Exploring Language Assessment Knowledge of In-Service English Language Teachers in K-12 Context[1]

Gamze Satgun [1] (iD) **Asuman Aşık [2]** (iD)

[1] Ministry of Education, Türkiye, gamzesatgun@gmail.com

[2] Gazi University, Gazi Faculty of Education, Department of English Language Teaching, Ankara, Türkiye, asuman.asik@gazi.edu.tr

Article Info	ABSTRACT
	Language Assessment Literacy (LAL) is recognized as a crucial literacy skill that English language teachers should continually enhance. More research is needed to investigate the levels of LAL and Language Assessment Knowledge (LAK) of language teachers in a variety of different contexts. Notably, there is a scarcity of studies within the K-12 contexts in the existing literature. In response to this gap, the present study delves into the LAL levels of in-service English teachers within the K-12 educational context in Türkiye. The data were collected with the Language Assessment Knowledge Scale (LAKS) from 202 in-service English teachers working at primary, middle, and high school levels. The results of the analyses revealed that the teachers' LAK in general was notably high. The teachers' knowledge level in assessing reading and speaking was significantly high while they had significantly lower levels of knowledge in assessing writing and listening. Regarding the effect of the demographic factors on LAK, no significant difference was found. Based on the results, pedagogical implications are made for the teachers and policy makers.

Citation: Satgun, G. & Aşık, A. (2023). Exploring language assessment knowledge of in-service English language teachers in K-12 Context. *Journal of Teacher Education and Lifelong Learning*, *5*(2), 610- 627

---

[1] This article is derived and compiled from the master's thesis completed by the first author under the supervision of the second author.

INTRODUCTION

Language Assessment Literacy (LAL) is related to a mastery of essential components, which might comprise of knowledge about language, assessment, the context, and the ability to facilitate assessment-related procedures such as designing, collecting, administering, or interpreting assessment information, all crucial for making informed and ethical decisions (Inbar-Lourie, 2013). Despite the expanding body of research on LAL, there is a need for further studies to enhance our understanding of LAL across diverse language teaching contexts. Since the knowledge base of LAL is a dynamic facet, recent research findings can be utilized for the enhancement of teachers' language assessment knowledge (Coombe et al., 2020). Besides, the role of the context regarding LAL is undeniable. However, in some studies, it is assumed that the conceptualization of LAL applies to all educational levels, such as primary, secondary, and tertiary, neglecting the specific needs and aims of that context (Tsagari, 2020). Thus, a distinctive discussion of LAL in various educational levels is required to offer a contribution to LAL literature, which was also aimed at with this study.

Lam (2014) highlights that existing LAL research predominantly focuses on language instructors at the tertiary level or language testers. Consequently, there's a pressing need for future studies to explore the classroom-based assessment of in-service language teachers at primary and secondary levels and its impact on LAL development in K-12 contexts as well. To fill this gap, this study aims to contribute to the field by examining the EFL teachers' LAK level in assessing four skills (reading, writing, speaking, and listening) and the relationship with demographic features of the teachers, in the context of K-12 education in Türkiye. With these objectives in mind, this study aims to address the following research questions:

1. What is the overall and skill-based language assessment knowledge (LAK) level of EFL teachers in the K-12 context in Türkiye?

2. Does EFL teachers' level of LAK change in terms of gender, years of experience, the BA program being graduated, educational background, educational level of the workplace, taking a separate testing and assessment course in BA, attendance to additional training in testing and assessment?

3. Does EFL teachers' perceived self-competency in assessing each language skill have an effect on their LAK level?

LITERATURE REVIEW

Language Assessment Literacy

LAL is described as "the level of knowledge, skills, and understanding of assessment principles and practice that is increasingly required by other test stakeholder groups, depending on their needs and context" (Taylor, 2009, p. 24). Teachers hold the dual role of teaching and assessing (Inbar-Lourie, 2008; Rea-Dickins, 2004). Therefore, this role requires teachers to be more knowledgeable of assessment-related concerns, which helps them not only pursue more effective assessment procedures, but also evaluate their own instruction and construct appropriate assessments that motivate learners in their learning process (Rogier, 2014). High levels of LAL indicate a mutual benefit for both teachers and learners in that it provides feedback about both teaching and learning (Popham, 2009). To improve instructional quality and learner achievement in addition to determining appropriate methods and techniques of assessment for a particular purpose, teachers need to develop their LAL (Coombe et al., 2012).

In a similar vein, Herrera and Macías (2015) suggest that teachers should possess the capacity to correlate their assessment practices to approaches to language teaching, construct proper assessments,

choose the best assessments that suit the purpose, and be aware of the impact of large-scale examination if they have high LAL competence. Moreover, LAL will contribute to higher test validity and more transparent procedures (Coombe et al., 2012). Even though the significance of LAL for teachers is established in the literature, they are not believed to be qualified in assessment-related procedures (Plake, 1993; Popham, 2009; Stiggins, 1991). Thus, it is highly important for teachers to receive trainings about testing and assessment procedures for improving their knowledge of assessment and developing their LAL (Popham, 2011). However, Herrera and Macías (2015) discuss that pre-service trainings on assessment are not enough since they may be highly dependent on textbooks. Therefore, they advise that continuous professional development programs for both pre- and in-service teachers should embrace alternative approaches to enhance LAL of teachers.

Many researchers sought to identify the EFL teachers' needs regarding their LAL. First, focusing on LAK, Tavassoli and Farhady (2018) proposed that most of the teachers needed to enhance their LAK although they were aware of the essential components of assessment and had an average degree of perceived self-competency. Farhady and Tavassoli (2018) developed a knowledge test based on the needs analysis that they conducted (Tavassoli & Farhady, 2018) to reveal EFL teachers' actual LAK level as opposed to their perceptions about their knowledge of language assessment. The results demonstrated that LAK level of EFL teachers was insufficient despite their perceptions about their knowledge. Demographic factors of gender and type of their undergraduate degree, whether English language teaching, English literature, or others, had a possible impact on teachers' LAK level. However, the teaching context (state/private) and experience level influenced EFL teachers' LAK level in favor of the more experienced teachers and state institutions.

