all of the students decided to take part in the study.

As previously stated, the participants have between an intermediate and advanced English proficiency level, according to the Program's placement criteria and courses' denomination. That is to say,

they have already undergone different assessments, whether in their English courses or non-language courses, and, therefore, could provide us with reliable opinions regarding the learning assessment processes.

Data Collection

All participants completed a questionnaire that inquired on their experiences with assessments, both in their English courses and the other non-language courses. Then researchers selected groups of five participants for the focus-group interviews. Finally, about ten students volunteered to complete a narrative form with personal concepts and suggestions about the learning assessment process. The instruments were administered in the classroom, the computer lab, and in a teacher's office. They were also delivered at different moments of their educational process; that is, participants completed the forms previous to, while, and after assessments (see Table 1). As previously stated, the authors designed and adjusted these instruments following the revision made by peer researchers.

Table 1

Research Instruments

Instruments	Area of Inquiry
Questionnaire	Definitions of learning assessment; feelings and attitudes towards assessment; test-taking strategies; preferences regarding assessment and testing
Focus-group Interviews	Perceptions of and experiences with learning assessment
Narratives	Feelings and suggestions regarding classroom assessment experiences

Following is a more detailed description of the instruments and procedures:

Questionnaire. This instrument consisted of 20 questions, 18 being selected-response type questions, and two personal-response questions. The personal or open-ended questions were intended to provide more descriptive information that closed-ended questions might not have inquired. The questions asked on their definitions of learning

assessment, their feelings and attitudes towards assessment, strategies they use to prepare for tests, types of evaluation instruments or modes they prefer, and the like. The questionnaire was online and did not take more than 20 minutes for students to complete. Almost all of the students completed the questionnaire in the computer lab during English class.

Focus-group Interviews. Two groups of students per academic term were interviewed. Each interview took about 30 minutes with equal opportunities for all participants (between five and six) to talk and respond to prompts. Between four and five prompts were presented to students inquiring about the topics mentioned in the previous instruments. With this instrument, students could expand on concepts and opinions they did not provide in the questionnaire.

Narratives. About ten students per academic term participated in this research stage. This instrument consisted of three guiding questions that inquired on definitions of learning assessment and suggestions regarding the evaluation process. Participants took about 20 minutes to complete this form.

Ethical Considerations

Their English professor informed students in this study about the type and purpose of the study. It was clarified to students that their names would never be revealed and that this study would affect their educational process in no way. All data, even contained in the online questionnaire, was kept confidential. The online questionnaire did not request their names.

Data Analysis

The data contained in each of the instruments will be presented mostly through qualitative descriptions

and some percentages. Firstly, we organized the qualitative data through charts with some initial categories. Secondly, we used Atlas.ti software to review the raw data and the data already organized in charts and either modify or add a new category. This program has the advantage of analyzing virtual documents simultaneously and show all the final data in a semantic layout easy to read and to be interpreted.

We analyzed questionnaires through pie charts with qualitative descriptions that gave us better insights regarding students' responses. The open-ended questions of the questionnaire (19 and 20) were separated in order to analyze them carefully and search for relevant information. In doing so, the students' responses were analyzed through code framing to identify categories and themes related to the research purpose (Bodgan & Biklen, 1998, cited in Muñoz et al., 2012). The categories obtained from the qualitative instruments were not pre-established, but they emerged throughout the analysis.

The focus group interviews were audio-recorded and transcribed to analyze participants' responses and give them an appropriate category. About students' narratives, we applied the same code framing procedure to provide consistent analysis and interpretation. See Table 2 for a quick view of raw information extracted from the narratives before categorizing.

