



Content list available at <http://ijltr.urmia.ac.ir>

*Iranian Journal  
of  
Language Teaching Research*



Urmia University

---

## Do Students Feel that they are Assessed Properly?

Ali Işık<sup>a,\*</sup>

*<sup>a</sup>Istinye University, Turkey*

---

### ABSTRACT

As each and every language learner is subject to assessment, a sound and valid assessment plays a pivotal role in foreign language education. This study focuses on how assessment is perceived by English as a foreign language (EFL) students in the Turkish EFL context with the participation of 481 EFL students from 24 K-12 level schools and 8 universities. A mixed-methods research design was adopted, and the data were collected through a questionnaire, follow-up interview, and observation. The results showed that the students were not satisfied with the assessment practices and they did not feel like they were assessed. It was also observed that the traditional approach focusing on the formal properties of English was commonly practiced while assessing the students. Moreover, it was found that the assessment quality in the schools was low and it was taken as a formal requirement to grade students. The final part of the paper suggests the need for a comprehensive teacher-training in assessment to increase the assessment literacy of English teachers.

**Keywords:** assessment; testing; teacher training; English teaching; English learners

© Urmia University Press

---

### ARTICLE HISTORY

**Received:** 8 July 2019

**Revised version received:** 14 Nov. 2019

**Accepted:** 20 Dec. 2019

**Available online:** 1 Jan. 2020

---

\* Corresponding author: Department of English Language & Literature, Istinye University, Istanbul, Turkey

Email address: [isikal@hotmail.com](mailto:isikal@hotmail.com)

© Urmia University Press

## Introduction

It has been widely accepted that assessment has a strong effect on foreign/second language education and its quality, and it has been a popular topic of interest in language education (Baker, 2016; Fulcher, 2012; Menken, Hudson, & Leung, 2014; Tsagari & Vogt, 2017). Consequently, the power of assessment in language education programs has started to receive considerable attention and the quest for constructing valid, fair, accountable, and carefully-designed assessment which is compatible with the language program and the adopted language teaching methodology has gained momentum (Brown, 2004; Brown, 2005; Johnson, Becker, & Olive, 1999; Malone, 2013; Tsagari, 2016). Accordingly, language teachers, as one of the major agents of assessment, have been the center of attention; moreover, what language teachers should know and how they should apply their conceptual knowledge in assessment have been studied extensively (Davison & Leung, 2009; Douglas, 2018; Farhady & Tavassoli, 2018; Hill & McNamara, 2011; Rea-Dickins, 2004; Tsagari, 2016). Another issue has to do with the relationship between assessment and curriculum. It has been argued that assessment is to be considered within the framework of language teaching program and methodology, not as an independent entity, and teachers are to be equipped with fundamental theoretical knowledge and practical skills and strategies to design and implement high quality, context-sensitive assessment (Brunfaut, 2014; Douglas, 2010; Jiang, 2017; Norris, 2016; Scarino, 2013; Tsagari, 2016; Tsagari et al., 2018).

Inspired by the discussions on the role of assessment in language education, this study attempts to depict how students of English as a foreign language (EFL) perceive assessment in English education in Turkey. It also aims to raise awareness among teachers and other authorities by shedding light on the essential role of assessment in English education and how it is perceived by students in order to improve the quality of assessment in EFL. Finally, it aims at contributing to the data on assessment in EFL regarding student perception.

## Review of Literature

Although assessment is highly emphasized in language education, there are different perspectives on how to realize it. Different approaches determine how assessment is conceptualized, practiced and exploited in language education (Alderson, Brunfaut, & Harding, 2017). The product-oriented traditional testing approach, which generally targets testing formal aspects of language and comprehension, adopts the psychometric tradition and gives priority to objectivity, which places a premium on testing the end product expressed in terms of numbers and statistics (Brown, 2005; Combee, Folse, & Hubley, 2011; Earl, 2012; Stolz, 2017; Tsagari & Banerjee, 2016; Tsagari et al., 2018; William, 2011). Moreover, the tests in this approach are geared towards “assessment of learning”, that is, the final product or the scores students get are meant to measure and determine the linguistic levels of students, teacher performance, and the efficiency of the language program and schools (Brunfaut, 2014; Combee et al., 2011; Earl, 2012; Inbar-Lourie, 2008; Malone, 2013; Purpura, 2016; Tsagari, 2016).

On the other hand, the current assessment approaches, which can also be named as alternatives in assessment or alternative assessment, strive to obtain the whole picture about students by considering their performance in and out of the classroom. They stress the importance of student-centeredness, the learning process, relevant and meaning-based activities, authenticity, and holistic approaches to language and stakeholders (Douglas, 2018). Using multiple data eliciting instruments, they look for triangulation to have a highly accurate evaluation of the students. They favor “assessment for learning” and exploit assessment to improve language education, not only to test student and teacher performance and the language program but also school efficiency (Fox, 2015; Green, 2018; Harsch, 2014; Linfield & Posavac, 2018; William, 2011). The data obtained from assessment are exploited to identify the strengths and weaknesses

in the language programs and to improve the quality of language education. Students are also informed about their performance and assume responsibility to better their learning. The primary focus is not only the end-product but the whole teaching and learning process considering all factors in and out of the language education context that have an impact on language education. Although they seem revolutionary and convincing, these ideas are slow to be adopted and put into practice (Barootchi & Keshavaraz, 2002; Brown, 2005; Brunfaut, 2014; Combee et al., 2011; Earl, 2012; Purpura, 2016; Tsagari, 2016; Tsagari & Vogt, 2017; Tsagari et al., 2018).

The discussion of the agents of classroom-based assessment has also received attention and their roles in assessment are depicted in accordance with the adopted approach of assessment. In the traditional approach, teachers are considered agents and students as the object of testing. Teachers have full control over preparing, administering and scoring the tests. On the other hand, alternative assessment considers learning as a social process and assigns a role in assessment to each and every stakeholder who has a share in the language education process. As the variety and amount of information coming from different sources help complete and see the whole picture of assessment, the contributions of teachers, students, peers, school administration, and families are highly appreciated (Inbar-Lourie, 2008; Leung, 2004; Luoma, 2002; Lynch, 2001; Lynch & Shaw, 2005; McNamara & Roever, 2006). The current approaches, emphasizing learner autonomy, naturally encourage and value student contribution to assessment. Through different alternative means, students are expected to assess their own performance, the performance of their peers, the teaching performance of their teachers, and each and every aspect of a language program such as materials, tasks, learning context, language teaching methodology, measurement, and the system of evaluation. Teachers provide students with guidance about how to assess and use the information obtained from the assessment to monitor their own learning and better their performance (Combee et al., 2011; Fox, 2015; Harsch, 2014; Little, 2005). Ultimately, alternative assessment can pave the way for a different perspective on assessment, "assessment as learning", to conceptualizing learning, teaching and assessment as the parts of a whole, not separate entities (Dann, 2014).

Although the approaches differ in how they conceptualize assessment, they concur on the role of assessment in language education, namely the washback effect. Assessment may foster or hinder the language teaching/learning process. It is likely that teachers, students, and administrators direct their efforts to the content and task types of the teaching/learning process as a means of getting good grades or becoming successful. Thus, assessment holds the power to frame the language education process in general; its goals, teaching and learning strategies, materials and tasks, target language knowledge and skills, and the relevant assessment and evaluation system (Agrawal, 2004; Spolsky, 2008). If there is one to one correspondence between what is done in and out of the language education context and what is assessed, the effect of assessment (washback) on teaching and learning is positive, and if not, it is negative. To manage the positive washback effect, the manner of assessment should be based on the careful analysis of language materials, type of tasks, content-focus, topic-focus, and time allotted to each skill, content, and topic. In other words, it is imperative to create assessments using similar materials, tasks, contents, topics, and skills and allotting the same amount of weight for each of these in accordance with the teaching and learning process. If it cannot be managed, assessment does not really reflect the teaching and learning process. In this case, there occurs a mismatch between the teaching and learning process and the assessment process. Consequently, stakeholders are likely to give priority to what is assessed to be successful, which creates a negative washback effect on the language education process. Generally speaking, high-stakes assessment, which assigns passive roles to both teachers and students, provides a good example of this negative situation. It focuses on the end-product and tests limited content, namely formal properties of language and comprehension, by using easily scorable, objective testing techniques. Both teachers and students spend their time and effort to be successful on such exams. Furthermore, language activities are

ultimately geared towards success in those language programs and may turn into exam preparation programs (Little, 2005). Hence, assessment outlines the language education process, and a carefully-designed assessment process provides a fine balance between what is taught and learned and what is assessed (Brown, 2005; Combee et al., 2011; Fox, 2015).