Fulcher (2012) found that the teachers were highly aware of their needs which are the skills, knowledge, and principles based on a procedural approach that balances the classroom and normative assessment within a wider social, historical, and ethical perspective. Additionally, Muhammad and Bardakçı (2019) aimed to demonstrate Iraqi EFL teachers' self-perceived LAL, their opinions about their pre-service education, and strengths and weaknesses in their knowledge regarding assessment. EFL teachers believed that their pre-service education trained them as good assessors and felt competent, contrary to their assessment literacy, which was found to be notably low. With a similar aim, Vogt and Tsagari (2014) identified the needs of EFL teachers in language testing and assessment. The teachers' LAL level was not satisfactory since no/limited training was received in this field. Besides, teachers needed trainings mostly for assessment purposes such as grading, placement of the students, and awarding certificates. Furthermore, they did not feel competent to review the quality of the assessment tools according to reliability and validity and to employ portfolio assessment, self- and peer-assessment in their practices.

The contextual and experiential aspects that contribute to the language assessment literacy have been emphasized by Yan et al. (2018). Their analysis regarding the needs and practices showed that the context of assessment, teachers' training experiences, assessment practices, and needs in knowledge and training contributed to EFL teachers' LAL development. Xu and Brown (2017) investigated Chinese EFL teachers' assessment literacy and its relationship with their experience, trainings, educational background, and gender. Thus, Chinese EFL teachers' assessment literacy was poor because of inadequate trainings in pre- and in-service education, not bearing a grounded criterion for assessment literacy, and lack of quality standards for assessment practices.

It is also seen that research on LAL in Türkiye has lately become a subject of interest since most of the studies have been conducted in recent years. In the studies, knowledge (Genç et al., 2020; Ölmezer-Öztürk & Aydın, 2019), perceptions and practices (Arslan & Üçok-Atasoy, 2020; Büyükkarcı, 2014; Işık, 2021; Kırkgöz et al., 2017; Önalan & Gürsoy, 2020; Öz & Atay, 2017; Tanyer & Susoy, 2018), training needs (Ballıdağ & İnan-Karagül, 2021; Mede & Atay, 2017), and problems (Büyükkarcı, 2014; Tuzcu-Eken, 2016) as well as the impact of demographic and contextual factors

(Büyükkarcı, 2016; İnan-Karagül et al., 2017) on language assessment have been analyzed. Higher educational context was mainly preferred by the researchers in Türkiye whereas the research regarding K-12 context is limited. Focusing on perceptions and practices in primary education context, Kırkgöz et al. (2017) revealed that EFL teachers considered reading and speaking as important skills to be assessed. Also, they believed that listening can be ignored in the assessment process while writing is considered somewhat more important. In a similar context, Tanyer and Susoy (2018) explored pre-service and in-service EFL teachers' perceptions and practices regarding the assessment of young learners. It was demonstrated that both pre- and in-service teachers had insufficient knowledge and skills in parallel with their perceptions since they did not feel competent and believed that assessing young learners is a challenging procedure to sustain. On the contrary, in another study with a wider context including primary, secondary, and higher educational context, Işık (2021) found that EFL teachers were confident with their current status even though they indicated low levels of LAL based on data obtained from a questionnaire, interviews, classroom observations, and sample exam evaluations. Concerned with the K-12 context, Ballıdağ and İnan-Karagül (2021) put forth that teachers needed further training due to their insufficient LAL. Finally, aiming to determine EFL teachers' knowledge level in assessing writing and speaking in the K-12 context, the findings of a study by Genç et al. (2020) showed that teachers had significantly lower scores in assessing writing while they had higher knowledge level when it comes to assessment of speaking skills. Additionally, no impact was observed on their LAK in terms of their level of experience, type of the BA program and taking additional trainings. As listed, previous studies mainly focused on self-reported perceptions of EFL teachers with a limited focus on language skills. Thus, the significance of the current study is that LAK levels were analyzed with knowledged-based investigation instead of self-reported perceptions. Additionally, the current study focused on skill-based LAK by including the participants from a variety of K-12 levels (primary, secondary, and high school contexts).

METHOD

Research Design

The aim of the study is to investigate in-service EFL teachers' overall and skill-based LAK level and whether their LAK changes based on their demographic characteristics. Accordingly, a correlational research design has been employed as a quantitative approach to identify potential relationships among two or more variables without any manipulation. (Fraenkel et al., 2012).

Participants and Research Context

In Türkiye, English is instructed as a foreign language. Within the K-12 educational framework, a primary focus of this study, English teachers employed in these institutions bear the responsibility for testing and assessment procedures. However, students in both lower and upper-secondary levels are required to undertake a standardized test, administered by the Ministry of National Education (MoNE), at the conclusion of each educational stage.

The teachers who work at the primary, middle, and high schools in Türkiye, whether private or state, in the 2021-2022 educational year constitute the universe of this research. After obtaining the necessary permissions from the MoNE and ethical committee approval from a state university in Türkiye, the scale is sent to EFL teachers all over the country via MoNE's electronic document management system. Among the whole population of EFL teachers, 202 participants took part in the scale on a voluntary basis. Demographic information of the participants is presented in Table 1 below.

Table 1. *Demographic Features of the Participants in the LAKS*

Demographic Features		Number	Percentage
Gender	Male	48	23.8

	Female	154	76.2
Years of experience	1-5 years	91	45.0
	6-10 years	34	16.8
	11-15 years	23	11.4
	16-20 years	27	13.4
	More than 20 years	27	13.4
BA program	ELT	181	89.6
	Non-ELT	21	10.4
Educational background	BA degree	161	79.7
	MA degree	37	18.3
	PhD Degree	4	20
Educational level of the institution	Primary school	35	17.3
	Secondary school	103	51.0
	High school	64	31.7
Type of the institution	State/Public	187	92.6
	Private	15	7.4
Having testing/assessment course in pre-service	Yes	113	55.9
	No	89	44.1
Attending professional development programs/courses on testing/assessment	Yes	105	52.0
	No	97	48.0

Table 1 shows that 75% of the participants consist of female teachers. Teachers with 1-5 years of experience constitute half of the participants, while the others have an approximately equal distribution. 90% of the participants graduated from the Department of English Language Teaching. Among 202 participants, only 4 of them had a PhD degree, while the number of participants with an MA degree is 37, and with BA degree is 161, which is 80% of the whole participants in the study. While almost all the teachers work at a state school, nearly half of the teachers work at the middle schools followed by high school with 30% and primary schools with 17%. Finally, teachers' pre- and in-service trainings regarding language assessment are analyzed. It is seen that more than half of the teachers have received a sort of assessment course in pre-service education. However, it can be deducted from the table that half of the participants did not attend any in-service training for their professional development.