Table 2

Narratives: Guiding questions and sample responses

Pregunta	Respuesta de los estudiantes
1. ¿Cómo concibe o define usted el proceso de evaluación de los aprendizajes, ya sea de inglés u otros contenidos?	<p>S-1¹: Se diseñaron teniendo en cuenta la temática vista, se conocía lo que se iba a evaluar. Se dio una respectiva retroalimentación.</p> <p>S-2: Se evalúan las cuatro habilidades.</p> <p>S-3: Se mira el desarrollo de las habilidades por medio de un proceso.</p> <p>S-4: Muy en contexto con lo que se ha hecho.</p>
2. ¿Cree usted que el proceso de evaluación de los aprendizajes corresponde generalmente a lo enseñado por el/la docente o sugerido en el microdiseño curricular, ya sea de inglés u otros contenidos?	<p>S-1, S-2, S-3, S-4, S-6, S-7, S-8, S-9: Lo que se evalúa es lo que se enseñó durante el semestre.</p> <p>S-5: El actual proceso evaluativo debería preocuparse por generar distintas formas de evaluar de una manera más creativa y dinámica y no solo papel y lápiz como el estilo tradicional.</p> <p>S-8: Muchos temas no corresponden a lo enseñado en otras materias y a la hora de evaluar el estudiante queda perdido.</p>
3. ¿Qué sugerencias daría al proceso de evaluación desarrollado durante este curso de Inglés u o durante otros cursos actuales o previos en la carrera?	<p>S-1: La metodología del profesor debería tenerse en cuenta para otros cursos ya que se da una respectiva retroalimentación de las fallas que tenemos.</p> <p>S-2: Aprender fuera de la universidad con otras personas.</p> <p>S-3: Más claros en algunas preguntas o ejercicios. Las entregas de trabajos son una buena forma de ayudar.</p> <p>S-4: No me parece que se evalúe el vocabulario y la gramática aprendida correctamente.</p>

The data gathered from the instruments mentioned above allowed us to triangulate information and obtain an overall view of the results.

Findings

The analysis conducted through the software Atlas.ti yielded three main categories: perception of assessment as *assessment of learning (henceforth AOL)*, perception of assessment as *assessment for learning (henceforth AFL)*, and *suggestions regarding the assessment process*. Within each of these categories, some sub-categories were identified, as described below.

Students' Perceptions of Assessment

Students' perceptions of learning assessment were classified according to two sub-categories: *Assessment of Learning (AOL)* and *Assessment for learning (AFL)*.

¹ The nine students who responded the three questions of the narrative forms were coded with the letter S plus a number from 1 to 9. Some items are grouped based on the commonality of the responses.

Perceptions of assessment as AOL

Assessment is intended to elicit rote memorization.

Concepts provided by participants evidence the perspective of learning assessment as students' reproduction of information through memorization. The following excerpts, extracted from the different research instruments, support this concept (they are presented in the original language with corresponding translation):

Se miden los conocimientos a nivel memorista esenciales para poder alcanzar otro nivel de conocimiento. (E.S-Q19)

[Knowledge is measured at the memory level, essential to reach a higher level of knowledge.]

Es el recurso mediante el cual se puede saber si aprendiste lo que te ofreció el curso, pero sobre todo si memorizaste cada cosa. (E.S-Q19)

[It's the resource through which you get to know if you learned what the course offered you, but especially if you memorized everything.]

Assessment is regarded as a measurement of students' knowledge. Participants view classroom assessment as an opportunity for both teachers and students to know if the latter has learned the concepts presented throughout the course. The following excerpts support this statement:

Puedo decir que es un elemento que ayuda a medir nuestros conocimientos. (E.S-Q19)

[I can say that it's an element that helps us measure our knowledge.]

Es una medida para valorar el desempeño y el aprendizaje de los estudiantes. (E.S-Q19)

[It's a measure to assess students' performance and learning.]