These discussions on assessment have also stimulated research on different aspects of assessment. Hasselgreen, Carlsen, and Helness (2004), Scarino (2013), Tsagari and Vogt (2017), and Vogt and Tsagari (2014) studied what English language teaching (ELT) teachers know about assessment and how they put their knowledge into practice. These studies, focusing on the assessment literacy of the teachers and their needs for training, indicated that the undergraduate programs did not prepare ELT teachers well enough in terms of assessment literacy and the teachers were not competent enough to carry out an effective assessment. Bailey and Brown (1996) and Brown and Bailey (2008) who compared the content of the assessment courses in 1996 and 2008, indicated that the content of assessment courses improved gradually. In another study, Heritage (2007) examined the assessment training of College EFL teachers in China which revealed a negative portrait. It was found that the EFL teachers were not given enough assessment training in their undergraduate and graduate programs and received no in-service assessment training. It was also found that the teachers were far from realizing the natural and indispensable mutual relationship between language education and assessment, and they considered assessment as an extra workload. Likewise, Shohamy, Inbar-Lourie, and Poehner (2008) indicated inadequate teacher training in assessment. They found out that teachers did not receive enough quality training related to assessment and consequently they lacked the fundamentals and the skills to practice quality assessment. They followed the traditional assessment and assessed the linguistic aspects of language employing traditional question types. Another study conducted by López Mendoza and Bernal Arandia (2009) in Colombia pointed out a positive correlation between assessment and teacher training in testing, but it was also observed that the teachers in Colombia did not receive enough quality training in testing and they employed traditional product-oriented tests. Consequently, testing had no effect to improve the quality of English teaching/learning. From a different yet a similar perspective, Xu and Liu (2009), who investigated teachers' knowledge of assessment and factors affecting their assessment practice through a narrative inquiry of a Chinese college EFL teacher's experience, found that the way teachers practiced assessment was shaped by their need to conform to how their colleagues practiced assessment. Similarly, Jin (2010) conducted a study to investigate the assessment-related teacher training and assessment practices in Chinese universities with the participation of 86 ELT teachers. It was found that the assessment courses offered to train teachers were adequate in the content; however, that theoretical knowledge was not put into practice adequately in language classrooms. Finally, Scarino (2013) summarized the examples of some aspects of developing language assessment literacy. She indicated that teachers needed to develop their theoretical knowledge base and be ready to implement context-sensitive testing practices.

In Turkey, assessment has not aroused enough attention among researchers. In one study, Saricoban (2011) examined the tests prepared by English teachers and pointed out validity problems, especially with respect to the target grammar items. Ozdemir-Yilmazer and Ozkan (2017) investigated the assessment practices in a university prep school and reported that assessment procedures were dominated by the proficiency exam which functioned as a gateway for students to continue their academic programs. In another study, Mede and Atay (2017) replicated the study carried out by Vogt and Tsagari (2014) with the participation of the English teachers at a university English preparatory program. They found that the assessment literacy of the teachers was limited with respect to classroom-based assessment and assessment-related concepts.

As in the studies summarized above, research on assessment focuses on ELT teachers and ELT teacher training. To the knowledge of this researcher, research on students regarding assessment

is non-existent or scarce. This study, inspired by the discussions about assessment, attempts to present a broader perspective about assessment by eliciting data from the students studying EFL in different types of language programs in different institutions in Turkey. This study attempted to study the perceptions of students from both public and private language teaching institutions at all levels of education to shed light on country-wide assessment practices in Turkey. Ultimately, it aims to investigate the perceptions of Turkish EFL students about both low-stake and high-stake assessment practices in the Turkish EFL context. Moreover, the study attempts to raise awareness on the role of assessment in English education to improve the quality of assessment-related practices. Finally, it aims to contribute to data regarding student perception of assessment in English education. More specifically the study focused on the following research questions:

1. How is assessment perceived by the Turkish EFL students?
2. Does the level of education affect how the Turkish EFL students perceive assessment?
3. Does the school type (public vs. private) affect how the Turkish EFL students perceive assessment?

## Methodology

### *The Turkish EFL Context*

In Turkey, EFL education is one of the major concerns of both government and parents in public schools, which are tuition-free, and private schools that require a tuition fee. Some parents send their children to private schools especially to have them receive more and better EFL education. Officially EFL education starts in the second grade, but private schools start EFL in kindergarten. At the K-12 level, (primary school, grades 1-4, secondary school, grades 5-8, and high school, grades 9-12) students must take at least two written exams a semester and the teachers must give at least one oral exam grade, a kind of performance grade, to their students. The average of the written exams and the oral exam(s) determines the English grade of the students for a given semester. At the end of the secondary school, all the students take a centralized exam administered by the government; their ranking on this exam determines what high school they can attend. As the students are also required to answer multiple-choice questions testing grammar, vocabulary, and reading comprehension on the high-stakes high school entrance exam, both English teachers and students gear their efforts especially in the eighth grade to prepare for the exam. Consequently, grammar, vocabulary, and reading comprehension get primary emphasis.

At the end of the 12<sup>th</sup> grade, all students must take a central university entrance exam administered by the government. Their performance on this test determines which university they can attend. The students who take the English component of the university exam must answer multiple-choice questions testing their knowledge of grammar, vocabulary, reading comprehension, and ability to do English-Turkish translation. Ultimately, both the students and teachers give priority to exam preparation and the English component of the exam is prioritized especially in the twelfth grade. Outcomes are contingent on student performance and the results of the university exam. Once admitted to English-medium programs, students are expected to pass the proficiency exams prepared by the respective testing office in order to begin their departmental studies. The proficiency exam tests their basic knowledge of English skills (grammar, vocabulary, reading, and writing).

Monthly achievement tests covering grammar, vocabulary, reading, and writing are administered in the prep class. At the end of each semester, students may take the proficiency exam. Once they

pass the prep class and start their academic programs, they take two English exams each semester in the first year. After the first year, English becomes an optional course and if they choose it, they continue taking two exams each semester.

Generally speaking, at the K-12 level, there is no testing office and English teachers prepare their own tests as a group. They either use the assessment component of their English coursebooks or refer to the internet to prepare their tests. At the university level, measurement and evaluation are the responsibility of the testing office. Volunteer English teachers who are approved by the university administration and not necessarily having any training and experience in assessment are accepted to the testing office not necessarily having any training and experience in assessment. Teachers administer the tests the testing office prepares. The evaluation is carried out and the results are announced by the testing office. The testing office does not share any further information with teachers and students.

### *Participants*

The target population of this study is comprised of primary to tertiary EFL students from the public and private schools offering English at all levels of education. Upon completing the bureaucratic procedures and receiving the official approval of the authorities, the researcher/s selected eight primary, eight secondary, eight high schools, and eight universities from two major cities in Turkey using random cluster sampling. The questionnaire was administered to 509 K-12 students and 297 university students, and among those, 284 K-12 students and 197 university students answered the questionnaire in the fall semester of 2018. In other words, 481 EFL students provided data about assessment. The distribution of the participants is summarized in Table 1.