Research Instrument and Processes

The data were collected with the Language Assessment Knowledge Scale (LAKS), developed by Ölmezer-Öztürk and Aydın (2018). The scale consists of two parts, and the first part involves questions about the demographic features of the participants. In the second part, there are four factors that aim to measure EFL teachers' language assessment knowledge level in assessing reading, listening, writing, and speaking. In the scale, consisting of 15 items in each factor and 60 items in total, the participants are asked to read the statements and select the most appropriate option for each statement among True/False/Don't Know options. The scale was found valid and reliable by Ölmezer-Öztürk and Aydın (2018). The validity was provided by 11 expert opinions and 18 practitioners. The scale has been developed after a thorough literature review process and consists of elements required for fundamental knowledge of language assessment (Ölmezer-Öztürk and Aydın, 2018). The reliability analysis of the scale has been completed by the researchers who developed the scale, and the reliability of the scale has been measured as .91 for the Cronbach Alpha coefficient. LAKS is prepared in an online format and sent to every EFL teacher throughout the country via the electronic document management system of MoNE. Ethics committee approval was obtained before the data collection from a state university ethics committee (16.02.2021, 2021-207). Additionally, necessary permissions were taken from the Ministry

of Education of Türkiye.

Data Analysis

The data collected with the LAKS were analyzed with descriptive and inferential statistics of IBM SPSS Statistics 22 program. The participants' answer to each statement was scored "1" point if their answer was correct, and "0" point if they selected the wrong answer or "Don't know" option. Consequently, the highest score that can be obtained from the scale is "60", consisting of "15" points for each factor, and the lowest score is "0". Firstly, the normality test was carried out, and the results indicated that the data were normally distributed. Therefore, parametric tests were conducted. To estimate reliability of the quantitative data, Cronbach alpha coefficient was determined as 0.58. Reliability rate between 0.50 and 0.70 indicates moderate reliability (Hinton et al., 2014). The participants' LAK level in general and skill-based was analyzed with one-sample t-tests. The impact of the demographic characteristics of the participants on their LAK level was analyzed with independent samples t-tests and one-way ANOVA and presented with inferential statistics. For the third research question of the study, which aims to find out whether there is a difference between the participants' perceived- self-competency in assessing language skills and their actual LAK level, Pearson Correlation and one-way ANOVA analyses were employed.

Table 2. *Reliability Analysis of the Scale and Sub-constructs*

Constructs	Cronbach Alpha
Reading	0.22
Listening	0.33
Writing	0.33
Speaking	0.43
LAKS	0.58

FINDINGS

EFL Teachers' General and Skill-based Language Assessment Knowledge Level

The findings in Table 3 (See Appendix) show that they had a mean score of 32.099 over 60, indicating that their general LAK level is higher than half of the total score. The highest score that can be obtained from the scale is "60" while the lowest is "0". Therefore, to decide whether the participants have low or high LAK level, the point of reference for the study is determined as the half of the total score, which is "30". The same reference point was utilized in line with the research which developed the scale (Ölmezer-Öztürk & Aydın, 2018). To explore whether the participants have significantly high level of LAK, one sample t-test is conducted. When the point of reference (30) is compared to the participants' mean scores, a difference of 2.099 is observed between the participants' mean scores and the point of reference. The results indicate that the teachers' general LAK level is significantly high in K-12 context.

Table 4. *One Sample T-test Results for General LAK Level*

	N	X̄	Mean diff.	df	t	P
General LAK	202	32.099	2.099	201	5.485	.000*

*p<.05; Test value=30

To analyze EFL teachers' mean scores in each language skill, one sample t-test is applied. On the

scale, the total score that can be obtained from each language skill is "15" while the lowest is "0". Thus, the test value is determined as half of the total score, which is "7.5". The participants' mean scores in each skill are X̄=9.5693 (reading), X̄=6.8317 (listening), X̄=7.0594 (writing), and X̄=8.6386 (speaking). It is observed that in assessing reading and speaking skills, the participants' mean scores are higher than the reference point (7.5), whereas in assessing listening and writing skills, their mean score seems to be lower than 7.5. To determine whether the difference between the participants' mean scores and test value (7.5) is significantly higher or lower, one sample t-test results are shown in the table below.

Table 5. *One Sample T-test Results for Skill-based LAK Level*

	N	X̄	Mean diff.	df.	t	P
Reading	202	9.5693	2.0693	201	14.821	.000*
Listening	202	6.8317	-.6683	201	-4.444	.000*
Writing	202	7.0594	-.4405	201	-2.900	.004*
Speaking	202	8.6386	1.1386	201	7.154	.000*

*p<.05; Test value=7.5

The results reveal that the participants' skill-based LAK level in assessing reading, the language skill that the participants displayed the highest knowledge, was significantly higher than the test value (7.5). However, the participants demonstrated the lowest knowledge level in assessing listening, and the mean difference between the test value (7.5) and the participants' mean scores (X̄=6.8317) is found to be statistically significant. As a result, it is concluded that the participants' LAK in assessing listening is insufficient. Another skill that has received a low mean score is assessing writing. Even though there was not a significant difference between the participants' mean scores (X̄=7.0594) and the test value (7.5), the one sample t-test results indicate that the participants' knowledge level in assessing writing is significantly low. The final skill, assessing speaking, has received a high knowledge level. Based on the table above, it is shown that the participants' mean scores in assessing speaking (X̄=8.6386) are significantly high when compared to the test value (7.5). When the correlation between the language skills is analyzed, the results indicate both based on the Pearson correlation analysis.

Table 6. *The Correlation Among the Skill-based LAK*

	LAK	Reading	Listening	Writing	Speaking
LAK	1	.594**	.673**	.712**	.568**
Reading		1	.246**	.241**	.089
Listening			1	.369**	.106
Writing				1	.197**
Speaking					1

** Correlation is significant at the 0.01 level (2-tailed); N=202

The findings revealed that EFL teachers have high LAK in general, in assessing reading and speaking, whereas low LAK is indicated for assessing writing and listening. To be able to make an in-depth discussion of these findings, the participants' correct and incorrect answers to each item on the scale are investigated and discussed. The table that presents the descriptive statistics of the LAK level of EFL teachers in the K-12 context in Türkiye is provided in the appendix.