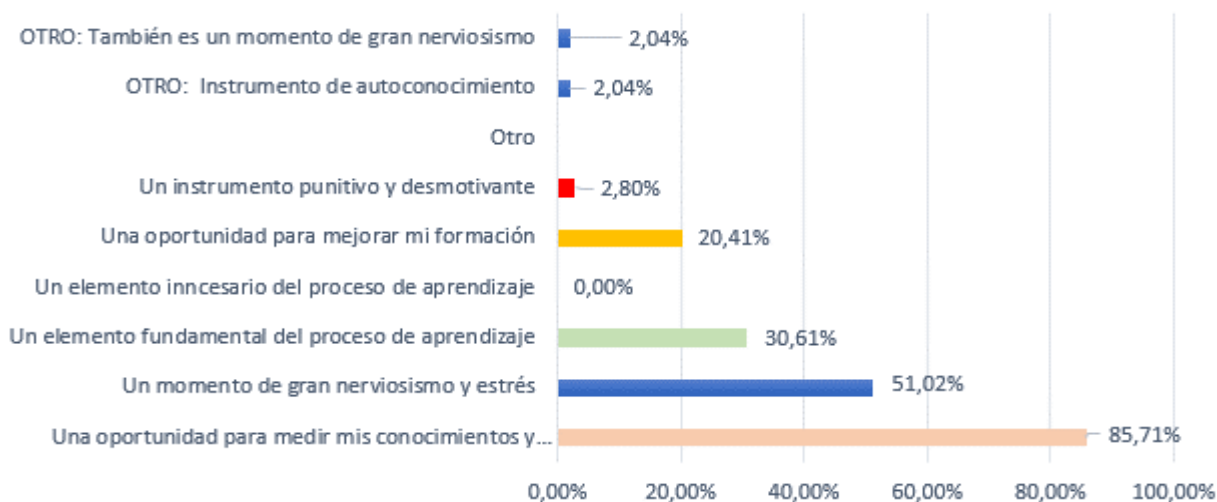
Mecanismos utilizados para evidenciar los niveles de apropiación de un tema. (E.S-Q19)

[Mechanisms used to show the levels of appropriation of a topic]

From the questionnaire, 85.7 % (see Figure 1) of participants described the assessment as “an opportunity to measure their knowledge and abilities” in response to the question “¿Con cuáles de los siguientes enunciados relacionas la evaluación?” (E.S-Q1).

Figure 1. Students' statements regarding assessment

Question 1: ¿Con cuáles de los siguientes enunciados relacionas la evaluación?
[Which of the following statements do you relate assessment with?]



Perceptions of assessment as AFL

Assessment as a multiple-perspective approach (Shohamy, 1998). According to participants, assessment entails evaluating students' knowledge through multiple ways that transcend the traditional paper-and-pencil method. The following excerpts support this statement:

Medir y corregir los errores que podemos presentar respecto a un tema, ya que no solo se puede evaluar

oral o en forma escrita sino de múltiples maneras. (E.S-Q19)

[Measure and correct the mistakes that we can make regarding a topic since we should not be assessed only through oral or writing but through multiple modes.]

El actual proceso evaluativo debería preocuparse por generar distintas formas de evaluar de una

manera más creativa y dinámica y no solo papel y lápiz como el estilo tradicional. (S.N-Q2:6-7)

[The current assessment process should focus more on generating different manners to assess in more creative and dynamic ways, and not only through the traditional paper and pencil style.]

Assessment as an opportunity to identify learners' strengths and weaknesses. Through assessment, both teachers and students can identify the students' weaknesses and strengths and take actions if necessary. Learners also see assessment as a moment to demonstrate their strengths and skills. The following excerpts support this statement:

Es conocer las habilidades y debilidades que un estudiante tiene y puede mejorar por medio de una evaluación. (E.S-Q19)

[It's to get to know the abilities and weaknesses a student has and can improve through assessment.]

La evaluación es el medio por el cual podemos identificar nuestras fortalezas y debilidades (E.S-Q19)

[Through assessment, we can identify our strengths and weaknesses.]

Assessment as a fundamental element in their learning process. This concept is taken from the questionnaire where 30.61 % of participants chose the statement "La evaluación es un elemento fundamental del proceso de aprendizaje" in response to the question "¿Con cuáles de los siguientes enunciados relacionas la evaluación?" (See Figure 1)

Assessment as an opportunity to enhance students' training. This concept is taken from the questionnaire where 20.41 % of participants chose the statement "La evaluación es una oportunidad para mejorar mi formación" in response to the question "¿Con cuáles de los siguientes enunciados relacionas la evaluación?" (See Figure 1)

Students' suggestions concerning the assessment process

Students provided the following suggestions regarding classroom assessments:

- implementing strategies that help students reduce their level of anxiety
- valuing the entire learning process rather than just the results of testing
- improving the quality of both the test preparation and some elements of the test
- working collaboratively with the teacher to improve classroom assessment practices
- designing assessments that transcend the elicitation of rote memorization