Table 1  
*The Distribution of the Participants*

Education level	Context	Number of students
Primary	Public	41
	Private	36
Secondary	Public	52
	Private	49
High school	Public	50
	Private	56
University	Public	103
	Private	94

For the follow-up interview, three students from each institution were chosen through random cluster sampling. Since the exams were completed the week before the final week and classes were informally over, the interview was conducted in the final week of the semester, which was the most appropriate time to carry out interviews. One participant was interviewed at a time and the interview was recorded to be transcribed for analysis.

### *The Questionnaire Development Process*

To the knowledge of the researcher, there is no available questionnaire focusing on student perceptions of assessment and there was a need to develop a questionnaire for this purpose. Thus, the researcher, who was also the trainer in the program, added a questionnaire development component to the 15-week training program on assessment. Throughout the course, 25 ELT trainees met three hours a week and covered the fundamentals of assessment, evaluated and developed items on specific skills, studied and designed alternatives in assessment, and designed a general assessment plan for a specific EFL course. In the final part of the program, they

developed questionnaires in groups. one of the requirements of the training was to conduct field research to delineate how assessment was planned, designed, implemented, evaluated and how it was perceived by the students, each trainee was to visit a school once a week and observe assessment practices there and report them at the end of the training. Thus, the questionnaire finalized at the end of the training was used to collect data that were exploited in this study.

After familiarizing themselves with the fundamentals of assessment and studying the questionnaires developed by the other researchers, the trainees worked in groups of three or four and started to develop their own questionnaire. The trainer, who had evaluated, adapted, and developed assessment for various English education programs, worked as an advisor and offered courses on assessment since 1999, collaborated with each group. The groups, who started working on the questionnaires, exploited the ones developed by other researchers to create their own. They either transferred items directly or adapted and organized them on their own. With the help of the e-group formed at the beginning of the training, the groups were able to share their questionnaires with the other groups and the trainer. In other words, they co-worked and provided/received feedback about their work, and this collaboration resulted in their revising their questionnaires as much as they wanted to. Thus, the task became an ongoing process that was carried out both in and out of the classroom. The groups shared the final versions of their questionnaires with others in the e-group to make sure that each trainee came to the class having examined all of them and ready to give feedback to their peers about their questionnaires during class presentations. In class, each group presented their checklists and got feedback both from their peers and the trainer. After finalizing their own, the groups got together and constructed one questionnaire. Finally, the trainer received and evaluated the questionnaires, which formed the first step toward achieving validity. The expert panel consisting of a group of three lecturers evaluated the content relevance and representativeness of the questionnaires. Upon receiving feedback from the panel, the trainer revised them again and piloted them with 33 EFL students. Cronbach's Alpha reliability coefficient of the questionnaire was found to be 0.765. The expert panel evaluated the data and made suggestions. Considering the feedback, the items with low-reliability coefficients were modified and the questionnaires were piloted again with the participation of 20 EFL students. The Cronbach's Alpha reliability coefficient of the modified teacher questionnaire was .815.

#### *Data Collection*

With the help of a mixed-methods research design, the data were collected through the questionnaire, follow-up interviews, observation from each type of institution (public and private) at each level of education (primary, secondary, high school, university), and the evaluation of a teacher-made test.

#### *Questionnaire*

The questionnaire was administered to the students in order to elicit their impressions of the assessment practices in their institutions (see Appendix 1). While the first item was for eliciting information about their current level of education, the rest of the items were for eliciting their perceptions on assessment. In December 2018, the trainees visiting each school presented the questionnaires to the students and explained what they were required to do. Then, the students were given the questionnaires and were requested to answer them by the following week. The internal consistency reliability of the questionnaire was computed by using Cronbach's alpha and found to be .93, indicating a high level of internal consistency.

*Follow-up Interviews*

To understand the assessment process better, three EFL students from each type of schools at different levels of education, 18 EFL students in total, were interviewed about the topics covered in the questionnaire. The interview process was carried out in Turkish (mother tongue) with the students to help them express themselves clearly without any language barrier.

*Observation*

As a requirement of the training process, each trainee was required to visit a school once a week for a semester to observe the assessment-related practice using the guiding questions for observation (see Appendix 2). They witnessed how assessment was planned, prepared, administered, and evaluated and how the students were provided with feedback. They reported their observations to the instructor. The researcher himself visited one school type, public and private, at each level of education three times a semester.

*The Evaluation of the Teacher-made Tests*

To glean insight into the teacher-made tests, a sample test was evaluated to convey the physical layout and the content of a typical exam (see Appendix 3).

*Data Analysis*

SPSS was used to analyze the data obtained from the questionnaires. The responses were analyzed through descriptive statistics to obtain frequencies and percentages. The interviews were recorded and transcribed. One-way ANOVA was used to check if the school type and the level of education affected the responses given by the students. To do these analyses, the total scores of the items in the questionnaire for each group were calculated separately and then using one-way ANOVA the mean scores of the groups were compared to each other. The data elicited from the interviews were categorized, coded and the frequencies and percentages were presented. The data obtained through observation were also categorized and reported.

**Results***The Questionnaire*

The results obtained from the student questionnaires are summarized in the following tables.

*Comparison of the Groups*

Table 2 and Table 3 illustrate the results of the comparison of the groups with respect to the type of school and level of education.

Table 2  
*The Comparison of the Learners with Respect to the Type of Schools (Public and Private)*

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.018	2	0.009	0.679	0.508
Within Groups	5.973	450	0.013		
Total	5.991	452			



As can be observed in Table 2, there was no significant difference between students attending the private and public institutions ( $F(452,2)=0.679, p>.05$ ).

Table 3  
*The Comparison of the Learners with Respect to the Level of Schools*

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.016	10	0.008	0.676	0.514
Within Groups	5.956	450	0.015		
Total	5.972	452			

Table 3 shows that the level of education did not affect significantly how the students perceive assessment practices ( $F(452,10)=0.676, p>.05$ ).

#### *The Frequency of Assessment*

Table 4 exhibits the findings of how frequently the students were assessed.

Table 4  
*The Frequency of Assessment*

What kind of assessment/tests do you take and how often?	Quiz		Exam		Final	
	F*	P**	F*	P**	F*	P**
once a week	21	4%	0	0%	0	0%
every 2 weeks	36	8%	0	0%	0	0%
once a month	27	6%	398	83%	0	0%
once a semester	15	3%	146	30%	1	0%

\*Frequency \*\*Percentage

The results regarding "quiz" indicated that 4% of the students received quizzes once a week. 8% of them reported that they took quizzes every two weeks, 6% once a month, and 3% once a semester. In terms of exams, 83% of the students said that they received exams once a month, and 30% once a semester.

#### *Satisfaction with the Frequency of Testing*

Table 5 summarizes the satisfaction level of the students regarding the frequency of assessment.

Table 5  
*Satisfaction with the Frequency of Assessment*

Is the frequency of testing enough?	Frequency	Percent
very satisfactory	0	0%
satisfactory	30	6%
partially satisfactory	301	63%
unsatisfactory	148	31%
very unsatisfactory	0	0%

It can be observed that the overwhelming majority of the students were not satisfied with the frequency of assessment.

*The Means of Assessment*

Table 6 depicts what other means teachers use to assess the performance of their students.

Table 6  
*Other Instruments Teachers Use in Assessment*

What other means does your teacher employ to evaluate your performance?	Frequency	Percent
project	104	22%
homework	204	43%
oral exam	9	2%

In terms of the means of assessment excluding exams, 33% of the students reported that no other means of assessment were used in addition to pen-and-paper assessment. Homework was marked as a means employed in assessment by almost half of the students. About one-fifth of the students stated that they were evaluated through their projects.

*Content Validity*

The findings of the content validity of the assessment the students received are summarised in Table 7.