Regarding assessing reading, among the items between 1 to 15, the highest scored item is 13, "Reading texts in a reading exam include various genres (essay, article, etc.)". Eighty seven percent of the participants answered this item correctly. The item which is the one that the teachers had the lowest mean score on the reading assessment is "When asking several questions about a reading text, all questions are independent of each other". Only 59 teachers (29%), by selecting the true option, have

answered this item correctly, which is quite low. When the number of the false (n=136) and do not know (n=7) options are compared, it is observed that nearly all the teachers had an idea about the characteristics of the questions about a reading exam regarding their relevance to each other. It is observed that 70% of the teachers have low level of knowledge about preparing sound reading questions.

The findings display that participants were mostly knowledgeable in these areas: utilizing various types of genres, summarization and top-down approach, the use of true/false items and cloze tests, the characteristics of multiple-choice questions as in grammatical correctness, not including the same word in the correct option, simplifying the language of the texts and questions, using more items and texts to gather more samples, hence increasing reliability. On the other hand, they were not informed about spelling in the scoring criteria and using a text that students have seen before. Teachers were mostly not aware of the fact that the questions in a reading exam should be independent of each other. In conclusion, it can be uttered that EFL teachers have high knowledge level in assessing reading considering that they have high mean scores in 12 among 15 items on the scale.

The findings related to assessing listening show that the highest mean score belongs to the item "In selective listening, learners are expected to look for certain information". Unlike extensive listening, acquiring a global understanding of the text is not the main concern. In total, 178 teachers (88%) provided a correct answer to this item. When the numbers of the participants with incorrect answers (n=12) and the participants who did not have an idea about the item (n=12) are examined, it is seen that they are the same. The mean scores of the teachers indicate that they were quite informed about the purposes of selective listening tasks. The item with the lowest mean score among the 15 items in this part, "Dictation is a kind of discrete-point testing". The number of the participants who gave a correct answer for this item is 36 (18%), while 31 (15%) did not have an idea about whether the item was true or false. Besides, 135 of the teachers (67%) have chosen the true option, giving a wrong answer. It is clearly observable that teachers have a significantly low level of knowledge in discrete-point testing techniques regarding dictation. Thus, EFL teachers in the K-12 context were mostly knowledgeable in selective listening, using listening cloze-tests for selective listening, characteristics of the listening text in terms of background knowledge, redundancy and authenticity involved. On the other hand, the teachers' knowledge level in assessing listening remained low regarding discrete-point testing, scoring, and purposes of discrete-point testing, intensive and integrative assessment approaches, scoring criteria whether to include spelling and grammar, and selection of listening text.

Furthermore, the participants' mean scores ($\bar{X}$=7,0594) related to assessing writing demonstrate that EFL teachers' knowledge level is lower than half of the total score, which puts assessing writing in third place. The mean scores obtained from the items, based on the participants' answers, are individually revealed, and discussed in this part from the highest to lowest. The item "Analytic scoring is used to see the strengths and weaknesses of learners" have the highest mean score in assessing writing part of the scale. The lowest mean score in this part is on the item "When there is a disagreement between the scores of the two raters, they score the written work again". The participants' answers to this item were interesting in that the number of teachers who gave a correct answer (n=32) was the same as those who do not know whether this item was true or false (n=31). Besides, 139 teachers (69%) believed that the raters should score the writings. It is highly clear that teachers are not knowledgeable in the scoring procedures of writing assessment when a discrepancy occurs.

All in all, the EFL teachers in the K-12 context were mostly aware of the analytic and holistic scoring in terms of their advantages and reliability, prompts that can be used such as visuals or a reading text, when to focus on mechanical errors, whether to adjust the scales based on the learners' needs or not. However, they were not knowledgeable enough in controlled writing, when to deal with irrelevant ideas, the number of writing tasks involved, using opinion-based prompts, giving options to learners, and scoring procedures when two raters are involved.

Lastly, it is observed that compared to reading, listening, and writing, speaking is the second highest regarding the participants' mean scores (X̄=8.6386), which is higher than half of the total score that can be obtained in this part. The highest mean score in assessing speaking part belongs to "Discussion among learners is a way of assessing speaking skills". The last item, which also got the lowest mean score, is "In peer interaction, random matching is avoided". Thus, EFL teachers in the K-12 context in Türkiye generally have sufficient knowledge in using role-plays to assess discourse and discussion tasks for speaking, gathering more samples of learners' performances by using more than one task, using checklists for grading, speaking constructs in terms of assessing comprehension and production together, and the role of the interlocutor concerning adapting the interview questions and showing interest during the interview. On the other hand, their knowledge level remained below the average on these issues: using repetition and reading aloud as a speaking task, the role of the interlocutor in terms of when to score, when to end the interview, and whether to express if they do not understand the learner or not, how many learners to involve and how to match them in interactive tasks, and finally, the advantages of using holistic and analytic scales together.

Effects of Demographic Features on EFL Teachers' LAK Level

The relationship of the demographic characteristics on EFL teachers' LAK level is also investigated. Independent samples t-test results, given in Table 7 below, show that there is no statistically significant difference in the LAK level of teachers in terms of gender, the BA program being graduated (ELT, non-ELT), the type of institution that the participants work(state/private), the testing course taken in BA or additional trainings on testing.

Table 7. *Independent Samples T-test Results According to Gender, BA Program, Type of Institution, Testing Course in BA, Additional Trainings and LAK*

	Categories	N	X̄	S	Sd	t	P
Gender	Female	154	32.06	5.01	200	-.159	.874
	Male	48	32.20	6.66			
BA Program	ELT	181	32.07	5.60	200	-.208	.835
	Non-ELT	21	32.33	3.77			
Type of Institution	Public/State	187	32.23	5.35	200	1.259	.209
	Private	15	30.40	6.29			
Testing Course in BA	Yes	113	32.41	4.90	200	.933	.352
	No	89	31.69	6.05			
Additional Trainings	Yes	105	32.77	5.73	200	1.839	.067
	No	97	31.37	5.02			

Furthermore, regarding the participants' level of experience, the educational background, and the educational level of the workplace, one-way ANOVA is conducted, and no significant difference is found between the participant groups. The findings are presented in Table 8 below.