Implementing strategies that help students reduce their level of anxiety

Students suggested that teachers should implement strategies that reduce the impact of external factors (nervousness, anxiety, stress, and the like) on students' performance during assessment. According to the responses to the questionnaire, 51.02 % of the participants regard assessment as a moment of high stress. The statement "La evaluación es un momento de gran nerviosismo y estrés" was selected by students in response to the question: "¿Con cuáles de los siguientes enunciados relacionas la evaluación?" (E.S-Q1). Students' response to the high level of nervousness has to do with what Krashen stated (cited in Du, 2009) in his affective filter hypothesis; when the affective filter is active, it acts as a barrier provoking anxiety and fear failure in the acquisition of the L2. In language testing, anxiety is regarded as "a psychological condition in which a person experiences distress before, during, or after an exam or other assessment to such an extent that this anxiety causes poor performance or interferes with normal learning" (Du, 2009, p. 163). That is why teachers play a meaningful role in deterring unfavorable attitudes on assessment and encouraging students to study and view testing as a positive aspect of their learning process, creating a relaxed atmosphere for them.

The following excerpts support the statement above:

Normalmente, observo que varios compañeros presentan al igual que yo, mucho estrés, ansiedad y nervios, durante los exámenes de inglés. (E.S-Q20)

[I usually observe that several classmates feel, like me, a lot of stress, anxiety, and nerves during English exams.]

Pues no todo se puede evaluar oral o escrito, por ejemplo un estudiante puede conocer mucho del idioma pero sus nervios o la inseguridad le juegan una mala pasada. Entonces sería mejor desarrollar una manera en perder en aquel estudiante el miedo para sí luego evaluarlo. (E.S-Q20)

[Well, not everything can be assessed orally or in writing; for example, a student can be very competent in the language, but their nerves or insecurity play tricks on them. So, it would be better to help that student cope with their fear and then proceed with the assessment.]

Valuing the entire learning process rather than just the results of testing

According to participants, teachers should assess performance throughout the learning process as a whole, rather than just the outcomes of a test, whether a midterm or final exam. In their opinion, a test cannot serve as the sole evidence of students' knowledge. Students value both the product and the process. They also mentioned the importance of the 'process' when teachers make decisions regarding their learning processes. The following excerpts support this statement:

Una nota no define lo que un estudiante sabe. Mi sugerencia sería que las evaluaciones fueran tenidas en cuenta pero también todo el proceso que se lleva a cabo (formativo). (N.S-Q3: S-6, S-7)

[A grade does not define what a student knows. My suggestion is that evaluations be taken into account, but also the entire process (formative).]

En el caso de inglés que los exámenes equivalen al 80 % y las demás actividades al 10 % da la impresión de que el proceso de aprendizaje en este caso no importa. El proceso asegura el verdadero

aprendizaje, los exámenes solo aseguran un aprendizaje momentáneo que dura lo que tarde el examen en ser presentado. (N.S-Q3: S-7)

[In the case of the English exams that weigh 80 % and the rest of the activities only 10 %, it seems that the learning process is not important. The process ensures true learning, while exams only ensure momentary learning that lasts the time it takes the exam to be taken.]

Improving the quality of both the test preparation and some elements of the test

The participants provided the following suggestions regarding parameters and quality of tests: allotting more time for the writing task; reducing the quantity of information (topics) inquired on the test; making instructions more concise and clearer; improving the quality of the audio listening tracks, and training students in the case of computer-based tests. The following excerpts support this statement:

Más tiempo en las evaluaciones escritas (E.S-Q20)

[More time for the writing exams]

Que se evalúen menos temas a la vez (E.S-Q20)

[That fewer topics are assessed altogether]

Mejor calidad en los audios de listening, ya que algunas veces no se escucha con claridad o suficiente volumen (E.S-Q20)

[Better quality for the listening tracks since they are sometimes not clear and loud]

Working collaboratively with the teacher to improve classroom assessment practices

Students regard assessment as a process in which teachers and students should work closely to improve learning. Students suggested increasing teacher–student communication in order to share ideas regarding classroom assessment. The following excerpts support this statement:

Que se genere un ambiente más de confianza y armonía. Para que sus alumnos no se sientan

angustiados y presionados por cada vez que tengan un examen. Así se puede disfrutar del aprendizaje y aprendan a amar el conocimiento. (S.N-Q3: S-5)

[That an environment of trust and harmony be created so that students don't feel distressed and pressured every time they take an exam. This way, they can enjoy the learning process and learn to love knowledge.]