Table 7  
*The Content Validity of Assessment*

Does it cover what you study in the class?	Frequency	Percent
very satisfactory	0	0%
satisfactory	2	0%
partially satisfactory	279	58%
unsatisfactory	198	41%
very unsatisfactory	0	0%

Table 7 indicates that students were not happy with what was covered in assessment.

*Satisfaction with the Assessment of Performance*

Table 8 summarizes to what extent the assessment was successful in assessing the performance of students.

Table 8  
*Satisfaction with the Performance Assessment*

Do tests measure your performance?	Frequency	Percent
never	4	1%
rarely	286	60%
occasionally	185	39%
often	4	1%
always	0	0%

The majority of the students were not satisfied with their assessment performance. About one-third of them stated that they were occasionally satisfied with their assessment performance.

*Assessment Instruments and Language Learning Goals*

The compatibility between the students' language learning goals and the assessment instruments is summarized in Table 9.

Table 9  
*The Compatibility between the Students' Language Learning Goals and the Assessment Instruments*

Which of the following serves your language learning goals?	Frequency	Percent
essay	79	17%
true-false	28	6%
multiple choice	233	49%
short answer	127	27%
cloze	43	9%
matching	12	3%
other	8	2%

About half of the students reported that the multiple-choice type of items served their language learning goals. About one-third of them found short answer type of items compatible with their language learning goals, and about one-sixth of them thought so for the essay type of items.

*Fairness of the Assessment Instruments*

Table 10 indicates how the students evaluate assessment instruments in terms of fairness.

Table 10  
*The Fairness of Assessment Instruments*

Which are fair tests?	Frequency	Percent
essay	68	14%
true-false	27	6%
multiple choice	204	43%
short answer	103	22%
cloze	24	5%
matching	5	1%
alternative assessment	26	5%
other	5	1%

Similar to the results reported in Table 9, about two-fifths of the students found the multiple-choice type of items fairer. About one-fifth of them indicated the short answer type of items as fairer, and about one-sixth went for the essay type of items in terms of fairness.

*Instruments Reflecting Student Performance*

Table 11 provides data about how the students evaluate different instruments in terms of their congruity reflecting their performance.

Table 11  
*Instruments Reflecting Student Performance*

Which helps you reflect your real performance?	Frequency	Percent
essay	58	12%
true-false	16	3%
multiple choice	105	22%
short answer	49	10%
cloze	8	2%
matching	1	0%
alternative assessment	75	16%
other	0	0%

The students ranked the item types regarding their effectiveness in reflecting student performance and marked multiple-choice, alternative assessment, essay, and short answer types of items in their respective order.

*The Way the Teachers Exploit Assessment Data*

Table 12 provides the results concerning what the teachers did with the assessment data.

Table 12  
*How the Teachers Benefited from Assessment Data*

How does your teacher use test results?	Frequency	Percent
to give feedback about learning	16	3%
to go over the problematic topics/skills	3	1%
to give grades	464	97%
to diagnose problems	14	3%
other	0	0%

The overwhelming majority of the students indicated that the data obtained from assessment were used to grade students.

*The Way the Students Exploit Assessment Data*

Table 13 provides the results concerning what the students did with the assessment data.

Table 13  
*How the Students Benefited from Assessment Data*

How do you use test results?	Frequency	Percent
to see my weaknesses and strengths	33	7%
to review what I have not learned well	49	10%
to learn my grades	469	98%
other	0	0%

Almost all the students reported that they exploited assessment results to learn their exam grades.

*The Student Role in the Assessment Process*

The findings of the role of the students in the assessment process are given in Table 14.

Table 14  
*The student role in the assessment process*

Do you have a voice in the test preparation process?	Frequency	Percent
Never	472	99.6%
Rarely	2	0.4%
Occasionally	0	0%
Often	0	0%
Always	0	0%

As the table indicates, almost all the students stated that they had no voice in the assessment process.

*The Student Role in Different Types of Assessment*

To what extent students took part in different types of assessment is summarized in Table 15.

Table 15  
*The Student Role in Different Types of Assessment*

How frequently do you take part in the following assessment types?	never		rarely		occasionally		often		always	
	F*	P**	F*	P**	F*	P**	F*	P**	F*	P**
self-assessment	478	100%	0	0%	0	0%	0	0%	0	0%
peer assessment	478	100%	0	0%	0	0%	0	0%	0	0%
co-assessment	478	100%	0	0%	0	0%	0	0%	0	0%

\*Frequency \*\*Percentage

The table clearly indicates that the students never took part in self-assessment, peer assessment, and co-assessment.

*The Assessment-related Difficulty*

Table 16 depicts the difficulty the students experienced in assessment.

The content of the assessment they took and the time given for assessment were reported as the main sources of the difficulty by the overwhelming majority of the students. Similarly, the number of questions on an exam was indicated as another source of difficulty by the majority of students. Moreover, the physical conditions of the classroom were believed to cause difficulty by one-fifth of the students. Finally, the page layout of the exam paper was marked as another challenge by one-tenth of the students.

Table 16

*The assessment-related difficulty*

What causes difficulty for you on the exam?	Frequency	Percent
content of the questions	381	80%
time allotted	394	82%
instructions	34	7%
number of questions	340	71%
sequence of questions	21	4%
page layout of the exam paper	47	10%
physical conditions (heat, light, noise, etc.) of the exam room	104	22%
Other	0	0%

***Student Interviews***

The data obtained during the interview were categorized and coded and the percentages were provided as follows:

*The Meaning of Assessment*

All the students indicated that assessment was a part of their school life. 84% of the students indicated that it creates frustration and tension. 24% of the students reported that it was associated with learned helplessness; they had done everything to get good grades, but never managed it. 37% of them said that it symbolized giving accounts to their parents for whom grades were the major indicators of success.

*The Role of Assessment*

All the students stated that assessment was a means of grading and determining their status among other students to be accepted to high schools or universities. In order to be placed in one of the prestigious high schools, secondary students had to answer multiple-choice questions testing the formal aspects of English and reading comprehension. Hence, assessment was regarded as a gateway to better high school education. The same role of assessment applied to the high school students planning to take the English component of the university entrance exam. For all the university prep students, it was the entryway to begin their academic programs in their departments. High school students had another serious concern about assessment. In Turkey, 60% of the high school GPAs is directly added to students' university entrance exam scores to calculate their final university entrance scores, which determines their percentile and standing as well as their acceptance to a university and department. Since English is one of the courses high school students must study for four years, it has a huge impact on their GPAs. Knowing this, 84% of the 481 students reported that they felt frustrated because there was a mismatch between their performance and their English scores, which definitely would affect their future careers.

*Frequency of Assessment*

Regarding the frequency of assessment, 69% of the total number of students said that there was no problem. However, 17% of the 197 university prep students reported that weekly quizzes were too much.

### *Means of Assessment*

In terms of the means of assessment, 56% of the students mentioned that they did not prefer the traditional pen-paper assessment, and 71% of the students said that they preferred projects.

### *The Type of Feedback the Students Received*

With respect to the type of feedback the students received, 83% of the students mentioned that they got feedback on their exam papers and their teachers explained their errors on the exam. Students accepted the answer keys and grades on their exam papers as feedback. Nothing was reported about the degree to which English objectives were realized; ultimately, no further training was planned, nor was any suggestion made to remediate or strengthen the weaknesses of students.

### *The Relevance between Assessment and Student Performance*

As for the relevance between assessment and student performance, 79% of the students mentioned that assessment did not reflect their performance, and 61% of them believed that the content of the exam was irrelevant to what they did in the classroom.