Table 8. *One-way ANOVA Results According to the Years of Experience, Educational Background, Educational Level of the Workplace and LAK*

	Categories	N	X̄	S	Source of Variance	Sum of Squares	df	Mean Square	F	p
Years of	1-5 years	91	32.50	6.14	Between	138.673	4	34.668	1.176	.323

					Groups					
experience	6-10 years	34	31.52	5.24	Within Groups	5807.347	197	29.479		
	11-15 years	23	30.30	4.97	Total	5946.020	201			
	16-20 years	27	31.81	3.92						
	More than 20 years	27	33.25	4.62						
Educational Background	BA degree	161	32.15	5.56	Between Groups	8.152	2	4.076	.137	.872
	MA degree	37	32.00	5.16	Within Groups	5937.868	199	29.839		
	PhD degree	4	30.75	2.21	Total	5946.020	201			
Educational level of the workplace	Primary	35	31.45	4.17	Between Groups	23.971	2	11.986	.403	.669
	Secondary	103	32.07	5.45	Within Groups	5922.049	199	29.759		
	High	64	32.48	6.03	Total	5946.02	201			

EFL Teachers' Perceived Self-competency and Actual LAK Level

To investigate whether the teachers' LAK level changed according to their perceived self-competency, the teachers were asked to select their level of competency in the scale for each language skills consisting of "very competent", "competent", "not very competent", and "not competent" options. Therefore, teachers' perceptions of their knowledge in assessing each skill have been investigated with Pearson Correlation and one-way ANOVA analyses, as shown in Table 9 below.

Table 9. *Pearson Correlation Results of the Relationship Between Perceived Self-competency and LAK*

Variable		LAK of Reading	LAK of Listening	LAK of Writing	LAK of Speaking
Perceived Self-competency of Reading	r	-0,65			
	p	,357			
	N	202			
Perceived Self-competency of Listening	r		-0,39		
	p		,585		
	N		202		
Perceived Self-competency of Writing	r			-,136	
	p			0,53	
	N			202	
Perceived Self-competency of Speaking	r				-.199
	P				,004*
	N				202

p<.05

From Table 9, it is evident that while teachers' perceptions of their knowledge negatively correlate with their actual LAK levels in assessing language skills, this correlation is only statistically

significant for speaking. A closer examination reveals a moderate negative correlation with a value of r=-.199. One-way ANOVA analysis is used to determine whether there is a significant difference among the perceived self-competency groups (very competent, competent, not very competent, not competent) regarding their LAK level in each skill.

Table 10. *One-way ANOVA Results of Perceived Self-competency and LAK in each Language Skill*

		N	X̄	S	Source of Variance	Sum of Squares	Df	Mean Square	F	P
Reading	Very competent	110	9.74	1.88	Between Groups	12.639	2	6.320	1.615	.202
	Competent	84	9.28	2.03	Within Groups	778.891	199	3.914		
	Not very competent	8	10.12	2.64	Total	791.530	201			
Listening	Very competent	55	7.00	2.49	Between Groups	5.748	3	1.914	.416	.742
	Competent	112	6.75	1.79	Within Groups	912.529	198	4.609		
	Not very competent	34	6.88	2.57	Total	918.277	201			
	Not competent	1	5.00							
Writing	Very competent	70	7.40	2.10	Between Groups	21.211	3	7.070	1.528	.208
	Competent	108	6.99	2.15	Within Groups	916.076	198	4.627		
	Not very competent	21	6.28	2.32	Total	937.287	201			
	Not competent	3	7.00	1.00						
Speaking	Very competent	49	9.40	2.32	Between Groups	73.453	3	24.484	5.075	0.002*
	Competent	122	8.54	2.09	Within Groups	955.166	198	4.824		
	Not very competent	29	7.58	2.41	Total	1028.618	201			
	Not competent	2	11.00	0						

p<.05

As shown in Table 10, teachers reported that they are competent in their knowledge of reading assessment. However, teachers who are not very competent in assessing reading performed the highest in the scale. When the participants' mean scores in assessing reading (X̄=9.5693) is considered, which is significantly high, it is possible to say that their perceptions and actual LAK level are in line with each other since assessing reading has received the highest mean score among all language skills. Besides, there is no significant difference between the competency levels of the teachers regarding their LAK mean scores. Thus, it can be suggested that the teachers are generally aware of their knowledge level in assessing reading.

In terms of assessing listening, it is seen that most of the teachers felt competent (N=112) and very competent (N=55) regarding their knowledge in assessing listening, who also have the highest mean scores compared to the teachers who selected the options not very competent (N=34) and not competent (N=1). Nevertheless, the mean scores that the participants had in assessing listening part on the scale (X̄=6.8317) was significantly lower than the half of the total score (7.5) in addition to being the lowest among other skills. Additionally, the difference between the participants' LAK mean scores based on their level of competency is not statistically significant. Therefore, it can be inferred that the

participants' perceptions about their knowledge may coincide with their actual knowledge level in assessing listening.

Pertaining to assessing writing, it is demonstrated that the teachers have a high level of self-competency since 88% of the teachers stated that they are either competent (N=108) or very competent (N=70). Besides, the participants who are very competent received the highest mean score in assessing writing, while the mean score of the participants in assessing writing ($\bar{X}$=7,0594) is found to be significantly low. Consequently, it might be suggested that their perceptions of their knowledge mirror their actual LAK in writing assessment since no significant difference is observed between the participants' level of competence and their LAK mean scores.

As for assessing speaking, the teachers' competency level is found to be considerably high as the number of the participants who selected the option very competent (N=49) and competent (N=122) constitutes 84% of the total participants. Even though their mean scores in assessing speaking in the scale ($\bar{X}$=8.6386) is significantly high, a significant difference is observed between the competency groups regarding their actual LAK and perceived self-competency. To investigate more into the differences between all competency groups, Tukeys' Test is conducted as a post hoc test. Even though the not competent group has the highest score compared to others ($\bar{X}$=11.00), the results indicate a significant difference between the groups of not very competent ($\bar{X}$=7.58) and very competent ($\bar{X}$=9.40) in favor of the very competent group. The reason behind this can be explained by the small number of the participants in the not competent group (N=2).