Podría mejorar si todos los estudiantes junto con el profesor comparten ideas para el proceso de evaluación. (E.S-Q20)

[It could improve if students, together with the teacher, share ideas regarding the assessment process]

Designing assessments that transcend the elicitation of rote memorization

Students suggested decreasing the types of assessments or questions aimed at eliciting rote memorization. The following excerpts support this statement:

Las evaluaciones que miden lo memorizado, deberían disminuir. (E.S-Q20)

[Assessments that require memorization should be discouraged.]

Algunos profesores quieren que nuestras respuestas sean memorizadas tal cual como aparecen en los textos. (E.S-Q20)

[Some teachers want our answers to be memorized as they appear in the texts.]

Conclusions and Discussion

The prevailing perception of classroom assessment in the present study participants is that of a summative process through which the teacher measures students' knowledge. The connotation of measurement, grades, errors, and memorization prevails in students' conceptualization of classroom assessment. These perceptions have been framed within the literature of classroom assessment and this study as *Assessment of Learning* (AOL). Thus,

there is a predominant perception of assessment as AOL with a clear summative function and the primary purpose of measuring what students have learned throughout a specific period of instruction.

To a lesser degree, according to our findings, there is a perception of assessment as a beneficial procedure through which students can demonstrate their abilities and skills; through which both teachers and students can identify strengths and weaknesses and take actions if necessary; through which teachers can improve the quality of instruction (washback effect), and as an opportunity to enhance students' formation. Thus, some participants do regard classroom assessment as a fundamental element in the learning process. Nevertheless, feelings of anxiety, nervousness, confusion, and stress are still present when facing the evaluation. These negative feelings seem to continue impacting learners' performance in terms of language use.

The assessment concept as a multiple-perspective approach was also evident, a perspective intended to take distance from an only mode of testing, usually characterized by a traditional paper-and-pencil type evaluation. Participants claim an assessment methodology that embraces the entire process, with multiple sources of learning evidence, and not just focused on a specific procedure. As Shohamy (1998) puts it, "the resulting multiplism means that in each of the assessment phases, it is possible to select features from a variety of options" (p. 242).

Regarding suggestions for classroom assessment, there is an emphatic call on making it a process that transcends the mere reproduction of factual information. In other words, students suggested that assessments elicit further analyses rather than just verbatim memorization. Teachers are called to switch the focus on classroom assessment from the product to the process. Participants also provided suggestions concerning technical aspects of testing—the quality of audio tracks, amount of content in exams, the time assigned for writing tasks, and the like—which may affect their performance and, therefore, their outcomes.

Thus, the present study provides us with a general picture of how a group of language students perceives classroom assessment based on their experiences. It is evident that when these learners were inquired about their perceptions of assessment, they tended to relate it to constructs such as testing, measurement, memorization, and errors, to mention some. Therefore, some of the suggestions they offered to make the assessment process fairer have to do with testing conditions, preparation, and test qualities. Nonetheless, learners also provided suggestions that can be interpreted as a move from formal testing to more formative assessments. These suggestions imply assessment as process-oriented rather than product-oriented, as a democratic construction between the teacher and the students, and an anxiety-free experience that may promote learning.

Additional to the suggestions provided by the participants of this study, we believe—as authors and researchers—that the incorporation of some assessment elements, which were not discussed in this study, can contribute significantly to the teaching–learning–assessment process. The implementation of timely and quality feedback, of self and peer assessments, portfolios, journals, and rubrics, and the permanent dialog between the teacher and students regarding assessment criteria and procedures, at the beginning and during the course, can add the formative assessment value implied by participants. The latter may foster a democratic approach to assessment, which seems to be absent in the process targeted in this study.