## **Observation**

### *Setting*

At the schools, the lessons were carried out according to the academic plan, which specified which units of the coursebooks were taught in which week. In other words, the academic plans were not organized around an overall goal and objectives, but the mathematical division of the total units in the coursebooks divided by the number of weeks in an academic year. Thus, the teachers stuck to the plan and tried to cover the units allotted for each week. Assessment-related activities were not observed until the week before the exam. However, one teacher from a private secondary school gave quizzes fortnightly. Two teachers from the public schools checked whether the units in the workbooks were done by the students and they were awarded pluses and minuses which counted as part of an oral grade at the end of the semester. In one private school, the department assigned the assessment tasks to each teacher from each grade two weeks before the exam date. Each teacher was responsible for a part and then they brought their parts together to form the exam. The other private schools went through the same process one week before the exam. Those schools either gave practice tests or worksheets focusing on the topics that would be covered on the exam. In the public schools, the exam preparation procedures were the same except for the practice tests and worksheets. Teachers informed the university students about the overall content of the exams before the exam week.

A different context was observed in the eighth grade and twelfth grade English classes in which the students and the teachers concentrated their ELT efforts on exam preparation. In other words, English classes turned into high school entrance and university entrance exam preparation classes. The topics tested on the entrance exam became the main interest: the exam preparation materials and practice tests replaced the ELT coursebooks.

*Steps of Assessment Procedures*

When preparing the exams, the teachers did not prepare any table of specifications guiding them about what to cover and in what weight. In addition, they did not pay attention to the basic principles of fair and valid assessment. They mainly focused on preparing tasks for the assigned parts of the exams. While preparing the texts and tasks, all the teachers at the schools exploited the assessment packs accompanying their coursebooks, tried to find texts and tasks from other coursebooks, or referred to the internet upon receiving their exam preparation date. Then the sections with the answer key prepared by each teacher were brought together, task scoring was determined and the process was finalized. The exams were printed and reproduced a day before the exam date or on the exam day. After administering the exams, the teachers got the exam papers of the classes they taught in order to score them. Normally, they announced the grades to the students the following week. For all schools, the same process was repeated for the second exams. In the week before the final week of the semester, the teachers gave at least one oral exam grade based on the checklists they kept, the participation of the students, and overall impression students made on the teachers. Five teachers from the private schools and two from the public schools gave two oral grades to increase the final grades of the students. Before the final week, the teachers entered the student grades into the software program and the administration announced them on the final day of classes.

At the university, the testing offices prepared the questions related to the tasks that would be asked on the proficiency exam. Thus, the exam served two purposes: achievement and proficiency preparation. However, there was no information on the criteria the testing office used while they prepared the exams. The testing offices worked in isolation. After the exams were administered, all the analyses were conducted by the testing office and announced to both the teachers and students online.

*Statistical Analysis*

Teachers at schools calculated the class averages and success percentages. They did not carry out any statistical analyses regarding the exam. At the universities, the assessment management software did all the statistical analysis on student performance and the quality of items. The analyses carried out on the exams by the testing office was not announced to the teachers and students.

*Feedback*

The teachers were observed to hurry to score the exam papers before the parent-teacher meetings, which were held once a semester in order to share the exam results with the parents. At the universities, the students learned their grades by logging into their accounts operated by student affairs. They had no chance to see their exam papers unless they officially applied to examine them.

*Assessment Tools*

The evaluation was based on pen and paper assessments. The knowledge of grammar and vocabulary and reading comprehension formed the basis of the exams. Listening and speaking were not tested, but writing was tested in all schools and universities for the students who were above the elementary level. Two of the universities also included listening on the proficiency exams. It was observed that the examinations focused only on formal aspects of English and declarative knowledge. The exams employed covered decontextualized items deprived of their functional use with no reference to where, when, how and for what purpose they were used for a



given context. They were not balanced with respect to the weight of linguistic elements to be tested. Comprehension questions were not reasonably balanced: They did not include referential/inferential questions and cognitive demand (lower/higher-order items) as described in Bloom's Taxonomy. Discrete items that required no or limited student production mainly dominated the exams. Thus, easy-to-score true/false, multiple-choice, fill-in-the-blank, and matching questions were commonly preferred on the exams.

#### *A Sample Exam Evaluation*

The exam was far from meeting the qualities of a well-formed, sound assessment tool (see Appendix 3). First, no information was provided about the exam, for example, which class it was prepared for, what topics or units it covered, the date of the exam, and the grading scheme. It was also poor in terms of face validity. The teacher himself/herself did not prepare the assessment items but found the theme from a different source and then cut them and pasted them on the exam paper. The exam paper looked like a collage, a mixture of typed and hand-written/drawn sections. The items on the exam were disoriented and disorganized. The exam started with a reading text with no instructions. The second task was chart reading. Both reading texts tested grammar and the usage of the simple present tense. Moreover, the learners might receive clues from the previous tasks that helped them to do task B. In part C, there was more than one possible answer. Moreover, there were unintentional clues from Text A again, such as "have lunch, watch a film". Part D was contextualized but only tested "am/is/are". The learners could easily find clues from other texts including "am/is/are". Task E tested two things at a time and lacked validity. The prerequisite task was to form the questions and then the next task was to match them with the appropriate answers. In addition, there were other possible answers to this task. Task F also tested two things at a time, both putting the sentences in the correct order and rewriting them considering the usage of the simple present tense. This task also provided unintentional clues for the other tasks and learners could find clues from other tasks to answer this item. Moreover, the grading was unfair, as the teacher awarded two points for each item. For example, a less demanding task, namely just writing "am, is, are" in Task D and comparatively more demanding tasks, such as creating the questions and matching them with the correct answers in Task E were given the same credit. Furthermore, the tasks were not organized from simple to complex. Task D was the easiest one and should have been the first task. In short, the exam employed a variety of tasks, but they were all created to test grammar, namely the correct usage of the simple present tense. To sum up, the test was not meaningful, valid, or fair.

#### **Discussion**

The findings obtained from the comparison of the groups indicated that the type of school and the level of education did not significantly affect how assessment was perceived by the EFL students in Turkey. Thus, no positive evidence was provided for the research questions 2 and 3. Generally speaking, it can be said that irrespective of their school type and level of education, all the Turkish EFL students had a common perception of assessment practiced in the Turkish EFL context.

Considering research question 1, the data obtained from the questionnaires, interviews, and observation corroborated one another and helped draw disciplined inferences about how assessment was perceived and practiced in the Turkish EFL context. Unlike the studies mentioned above, this study also directed attention to the perspective of the students. The data obtained from the Turkish EFL students indicated that the assessment practice in Turkey is problematic. The students were generally assessed twice a semester via a traditional pen and paper

exam, which cannot cover each and every English task they might encounter. Homework and to some extent officially assigned term projects were the other means of assessment. The frequency of assessment did not satisfy the students and they also believed that the content of the assessment was far from reflecting what was covered and achieved in the classroom. Consequently, they thought that the assessment they were administered did not measure their English performance. They were not pleased with assessment and believed that they were not assessed properly. They mainly indicated that multiple-choice items, and to some extent, short answer and essay tests were compatible with their language learning goals, as they had been and would be exposed to these types of items on the local exams offered by their institutions and on high-stakes exams administered by the government. Likewise, they found these kinds of items fairer. In addition to these items, they indicated alternative assessment as another means of demonstrating their real performance. They believed that the sole function of test results for their teachers was to give them grades. Regarding their roles in the assessment process, the students had no active involvement in the process and there was no place for self-assessment, peer-assessment, and co-assessment. In terms of the difficulty they experienced, the content of the questions, the allotted time, and the number of questions were three factors they had to cope with in the assessment process. In short, the students were not happy with how they were assessed.