DISCUSSION, CONCLUSION, RECOMMENDATIONS

This study investigated the LAL of the EFL teachers in their knowledge base in the K-12 context of Türkiye. The results indicated that EFL teachers had sufficient knowledge in language assessment supporting the findings of other studies in the higher education context (Hakim, 2015; Jannati, 2015). Related to each language skill, the findings illustrated that the teachers had sufficient knowledge respectively in assessing reading and speaking, reading being the highest, and insufficient knowledge in assessing writing and listening, listening being the lowest. These results are in line with Ölmezer-Öztürk and Aydın (2019). Similarly, Genç et al. (2020) investigated the LAK level of high school EFL teachers in assessing writing and speaking and found that teachers were knowledgeable in assessing speaking, whereas their knowledge in assessing writing was significantly low. Thus, it can be interpreted that the EFL teachers in the primary, middle, and high schools in Türkiye are more knowledgeable in reading and speaking assessment, while their knowledge is in writing and listening assessment even though they demonstrated a higher degree of LAK in general. A plausible explanation might be the teachers' perceived emphasis on certain skills during the assessment process. Kırkgöz et al. (2017) highlight that while teachers prioritize speaking and reading skills, they often deemphasize listening in their evaluations. In addition, only a minority reported that writing is a highly important skill. Consequently, it is possible to state that there is a relation between the teachers' perceptions of assessment and their LAK. However, the direction of this relation needs to be analyzed with further investigation.

When the impact of the teachers' demographic characteristics is examined, it is revealed that no significant effect is found on the teachers' LAK according to their gender, level of experience, the BA program being graduated, whether having a BA, MA, and PhD degree, working at a private or a state school, the educational level of the workplace whether primary, middle, or high school, taking a language assessment course in pre-service education, and attending professional development programs. These findings support the results of other studies regarding gender, years of experience, educational

background in different educational contexts (Büyükkarcı, 2016; Jannati, 2015; Öz & Atay, 2017; Xu & Brown, 2017). On the other hand, a contradiction is displayed in a study in terms of the effect of the type of the institution being private or state and experience level by Farhady and Tavassoli (2018). It was found that teachers with more experience and work at a state institution had higher levels of knowledge in language assessment. The reason behind this difference in findings can be related to the number of the participants in this study, in which only a minority of the teachers worked at a private institution. Consequently, it can be stated that the EFL teachers' LAK is possibly affected by the type of their institution. In conclusion, the contextual and experiential factors are not influential in shaping the EFL teachers' LAK in the K-12 context of Türkiye.

In the literature, a discrepancy is observed between the teachers' perceptions and their LAL-LAK (Farhady & Tavassoli, 2018; Işık, 2021). Regarding the competency level of the teachers as they perceive themselves to be in assessing each language skill and their actual LAK, the results indicated that, in general, teachers' perceptions are reflected in their knowledge in that they generally felt competent and indicated high LAK level. However, when the language skills are focused individually, it is shown that the teachers' perceived self-competency in assessing reading, writing, and listening is in accordance with their actual knowledge level, while a discrepancy is observed between their self-competency and their actual knowledge level in that the teachers with lower self-competency level received higher LAK levels in assessing speaking.

LAL typically comprises both theory and practice. In other words, teachers need to have a degree of knowledge regarding language assessment and the skills to reflect their knowledge in their assessment practices (Inbar-Lourie, 2008). Besides, the knowledge base of LAL is considered a prerequisite for teachers to mirror their knowledge of theoretical considerations into their classroom practices appropriately to be able to make proper judgements about the learners (Fulcher, 2012; Xu & Brown, 2016). In line with these considerations and the abovementioned conclusions, teachers should improve their knowledge about assessing listening and writing since lower levels of knowledge may result in inappropriate assessment practices hence shadowing the quality of foreign language education. A possible suggestion that can be made in this regard is that in-service trainings can be provided for the teachers on assessing language skills individually, namely writing and particularly listening, rather than broadly touching upon general language assessment theory with a focus on specific areas of assessing language skills such as scoring procedures, the use of scoring scales, alternative methods of assessment.

Based upon the results, pedagogical implications are drawn upon the problematic areas in teachers' LAL based on their knowledge. Due to inadequate teaching hours and large class sizes, it may be interpreted that reading, grammar, and vocabulary skills are typically assessed with written exams as a formal procedure while speaking, writing, and listening are assessed informally based on the classroom observations without incorporating a scoring scale, which may lead to inappropriate judgements about the learners' performances and questioning of the reliability of the assessment without scoring criteria. Therefore, in order to make sound interpretations of the learners' performances, teachers need to adopt more reliable procedures by utilizing checklists and rubrics. Additionally, a further suggestion on assessing speaking skill is that the institutions can support the use of technological applications and tools where learners' speaking performances are assessed, so that more practical applications in terms of time can be included in the assessment process. Thus, in-service trainings can be implemented in technology integrated assessment and the use of proper technological tools using automated feedback tools or artificial intelligence tools. Besides, MoNE can provide these technological applications for teachers to utilize in their classrooms so that speaking skills are incorporated into classroom activities more systematically to collect assessment data on learners' speaking skills.

Another implication can be made upon the listening skills, in which teachers' knowledge and practices was insufficient though indicated high perceptions of their knowledge and practices. Therefore, teachers need to develop their theoretical knowledge in assessing listening and their practices

accordingly. As a result, it is crucial that in-service trainings are provided for the teachers with a focus on listening assessment. It can be suggested that the content of the trainings should focus on a specific area of language assessment rather than in general such as giving feedback, scoring procedures and use of scoring scales, alternative assessment methods, technology-integrated assessment, and self- and peer-assessment that incorporates more applicable and practical concerns.

The study has some limitations. Data collection tools for LAL are limited to EFL teachers' knowledge base and classroom practices. Since the number of the participants in the scale is limited to 202 EFL teachers, the results may not be generalized to the larger population. Thus, further research can be performed with more participants to obtain more reliable results regarding EFL teachers' LAK. Also, to be able to make more in-depth discussion of the teachers' assessment practices, further research can incorporate classroom observations as another data collection tool.

REFERENCES

Arslan, R. Ş., & Üçok-Atasoy, M. (2020). An investigation into EFL teachers' assessment of young learners of English: Does practice match the policy? *International Online Journal of Education and Teaching (IOJET), 7*(2), 468-484.

Ballıdağ, S., & İnan-Karagül, B. (2021). Exploring the language assessment literacy of Turkish in-service EFL teachers. *Balıkesir University the Journal of Social Sciences Institute, 24*(45), 73-92. https://doi.org/10.31795/baunsobed.909953

Büyükkarcı, K. (2014). Assessment beliefs and practices of language teachers in primary education. *International Journal of Instruction, 7*(1), 107-120.