Despite the subjectivity of perceptions and beliefs regarding a specific subject matter, in this case, learning assessment, learners' perspectives offer teachers and researchers valuable reference of how we conduct the assessment in today's classrooms. These perspectives inform the field about assessment criteria, strategies, procedures, instruments, and practices. Nevertheless more research is needed to corroborate the perceptions, beliefs, and suggestions herein provided towards the consolidation of a formative approach to language learning assessment; an

approach that fosters democratic and fair evaluative practices; or as Poehner, Davin and Lantolf (2017) suggest, “assessment that promotes learner development” (p. 10).

References

- Areiza, H. N. (2013). Role of systematic formative assessment on students' views of their learning. *Profile*, 15(2), 165-183.
- Bachman, L., & Dambock, B. (2017). *Language assessment for classroom teachers*. Oxford University Press.
- Barkley, E., & Howell, C. (2016). *Learning assessment techniques*. Jossey-Bass.
- Brookhart, S. M., & Nitko, A. J. (2019). *Educational assessment of students*. Pearson.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.
- Carr, N. (2011). *Designing and analysing language tests*. Oxford University Press.
- Chappuis, J., & Stiggins, R. (2020). *Classroom assessment for student learning, doing it right – using it well*. Pearson.
- Creswell, J. (2012). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. SAGE.
- Du, X. (2009). The Affective Filter in Second Language Teaching. *Asian Social Science*, 5(8), 162–165. www.ccsenet.org/journal/index.php/ass/article/viewFile/3457/3131
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment. An advanced resource book*. Routledge.
- Hancock, C. (2006). Language teaching and language testing: A way and ways to make appropriate connections. *Acción pedagógica*, 15, pp. 6–13.
- Harmer, J. (2007). *The Practice of English Language Teaching*. Longman.
- Herrera, L., & Zambrano, L. (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>
- Hughes, A. (2003). *Testing for language teachers*. Cambridge University Press.
- Iseni, A. (2011). Assessment, Testing and Correcting Students' Errors and Mistakes. *Language Testing in Asia*, 1(3), 60-90. <https://doi.org/10.1186/2229-0443-1-3-60>

- Jang, E. (2014). *Focus on assessment. Oxford Key Concepts for the Language Classroom*. Oxford University Press.
- López, A., & Bernal, R. (2009). Language Testing in Colombia: A Call for More Teacher Education and Teacher Training in Language Assessment. *Profile*, 11(2), 55–70. <http://www.revistas.unal.edu.co/index.php/profile/article/view/11442>
- Moss, C. M., & Brookhart, S. M. (2009). *Advancing formative assessment in every classroom. A guide for instructional leaders*. Alexandria: ASCD.
- Muñoz, A., Palacio, M., & Escobar, L. (2012). Teachers' Beliefs about Assessment in an EFL Context in Colombia. *Profile*, 14(1), 143–158. <http://www.revistas.unal.edu.co/index.php/profile/article/view/29064>
- Mussawy, S. (2009). *Assessment Practices: Student's and Teachers' Perceptions of Classroom Assessment* [Master thesis]. University of Massachusetts, School of Education. http://scholarworks.umass.edu/cie_capstones/9/?utm_source=scholarworks.umass.edu%2Fcie_capstones%2F9&utm_medium=PDF&utm_campaign=PDFCoverPages
- Palacio, M., Gaviria, S., & Brown, J. (2016). Aligning English language testing with curriculum. *Profile Issues in Teachers' Professional Development*, 18(2), 63–77. <http://dx.doi.org/10.15446/profile.v18n2.53302>
- Poehner, M. E., Davin, K. J., & Lantolf, J. P. (2017). Dynamic Assessment. *Language Testing and Assessment*, 243–256. https://doi.org/10.1007/978-3-319-02261-1_18
- Popham, W. (2008). *Transformative assessment*. Association for Supervision and Curriculum Development. Alexandria, VA. <https://titleiidgrants.wikispaces.com/file/view/Transformative+Assessment.pdf>
- Shohamy, E. (1998). Evaluation of learning outcomes in second language acquisition; a multiplism perspective. In H. Byrnes (Ed.), *Learning foreign and second languages: perspectives in research and scholarship* (238–261). Modern Language Association of America.
- Suskie, L. (2018). *Assessing student learning. A common sense guide*. Jossey-Bass.
- Turner, C. E. (2017). Classroom assessment. In G. Fulcher, & F. Davidson, *The Routledge handbook of language testing* (pp. 65–78). Routledge.
- William, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade, & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18–40). Routledge.