In terms of conceptualization of testing, as strongly suggested by the data obtained from observation, the narrow definition, and function of assessment, an official procedure to test the performance of students at certain intervals to provide their grades was widespread. Assessment was utilized as a vehicle to assess student performance. Similarly, as illustrated by the analysis of the sample exam, traditional assessment techniques were implemented, which verified the findings of Tsagari and Vogt (2017). Pen and paper exams and homework were the main tools to test student performance. The exam was considered mainly as an official procedure, which was marked on the academic plan. The exams were not prepared considering the table of specifications and there was no reference to the fundamental issues regarding principles of assessment, how it was prepared, administered, evaluated and utilized to improve ELT. In terms of the focus of assessment, while administering the exams on time and grading students formed the center of gravity at the K-12 level, ELT and assessment practices were geared for the proficiency exam in the university prep classes. Hence, teaching for testing controlled the program, and the impacts of assessment were felt thoroughly. Both the teachers and students concentrated their efforts on getting the required grade on the proficiency exam to pass the prep class. As in the K-12 schools, the efficient and functional use of English in real-life did not receive the primary stress. This finding supported that of Ozdemir-Yilmazer and Ozkan (2017). In short, the research indicated that administering assessment in line with procedures and announcing the grades to the system received the primary stress in the Turkish EFL context.

The data obtained from the observation and the exam evaluation suggest that the ELT teachers either lacked the fundamental knowledge, skills, and basic concepts that govern assessment policy and practice or did not put what they knew into practice. Heritage (2007), Mede and Atay (2017), Scarino (2013), Tsagari and Vogt (2017), and Vogt and Tsagari (2014) also reported similar findings. The assessment practices adopted by the teachers were traditional and form-focused which were similar to the findings of Tsagari and Vogt (2017). Ultimately, they developed their assessment practice at schools. They tended to follow the assessment tradition and imitated their colleagues to conform to the group norms while carrying out the assessment practice.

## **Conclusion and Implications**

This wide-spectrum study pioneered and attempted to delineate a thorough picture of assessment to shed light on how it is conceptualized by the EFL students in Turkey by eliciting data from the students studying in different types of schools at different levels of education and universities.

The findings of the study demonstrated a clear dissatisfaction on the part of the Turkish students with the assessment in EFL classes per se. A considerable percentage of the students found the content of assessment unsatisfactory, which meant that they failed to see the relevance of the testing material to the actual language education that they had received. It can be argued that if the students had not seen the clear link between the assessment material and the educational material, the value of the assessment would have been called into question. The same trend is observable in the findings of the personal interviews. Because students did not see the relevance of the assessment material to the actual learning process, they did not relate positively to the testing processes. In other words, they did not see the tests as an integral component of their learning process but as an obstacle that they needed to overcome.

The study traced the roots of this dissatisfaction to two main causes: firstly, the lack or even the overall absence of the relevance of the assessment material and processes to the actual learning process in EFL classes that arose out of the mechanical assessment approaches traditionally employed in the Turkish language education setting and secondly, the inadequate assessment literacy of the teachers and test makers in designing and implementing fruitful assessment strategies and providing useful feedback to the students based on the students' performances. It was also intriguing to find that the type of the schools and the level of education did not significantly affect how assessment was perceived by the EFL students in Turkey. Based on the findings of the study, it can be seen that irrespective of the school type and level of education, all the Turkish EFL students shared a similar perception of assessment in the Turkish EFL context.

In terms of the overall assessment approach and practice, the study clearly showed that EFL assessment in the Turkish setting was almost exclusively directed at testing the formal properties of the English language, i.e. the knowledge of grammar, vocabulary, and reading comprehension, rather than the communicative and productive aspects of language such as listening and speaking. The data gathered in the study demonstrated how teachers, test makers, and students alike had developed a propensity for favoring easy-to-score and easy-to-analyze objective tests in the forms of multiple-choice questions or fill-in-the-blanks and how this tendency had, in turn, forced the assessment strategies away from employing a natural and holistic approach and instead, towards adopting a narrow and mechanical one. The propensity to see the multiple-choice question as a compatible testing method by the students indicated a sort of addiction on the part of students towards solving mechanical multiple choice questions, rather than engaging in a more natural, and integrated language practice. In short, they found it easier to choose an answer rather than produce language. This can be seen as proof of the failure of the assessment system in Turkey to push the language learners towards developing more fruitful language learning habits.

The consequences of employing this mechanical approach have been several-fold: First, the content of assessment has gradually moved away from reflecting what was covered or achieved in the learning environment, i.e. the classroom. As a result, the students had inadvertently developed the feeling that the assessment did not measure their actual performance or the level of their achievement but was merely there to issue a pass/fail grade as a part of the bureaucratic apparatus. Moreover, this narrow notion of assessment downgraded the role of students from active agents of assessment to passive participants. The total absence of interactive assessment strategies such as self-assessment, peer-assessment, and co-assessment, as observed in the study findings, was clear proof that students were considered only as passive participants in the process.

As a result of employing narrow, mechanical and inefficient assessment strategies, students had failed to see the relevance of any positive effect of the assessment system on their language learning progress but had been accustomed to seeing it as an official procedure of giving grades and determining pass/fail statuses. This notion of assessment as being independent of the language learning process had far-reaching repercussions in various EFL related areas of study,

which remain outside the scope of this current study. However, the findings here clearly corroborated the students' inability to relate to the assessment system as an indispensable component of a fruitful language education system and it focused their efforts on bypassing or avoiding the virtual hurdle that it created. A clear majority of the students asserted that the teachers used the assessment for the mere purpose of giving grades. The students used the exams for the very same purpose. This resonated with the earlier suggestion made above, that grades were the only thing students were trained to see on the exams. Consequently, they failed to see the relevance of the exams to actual language learning, and instead of seeing exams as an integral component of the language learning process, they considered them to be mechanical fail/pass hurdles that need to be bypassed or overcome.

These findings clearly pointed to the need for reconceptualizing assessment and designing a context-driven assessment plan that establishes assessment as an integral part of the language education system by aiming to develop and implement sound and valid assessment strategies. This reconceptualization which will doubtless have a positive effect on the students' perception of assessment can only be realized through teacher training in assessment. A careful program of teacher training in assessment with comprehensive coverage will lead to the implementation of informed assessment practices by the teachers, which in turn, will have a constructive effect on the overall perception of students of assessment. With the help of training programs, teachers can be expected to be competent and be resourceful enough to design and implement appropriate and adaptable assessments, fit to be tailored to the nature of their specific language program needs. Empowered with theoretical knowledge and practice in assessment, they will also know what to do with the data they obtain. Consequently, the teachers' and the students' awareness about the relationship between assessment and language education can be raised, which will undoubtedly have a positive impact on language education. Ultimately, valid and relevant assessment is likely to lead to efficient assessment and make students feel more satisfied with the assessment process.

Through training, teachers can redefine the roles of their students in the assessment process. As students are the active participants of the English education process, the data obtained from assessments are to be effectively exploited by students. In other words, the data must not be used only by teachers for the limited purpose of giving grades, but also to provide students with valuable data on the level of their achievements and probable shortcomings in the form of instructive feedback. The concept of feedback in Turkish language schools is reduced to giving correct answers to questions and clarifying what the students did wrong. Fruitful feedback must focus on pushing students toward a more productive use of language by providing the necessary training and further instruction, practice, or interactive communicative activities. To summarize, exams are not to be seen as pass/fail determiners but as markers of the students' linguistic achievements, and the feedback is elevated from mechanical grades and error-correction practices to organic instructions for the students that will help them improve their language skills.

With this interactive concept of assessment, students are granted a far more active role in assessment. If students are granted the responsibility and liberty to evaluate themselves, they may feel like part of the assessment process and benefit from it and develop a valid opinion about their English performance. Accordingly, they can redefine their roles in the assessment process. They can consider themselves not as the outsiders or the objects of assessment, but as active agents who have a voice in the process. As one of the agents of the assessment process, they can use it as a springboard to improve their English learning process. Such an active involvement may increase students' sense of agency in the assessment process and perceive it more positively.

Besides assigning an active role to the students, teachers' collaboration with their students to evaluate assessment data helps students reconceptualize assessment and examine their performance, which can improve their English learning process. This collaboration may help students establish a complementary relationship between learning and assessment. The

cooperative work that teachers carry out with their students also guides students on how to evaluate themselves and what to do with the data they obtain. Supporting them as assessors definitely creates a different and yet a more positive perspective of assessment.