Büyükkarcı, K. (2016). Identifying the areas for English language teacher development: A study of assessment literacy. *Pegem Eğitim ve Öğretim Dergisi, 6*(3), 333-346. https://doi.org/10.14527/pegegog.2016.017

Coombe, C., Troudi, S., & Al-Hamly, M. (2012). Foreign and second language teacher assessment literacy: Issues, challenges and recommendations. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoynoff (Eds.), *The Cambridge guide to second language assessment* (pp. 20-29). Cambridge University.

Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn? *Language Testing in Asia, 10*(3), 1-16. https://doi.org/10.1186/s40468-020-00101-6

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. https://doi.org/10.30466/IJLTR.2018.120602

Fraenkel, J.R., Wallen, N.E., & Hyun, H.H. (2012). *How to design and evaluate research in education.* (8th ed.). Mc Graw Hill.

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

Genç, E., Çalışkan, H., & Yüksel, D. (2020). Language assessment literacy level of EFL teachers: A focus on writing and speaking assessment knowledge of the teachers. *Sakarya University Journal of Education, 10*(2), 274-291. doi: https://doi.org/10.19126/suje.626156.

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

Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education & Literacy Studies, 3*(4), 42-48. https://doi.org/10.7575/aiac.ijels.v.3n.4p.42

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

Hinton, P.R., McMurray, I., & Brownlow, C. (2014). *SPSS explained* (2nd ed.). Routledge.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing, 25*(3), 385-402. https://doi.org/10.1177/0265532208090158

Inbar-Lourie, O. (2013). Guest Editorial to the special issue on language assessment literacy. *Language Testing, 30*(3), 301-307. https://doi.org/10.1177/0265532213480126

Işık, A. (2021). Exploring how ELT teachers perceive and practice English language assessment. *Journal of Language and Education, 7*(1), 109-126. https://doi.org/10.17323/jle.2021.10296

İnan-Karagül, B., Yüksel, D., & Altay, M. (2017). Assessment and grading practices of EFL teachers in Turkey. *International Journal of Language Academy, 5*(5), 168-174. https://doi.org/10.18033/ijla.3656

Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. *The International Journal of Research in Teachers Education, 6*(2), 26-37.

Kırkgöz, Y., Babanoğlu, M. P., & Ağçam, R. (2017). Turkish EFL teachers' perceptions and practices of foreign language assessment in primary education. *Journal of Education and e-Learning Research, 4*(4), 163-170. https://doi.org/10.20448/journal.509.2017.44.163.170

Lam, R. (2014). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197. https://doi.org/10.1177/0265532214554321

Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi-Ankara Üniversitesi TOMER [Language Journal-Ankara University TOMER], 167*(1), 43-60.

Muhammad, F. H. N., & Bardakçı, M. (2019). Iraqi EFL teachers' assessment literacy: Perceptions and practices. *Arab World English Journal (AWEJ), 10*(2), 431-442. https://doi.org/10.24093/awej/vol10no2.33

Ölmezer-Öztürk, E., & Aydın, B. (2018). Toward measuring language teachers' assessment knowledge: Development and validation of language assessment knowledge scale (LAKS). *Language Testing in Asia, 8*(20), 1-15. https://doi.org/10.1186/s40468-018-0075-2

Ölmezer-Öztürk, E., & Aydın, B. (2019). Investigating language assessment knowledge of EFL teachers. *Hacettepe University Journal of Education, 34*(3), 602-620. https://doi.org/10.16986/HUJE.2018043465

Önalan, O., & Gürsoy, E. (2020). Primary and secondary level EFL teachers' use of assessment and assessment results in Turkey. *i-managers' Journal on English Language Teaching, 10*(3), 1-11. https://doi.org/10.26634/jelt.10.3.16387

Öz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal, 6*(1), 25-44.

Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher, 6*(1), 21-24.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? Theory into Practice, 48(1), 4-11. https://doi.org/10.1080/00405840802577536

Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator, 46*(4), 265-273. https://doi.org/10.1080/08878730.2011.605048

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

Rogier, D. (2014). Assessment literacy: Building a base for better teaching and learning. *English Language Teaching Forum, 52*(3), 2-13.

Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Working Papers in Applied Linguistics & TESOL, 18*(1), 1-22.

Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan, 72*(7), 534-539.

Tanyer, S., & Susoy, Z. (2018). Assessment of young language learners: Perceptions and practices of Turkish EFL pre-service and in-service teachers. *The Journal of Language Teaching and Learning, 8*(2), 127-149.

Tavassoli, K., & Farhady, H. (2018). Assessment knowledge needs of EFL teachers. *Teaching English Language, 12*(2), 45-65.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics, 29*, 21-36. https://doi.org/10.1017/S0267190509090035

Tsagari, D. (2020). Language assessment literacy: Concepts, challenges, and prospects. In S. Hidri (Ed.), *Perspectives on language assessment literacy: Challenges for improved students learning* (pp. 13-32). Routledge.

Tuzcu-Eken, D. (2016). Testing in English classes in vocational schools of higher education. *Electronic Journal of Vocational Colleges, 6*(1), 38-47.

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly, 11*(4), 374-402. https://doi.org/10.1080/15434303.2014.960046

Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education, 58*, 149-162. https://doi.org/10.1016/j.tate.2016.05.010

Xu, Y., & Brown, G. T. L. (2017). University English teacher assessment literacy: A survey-test report from China. *Papers in Language Testing and Assessment, 6*(1), 133-158. https://doi.org/10.58379/UZON5145

Yan, X., Zhang, C., & Fan, J. J. (2018). "Assessment knowledge is important, but …": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System, 74*, 158-168. https://doi.org/10.1016/j.system.2018.03.003

APPENDIX

Table 3. *Descriptive Statistics of Language Assessment Knowledge Level of EFL Teachers in the K-12 Context*