Annex 1. Students' Answers

Pregunta	Respuesta de los estudiantes
1. ¿Cómo concibe o define usted el proceso de evaluación de los aprendizajes, ya sea de inglés u otros contenidos?	<p>S-1: Se diseñaron teniendo en cuenta la temática vista, se conocía lo que se iba a evaluar. Se dio una respectiva retroalimentación.</p> <p>S-2: Se evalúan las cuatro habilidades.</p> <p>S-3: Se mira el desarrollo de las habilidades por medio de un proceso.</p> <p>S-4: Muy en contexto con lo que se ha hecho.</p> <p>S-5: No estoy de acuerdo con el writing, ya que esta habilidad representa mucha importancia. Considero que el tiempo que se da para escribir un ensayo es muy corto. Nos garantiza que expresemos de una manera más confiada y segura nuestras ideas.</p> <p>S-6: Es un proceso tedioso que requiere un docente capacitado y preparado.</p> <p>S-7: Considero que es apropiado y se adecua con lo visto (inglés). En caso de otras materias es muy inapropiada y tradicional.</p> <p>S-8: En otras materias no fue claro, no cumplen con el microdiseño curricular.</p> <p>S-9: Es una forma de demostrar que sí hemos adquirido un conocimiento.</p>

Pregunta	Respuesta de los estudiantes
<p>2. ¿Cree usted que el proceso de evaluación de los aprendizajes corresponde generalmente a lo enseñado por el/la docente o sugerido en el microdiseño curricular, ya sea de inglés u otros contenidos?</p>	<p>S-1, S-2, S-3, S-4, S-6, S-7, S-8, S-9: Lo que se evalúa es lo que se enseñó durante el semestre S-5: El actual proceso evaluativo debería preocuparse por generar distintas formas de evaluar de una manera más creativa y dinámica y no solo papel y lápiz como el estilo tradicional. S-8: Muchos temas no corresponden a lo enseñado en otras materias y a la hora de evaluar el estudiante queda perdido. S-9: Cuando las evaluaciones son bien diseñadas y llevan contenidos que el estudiante ha aprendido los resultados se ven reflejados.</p>
<p>3. ¿Qué sugerencias daría al proceso de evaluación desarrollado durante este curso de inglés u o durante otros cursos actuales o previos en la carrera?</p>	<p>S-1: La metodología del profesor debería tenerse en cuenta para otros cursos ya que se da una respectiva retroalimentación de las fallas que tenemos. S-2: Aprender fuera de la universidad con otras personas. S-3: Más claros en algunas preguntas o ejercicios. Las entregas de trabajos son una buena forma de ayudar. S-4: No me parece que se evalúe el vocabulario y la gramática aprendida correctamente. S-5: Que se genere un ambiente más de confianza y armonía. Para que sus alumnos no se sientan angustiados y presionados por cada vez que tengan un examen. Así se puede disfrutar del aprendizaje y aprendan a amar el conocimiento. S-6, S-7: Una nota define lo que un estudiante sabe. Mi sugerencia sería que las evaluaciones fueran tenidas en cuenta pero también todo el proceso que se lleva a cabo (formativo). S-7: En el caso de inglés que los exámenes equivalen al 80 % y las demás actividades al 10 % da la impresión de que el proceso de aprendizaje en este caso no importa. El proceso asegura el verdadero aprendizaje, los exámenes solo aseguran un aprendizaje momentáneo que dura lo que tarde el examen en ser presentado. S-8: El curso de inglés 5 fue acorde a lo enseñado. S-9: Más importancia al uso del inglés con propósitos comunicativos y que se combinara con el uso de estructuras gramaticales.</p>