Another issue teacher training can help teachers with is the value of feedback. If feedback is not limited to announcing the assessment results and pointing out the mistakes of students on their exam papers, a more comprehensive and fruitful bridge between assessment and English education and between teachers and students can be built. Continuous feedback turns assessment into an invaluable teaching and learning tool. Feedback obtained from self-evaluation and other stakeholders can create an ongoing cycle to feed English Education. Feedback which becomes an interface between learning and assessment is likely to make students redefine assessment and appreciate its value in English Education.

Another serious issue that needs to be considered is washback. Language education is directly affected by assessment. Students' study habits are shaped by how they are assessed, and teachers disrupt study by testing, especially in the case of high-stakes assessment. The assessment instruments focusing on limited aspects of English preserve traditional methodology, which equates learning a language to learning about its formal aspects. Students naturally study in relation to what is covered on exams in order to be successful and may ignore the meaningful, communicative activities, which are of limited or no help on form-focused exams. The teachers at the university level teach to the test and focus on language forms and short texts, thereby ignoring communicative tasks even if they may not want to do so. Thus, the traditional method focused on teaching about the language and aiming at the mastery of target language forms is nurtured by assessment, which leads to a vicious cycle. In short, the assessment tradition not only falls short of reflecting the performance of students and providing data to improve language programs, but also fostering an effective language education methodology. Hence, enhancing or upgrading assessment would definitely contribute to the quality of ELT.

Hopefully, the study can spark numerous other scholarly studies to gather rich amounts of data in order to make data-driven, disciplined decisions to improve the quality of assessment. Likewise, it is hoped that the study contributes to popularizing assessment in English among academicians and ELT teachers and attracts the attention of decision-makers. Such a focus on assessment, hopefully, ignites action to better assessment and English education in Turkey.

This study tends to have several limitations. The data gathering was confined to two major cities and thus, may not be taken as representative of all regions in Turkey. Consequently, it may fail to reflect how the EFL learners perceive assessment in general on a country-wide basis. Moreover, the perceptions of the Turkish ELT teachers were not included in the study. It would have been better to compare the perceptions of teachers and students in order to present a holistic picture of assessment in the Turkish EFL context. Thus, for further research, it is recommendable to expand the scope of the data gathering to include teachers and students from across the country.

On the other hand, a country-wise holistic analysis may fail to represent a detailed picture of the dominant viewpoint regarding assessment; thus, a schoolwise comparison is suggested.

## References

- Agrawal, M. (2004). Curricular reform in schools: the importance of evaluation. *Journal of Curriculum Studies*, 36(3), 361–379.

- Alderson, J. C., Brunfaut, T., & Harding, L. (2017). Bridging assessment and learning: A view from second and foreign language assessment. *Assessment in Education: Principles, Policy & Practice*, 24(3), 379-387.
- Bailey, K. M., & Brown, J. D. (1996). Language testing courses: What are they? In A. Cumming & R. Berwick (Eds.), *Validation in language testing* (pp. 236–256). Clevedon, UK: Multilingual Matters.
- Baker, B. (2016). Language assessment literacy as professional competence: The case of Canadian admissions decision makers. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée*, 19(1), 63-83.
- Barootchi, N., & Keshavaraz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests. *Educational Research*, 44(3), 279–288.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Longman.
- Brown, J. D. (2005). *Testing in language programs*. NY: McGraw-Hill.
- Brown, J. D., & Bailey, K. M. (2008). Language testing courses: what are they in 2007? *Language Testing*, 25(3), 349-383.
- Brunfaut, T. (2014). A lifetime of language testing: An interview with J. Charles Alderson. *Language Assessment Quarterly*, 11(1), 103-119.
- Combee, C., Folse, K., & Hubley, N. (2011). *A practical guide to assessing English language learners*. Michigan: Michigan University Press.
- Dann, R. (2014). Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice*, 21(2), 149-166.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43(3), 393-415.
- Douglas, D. (2010). *Understanding language testing*. London: Hodddler Education.
- Douglas, D. (2018). Introduction: An overview of assessment and teaching. *Iranian Journal of Language Teaching Research*, 6(3), 1-7.
- Earl, L. M. (2012). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.
- Farhady, H., & Tavassoli, K. (2018). Developing a language assessment knowledge test for EFL teachers: A data-driven approach. *Iranian Journal of Language Teaching Research*, 6(3), 79-94.
- Fox, J. (2015). Trends and issues in language assessment in Canada: A consideration of context. *Language Assessment Quarterly*, 12(1), 1-9.

- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132.
- Green, A. (2018). Assessment for learning in language education. *Iranian Journal of Language Teaching Research*, 6(3), 9-18.
- Harsch, C. (2014). General language proficiency revisited: Current and future issues, *Language Assessment Quarterly*, 11(2), 152-169.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). *European survey of language testing and assessment needs. Part one: General findings*. Retrieved January 23, 2019, from: [www.ealta.eu.org/resources](http://www.ealta.eu.org/resources)
- Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Pbi Delta Kappan*, 89, 140-145.
- Hill, K., & McNamara, T. (2011). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395-420.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402.
- Jiang, J. (2017). Language assessment: Critical issues—An interview with Antony John Kunnan. *Language Assessment Quarterly*, 14(1), 75-88.
- Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555-584.
- Johnson, R., Becker, P., & Olive, F. (1999). Teaching the second-language testing course through test development by teachers-in-training. *Teacher Education Quarterly*, 26(3), 71-82.
- Leung, C. (2004). Developing formative teacher assessment: knowledge, practice and change. *Language Assessment Quarterly*, 1(1), 19-41.
- Linfield, K. J., & Posavac, E. J. (2018). *Program evaluation: Methods and case studies*. New York, NY: Routledge.
- Little, D. (2005). The common European framework and the European language portfolio: Involving learners and their judgment in the assessment process. *Language Testing*, 22(3), 321-336.
- López Mendoza, A. A., & Bernal Arandia, 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.
- Luoma, S. E. (2002). What is happening in language testing? *Sprogforum*, 23(8), 7- 15.
- Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 351-372.

- Lynch, B., & Shaw, P. (2005). Portfolios, power and ethics. *TESOL Quarterly*, 39(2), 263–297.
- Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344.
- McNamara, T., & Roever, C. (2006). *Language testing: the social dimension*. Oxford: Blackwell Publishing.
- Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 43–60.
- Menken, K., Hudson, T., & Leung, C. (2014). Symposium: Language assessment in standards based education reform. *TESOL Quarterly*, 48(3), 586–614.
- Norris, J. M. (2016). Language program evaluation. *The Modern Language Journal*, 100(S1), 169–189.
- Ozdemir-Yilmazer, M., & Ozkan, Y. (2017). Classroom assessment practices of English language instructor. *Journal of Language and Linguistic Studies*, 13(2), 324–345.
- Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal*, 100, 191–2018.
- Rea–Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21, 249–258.
- Saricoban, A. (2011). A study on the English language teachers' preparation of tests. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 41(41), 398–410.
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309–327.
- Shohamy, E., Inbar-Louric, O., & Poehner, M. (2008). Investigating assessment perceptions and practices in the advanced foreign language classroom. *Report n, 1108*.
- Spolsky, B. (2008). Language testing at 25: maturity and responsibility? *Language Testing*, 25(3), 297–305.
- Stolz, S. A. (2017). Can educationally significant learning be assessed? *Educational Philosophy and Theory*, 49(4), 379–390.
- Tsagari, D. (Ed.). (2016). *Classroom-based assessment in L2 contexts*. Newcastle: Cambridge Scholars Publishing.
- Tsagari, D., & Banerjee, J. (Eds.). (2016). *Handbook of second language assessment* (Vol. 12). Boston: DeGruyter Mouton.
- Tsagari, D., & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 18–40.