Items	N	True	False	Don't Know	X̄	S
Assessing Reading						
1. Asking learners to summarize the reading text is a way of assessing their reading skills.	202	170*	17	15	,8416	,36604
2. When asking several questions about a reading text, all the questions are independent of each other.	202	59*	136	7	,2921	,45585
3. Cloze test is used for assessing the main idea of the text.	202	77	106*	19	,5248	,50063
4. In a reading exam, using a text learners have encountered before is not a problem.	202	78	96*	28	,4752	,50063
5. One reading text is enough to be included in a reading exam.	202	70	113*	19	,5594	,49769
6. The language of the questions is simpler than the text itself.	202	140*	43	19	,6931	,46237
7. Errors of spelling are penalized while scoring.	202	78	97*	27	,4802	,50085
8. Taking vocabulary difficulty into consideration is necessary in assessing reading skills.	202	160*	36	6	,7921	,40683
9. Including not stated/doesn't say along with true/false items has advantages over true/false items.	202	144*	25	33	,7129	,45355
10. The more items a reading text is followed, the more reliable it becomes.	202	140*	41	21	,6931	,46237
11. Using the same words in the correct option as in the text is not a problem.	202	61	123*	18	,6089	,48921
12. Simplification of reading texts is avoided.	202	60	111*	31	,5495	,49878
13. Reading texts in a reading exam include various genres (essay, article, etc.).	202	176*	15	11	,8713	,33571
14. In top-down approach, assessment is on overall comprehension of the reading text.	202	151*	21	30	,7475	,43551
15. Using ungrammatical distractors in multiple choice questions in a reading exam is a problem.	202	147*	40	15	,7277	,44624
Reading-Total	202				9,5693	1,98443
Assessing Listening						
16. Using reading texts for listening purposes poses a problem.	202	74*	102	26	,3663	,48300
17. Including redundancy (e.g., what I mean to say is that ….) in a listening text poses a problem.	202	56	129*	17	,6386	,48160
18. Any type of listening text is used for note-taking.	202	98	76*	28	,3762	,48564
19. Spelling errors are ignored in scoring the dictation.	202	95*	96	11	,4703	,50036
20. Errors of grammar or spelling are penalized while scoring.	202	98	83*	21	,4109	,49322
21. A listening cloze test is a way of selective listening.	202	159*	20	23	,7871	,41035
22. Phonemic discrimination tasks (e.g., minimal pairs such as sheep-ship) are examples of integrative testing.	202	120	39*	43	,1931	,39569

23. Scoring in note-taking is straightforward.	202	107	41*	54	,2030	,40321
24. In discrete-point testing, comprehension is at the literal/of local level.	202	100*	26	76	,4950	,50122
25. Using dictation diagnostically in assessing listening skills does not pose a problem.	202	90*	78	34	,4455	,49826
26. Giving learners a transcript of the listening text is a valid way of assessing listening skills.	202	75	115*	12	,5693	,49640
27. Dictation is a kind of discrete-point testing.	202	135	36*	31	,1782	,38365
28. Inference questions based on intelligence are avoided in listening tests.	202	111*	59	32	,5495	,49878
29. Asking learners to listen to names or numbers is called intensive listening.	202	112	54*	36	,2673	,44366
30. In selective listening, learners are expected to look for certain information.	202	178*	12	12	,8812	,32437
Listening-Total	202				6,8317	2,13742
Assessing Writing						
31. Giving two options to learners and asking them to write about one ensure reliable and valid scoring.	202	147	35*	20	,1733	,37942
32. Analytic scoring is used to see the strengths and weaknesses of learners.	202	162*	17	23	,8020	,39950
33. The parts of a scoring scale and the scores in each part do not change for different levels of learners.	202	61	123*	18	,6089	,48921
34. When there is a disagreement between the scores of the two raters, they score the written work again.	202	139	32*	31	,1584	,36604
35. Learners are required to write about at least two tasks in the exam rather than one task.	202	93*	89	20	,4604	,49967
36. Giving restrictive prompts/guidelines to learners for the writing task is avoided.	202	72	100*	30	,4950	,50122
37. Giving learners an opinion and asking them to discuss it is a valid way of assessing their writing skills.	202	150	34*	18	,1683	,37508
38. Using visuals which guide learners for writing poses a problem.	202	53	135*	14	,6683	,47199
39. Holistic scoring is used to see whether the learner is proficient or not at the end of the term.	202	117*	28	57	,5792	,49491
40. Analytic scoring leads to greater reliability than holistic scoring in writing.	202	119*	34	49	,5891	,49322
41. In controlled writing, learners have the chance to convey new information.	202	93	75*	34	,3713	,48435
42. Classroom evaluation of learning in terms of writing is best served through analytic scoring rather than holistic scoring.	202	89*	49	64	,4406	,49769
43. Irrelevant ideas are ignored in the assessment of initial stages of a written work in process writing.	202	109	67*	26	,3317	,47199
44. Providing a reading text for writing is a way of assessing writing skills.	202	116*	60	26	,5743	,49568
45. Mechanical errors (e.g., spelling and punctuation) are dealt with in the assessment of later stages of a written work.	202	129*	54	19	,6386	,48160
Writing-Total	202				7,0594	2,15943
Assessing Speaking						
46. When the interlocutor does not	202	111	76*	15	,3762	,48564

626

understand the learner, giving that feeling or saying it poses a problem.						
47. Giving learners one task is enough to assess speaking skills.	202	49	145*	8	,7178	,45118
48. Interlocutors' showing interest by verbal and non-verbal signals poses a problem.	202	66	111*	25	,5495	,49878
49. When it becomes apparent that the learner cannot reach the criterion level, the task is ended.	202	78*	83	41	,3861	,48807
50. Using holistic and analytic scales at the same time poses a problem.	202	67	74*	61	,3663	,48300
51. Reading aloud is a technique used to assess speaking skills.	202	96*	94	12	,4752	,50063
52. In interlocutor-learner interviews, the teacher has the chance to adapt the questions being asked.	202	163*	19	20	,8069	,39569
53. In interactive tasks, more than two learners pose a problem.	202	72*	113	17	,3564	,48014
54. The interlocutor gives the score when the learner is in the exam room.	202	87	89*	26	,4406	,49769
55. In a speaking exam, production and comprehension are assessed together.	202	166*	26	10	,8218	,38365
56. Asking learners to repeat a word, phrase or a sentence is a way of assessing speaking skills.	202	95*	90	17	,4703	,50036
57. Discussion among learners is a way of assessing speaking skills.	202	179*	13	10	,8861	,31843
58. A checklist is a means of scoring oral presentations in in-class assessment.	202	170*	20	12	,8416	,36604
59. When the focus is to assess discourse, role plays are used.	202	166*	17	19	,8218	,38365
60. In peer interaction, random matching is avoided.	202	65*	109	28	,3218	,46832
Speaking-Total	202				8,6386	2,26219
LAK-TOTAL	202				32,0990	5,43895

* Refers to the corrects answers.