- Tsagari, D., Vogt, K., Froelich, V., Csépes, I., Fekete, A., Green A., Hamp-Lyons, L., Sifakis, N., & Kordia, S. (2018). *Handbook of assessment for language teachers*. Nicosia: University of Nicosia Press.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402.
- William, D. (2011). Formative assessment: Definitions and relationships. *Studies in Educational Evaluation*, 37(1), 3-14.
- Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *TESOL Quarterly*, 43(3), 493-513.

## Appendices

### Appendix 1

#### Student Questionnaire

We would like to collect data about assessment-related issues in English language education. Please answer the questions below Thank you.

Your institution:  primary school  private primary school  secondary school  
 private secondary school  university  private university

1. What kind of assessment/tests do you take and how often? (you can mark more than one)

	quiz	exam	final
once a week			
every 2 weeks			
once a month			
once a semester			

2. What other means does your teacher employ to evaluate your performance?

3. Is the frequency of testing enough?

very satisfactory  satisfactory  partially satisfactory  unsatisfactory  very unsatisfactory

4. Does it cover what you study in the class?

very satisfactory  satisfactory  partially satisfactory  unsatisfactory  very unsatisfactory

5. Do tests measure your performance?

never  rarely  occasionally  often  always

6. Which of the following serves your language learning goals? (you can mark more than one)

essay  true-false  multiple choice  short answer  cloze  matching  other

7. Which are fair tests?

essay  true-false  multiple choice  short answer  cloze  matching

alternative assessment  other

8. Which helps you reflect your real performance? (you can mark more than one)

essay  true-false  multiple choice  short answer  cloze  matching

alternative assessment  other

9. How does your teacher use test results? (you can mark more than one)

- to give feedback about learning  
 to go over the problematic topics/skills  
 to give grades  
 to diagnose problems  
 other .....

10. How do you use test results? (you can mark more than one)

- to see my weaknesses and strengths  
 to review what I have not learned well  
 to learn my grades  
 other .....

11. Do you have a voice in the test preparation process?

never  rarely  occasionally  often  always

12. How frequently do you take part in the following assessment types?

	never	rarely	occasionally	often	always
self-assessment					
peer assessment					
co-assessment					

13. What causes difficulty for you on the exam? (you can mark more than one)

- content of the questions  time allotted  instructions  number of questions  
 sequence of questions  page layout of the exam paper  
 physical conditions (heat, light, noise, etc.) of the exam room  other .....

## Appendix 2

### **Guiding Questions for Observation**

What is the approach to assessment?

What is the aim of assessment?

What is the relationship between assessment and curriculum?

What is assessed?

What criteria are considered while preparing an exam?

When and how often is assessment carried out?

Who prepares the exams? Who are the agents of assessment?

How is assessment carried out? Which instruments are used in assessment?

How are the assessment results exploited?

## Appendix 3

Name: \_\_\_\_\_

Number: \_\_\_\_\_

Class: \_\_\_\_\_

**— Answer Key —**

**B**

**2008-2009 FIRST TERM FIRST WRITTEN EXAM**

When we are in space, we eat, sleep, wash, work and play just like on the earth. Every day is different, everyday is interesting and none of the days are typical.

I usually get up at six o'clock in the morning. I have breakfast with my friends. After breakfast, I do my regular exercise. Doing exercise is very important because in space we can't use our legs and it is very dangerous for us.

Then, I work on my research about space. I do some experiments. After we finish our work, we have lunch. We eat our meals very carefully because food can also float away.

In the evening, I sometimes watch films or play games with my friends in my free time. I always write e-mails to my friends and talk with my family on the radio.

At night, I usually sleep in a sleeping bag. Some of my friends like floating on air, but I hate it. I tie my sleeping bag to a wall so that I don't float during the night.

This is sometimes difficult but mostly enjoyable life. It is interesting to watch the world from the space.

**A) Read the passage again and answer the questions. (5x2=10pts)**

1. What does she do after breakfast?  
*After breakfast, she does her regular exercise.*
2. Why does she do exercise?  
*because in space they can't use their legs.*
3. What does she do after doing exercise?  
*She works on her research about space after doing exercise.*
4. What does she do in her free time?  
*she watches films or play games with her friends in her free time.*
5. How does she communicate with her family?  
*she talks with her family on the radio.*

**B) Answer the questions according to the information chart. (8x2=16 pts)**

1. Where does he live?  
*He lives in Manchester*
2. What does he do?  
*He is a musician*
3. How many children has he got?  
*He has got two children*
4. Which languages does he speak?  
*He speaks English and French*
5. Does he smoke?  
*He doesn't smoke*
6. Which day doesn't he work?  
*He doesn't work on Sunday*
7. What does he like doing?  
*He likes swimming and music*
8. What does he hate doing?  
*He hates football and walking*

Name	: Jack Morgan
Hometown	: Manchester
Work	: Musician
Marital status	: married
Children	: 2 girls
Job detail	: - play guitar - sing songs
Language	: English, French
Smoking	: X
Driving licence	: ✓
Free days	: Sunday
Likes	: music, swimming
Dislikes	: football, walking

c) Match the free time activities with the verbs (7x2=14pts)

do	for a walk	have	: ... a cup of tea, lunch with friends
have	a film	go	: shopping, for a walk
watch	lunch with friends	watch	: television, a film
go	shopping	do	: washing, yoga
	the washing		
	television		
	a cup of tea		
	Yoga		

d) Complete the paragraph with am, is, or are. (12x2=24pts)

My name is John Ponner. I am the manager of a big company. My company's name is Gardens. Gardens is a baby clothes company. Our workers are from different countries. I am from London. Alice Duell is one of my best friends in the company. She is from Paris. George Sipper and Jim Lucky are our new friends. They are from the United States. George is very young and energetic. I am very happy to work with him in Gardens.

e) Complete the questions with the verbs given; then match them with the answers (10x2=20)

go go do do help start have read come enjoy

1. What <u>do you do</u> ?	d.	a. No, not yet.
2. <u>Do you enjoy</u> your job?	e.	b. Usually by car.
3. What time <u>do you start</u> to work in the morning?	f.	c. She's a doctor.
4. How <u>do you go</u> to work?	b.	d. I work in a bank.
5. And your wife. What <u>does she do</u> ?	c.	e. Yes, I do.
6. Where <u>does she come</u> from?	d.	f. At half past eight.
7. <u>Does she read</u> a lot in her free time?	e.	g. Rarely.
8. <u>Do you help</u> your wife to cook?	f.	h. Yes, she likes reading.
9. <u>He is</u> often <u>do you go</u> out together?	g.	i. No, she doesn't let me cook.
10. <u>Do you have</u> a baby?	a.	j. She comes from Buenos Aires.

f) 1. Make sentences with these words and put the verb in the correct form: (8x2=16)

1. (never / I / stay up / late)	I never stay up late.
2. (the dusting / we / do / always / on Fridays)	We always do dusting on Fridays.
3. (my mother / get / sometimes / angry with us)	My mother sometimes gets angry with us.
4. (late / come / the bus / never)	The bus never comes late.
5. (my friends / go to / often / the cinema)	My friends often go to the cinema.
6. (our neighbours / are / usually / kind / to us)	Our neighbours are usually kind to us.
7. (sell / never / fruit / a baker)	A baker never sells fruit.
8. (practice / new words / often / she)	She often practices new words.

GOOD LUCK!

**Ali Işık** got his BA, MA, and Ph.D. degrees from the Foreign Language Education Program, Bogazici University (Istanbul, Turkey). He worked as a part-time lecturer at Bogazici University, METU, Hacettepe University, Yeditepe University, and Bogaziçi University consecutively. He is currently the Head of English Language and Literature Department and also serves as an ELT program and materials development advisor and teacher trainer at Istinye University, Istanbul, Turkey. His research interests include language teaching methodology, teacher training, content-based instruction, materials development, measurement and evaluation.