

Turkish EFL Teachers' Beliefs, Practices and Challenges on Assessment

Ajda MUTLU^a

Abstract

Abstract: The purpose of this study was to explore Turkish EFL teachers' assessment beliefs, assessment practices, and the challenges they encounter during the assessment process. In this mixed methods study, the data were collected on both a quantitative and qualitative basis. For the quantitative component, a random sampling method was used to collect data from English teachers through a questionnaire (n=257). For the qualitative component (n=25), the data were collected through interviews with English teachers. The findings revealed English teachers' assessment purposes, practices, use of technology, and the four skills assessment practices. In addition, the qualitative data revealed the barriers and difficulties teachers face in the assessment process, including students' low motivation, low language proficiency, being unready, and insufficient technology in the classrooms.

Keywords

English language teaching
Assessment beliefs
Assessment strategies

Cite as: Mutlu, A. (2025). Turkish EFL teachers' beliefs, practices and challenges on assessment. *Manisa Celal Bayar University International Journal of English Language Studies*, 4(1), 1-19. <https://doi.org/10.5281/zenodo.14767334>

^a Corresponding Author: Ajda MUTLU, Manisa, Türkiye, ajdamutlu87@gmail.com, Orcid: 0000-0002-0604-0527.

Introduction

Assessment in English teaching has always attracted the interest of researchers as teachers are constantly striving to improve methods and techniques for assessing student performance and providing feedback to enhance the learning and teaching process (Brown, 2004). Munoz et al. (2012) argued that language teachers' assessment practices are based on their assessment beliefs. It is important to examine teachers' beliefs to uncover the congruence or discrepancy between teachers' thoughts and actions in the classroom. Similarly, investigating English teachers' assessment beliefs may help teachers adjust their assessment methods in the classroom. In addition, Işık (2021) claimed that although assessment has attracted interest from researchers, assessment research is relatively scarce and insufficient to reveal the assessment practices of Turkish EFL teachers. Moreover, Ölmezer and Öztürk (2018) emphasized that teachers would evaluate their assessment methods through such research which might contribute to updating their assessment practices according to their student's needs. Therefore, this mixed methods study is expected to shed light on language teachers' beliefs, practices, and challenges in the field of language assessment. In addition, similar research may be valuable for providing data to educational institutions (e.g., K-12 levels) to help them revise assessment policies and encourage teachers to use effective assessment practices in their classrooms.

EFL teachers' assessment methods vary according to students' learning needs, school environment, and curriculum (Ezir, 2013). In Yetkin's (2015) classification, various assessment methods such as formative and summative, formal and informal, continuous and final, process-based and product-based, divergent and convergent, criterion-based and norm-based, discrete and comprehensive, objective and subjective were involved. However, since researchers have different classifications of assessment types, it might be concluded that the assessment methods used cannot be considered unique for each class and learner. Therefore, teachers must develop their assessment skills and adapt them to their teaching.

Four Skills Assessment in EFL Classrooms

Language assessment typically involves four skills assessment. These skills cover reading, writing, speaking, and listening skills assessment. Thus, it might be difficult for language teachers to assess all four skills. Since each skill requires a specific assessment approach, method, and strategy, teachers must decide how to select and apply different assessment tools, practices, and techniques in an educational setting (Öz, 2017). Yetkin (2018) pointed out that educational assessment systems only focus on certain characteristics of the four skills during the assessment process. Therefore, researchers have categorized the sub-skills as reading, writing, listening, and speaking (Paker & Höl, 2012). According to Paker and Höl (2012), listening skills can be divided into sub-skills such as skimming, information transfer, dictation, and note-taking. On the other hand, reading skills include sub-skills such as skimming, information transfer, quoting, inference, and deductive information. Writing skills include sub-skills such as description, comparison and contrast, cause and effect, problem-solving and reasoning, and speaking skills are related to structural and functional use in Paker and Höl's (2012) classification.

Various methods have been employed by researchers for reading skills assessment. Brown (2004) argues that in order for students to fully grasp the content they are reading, they must possess a blend of syntactic knowledge, morphological knowledge, genre knowledge, and general word knowledge. Similarly, Hughes (2003) defines macro reading skills as the ability

to scan text for information, skim material for the main idea, identify the different stages of an argument, locate supporting evidence, determine pronoun referents, use context to understand unfamiliar words and recognize connections between elements. By considering these classifications, teachers can establish their own assessment criteria for assignments according to their understanding of both micro- and macro-skills in reading.

Among the four language skills, speaking skills are generally believed to be the most challenging language skill to be assessed. This is because it requires the careful monitoring and evaluation of oral performance. Brown (2004) classifies oral performance into five distinct categories as imitative speaking, intensive speaking, responsive speaking, interactive speaking, and extensive speaking. Hughes (2003) further explains that responsive speaking involves small talk, question-and-answer sessions, and giving directions, while interactive speaking includes discussions, role-plays, interviews, and games. In addition, researchers identified different criteria for speaking skills assessment. For instance, Louma (2004) emphasizes the importance of practicality, validity, reliability, and authenticity in speaking assessment. In addition, Hughes (2003) gives importance to various aspects such as accuracy, appropriateness, a diverse range of language usage, flexibility in starting and directing conversations, speaking time, pronunciation, grammatical correctness, vocabulary usage, fluency, and understanding.

Writing skills are another critical skill for language assessment. The researchers put forward various ideas regarding the assessment of writing skills (Hughes, 2003; Munoz et al., 2012; Weigle, 2002). Munoz et al. (2012) classified writing skills assessment into five forms: narrative, descriptive, expository, persuasive, and argumentative. Weigle (2002) emphasizes the importance of content, rhetorical structure, organization, vocabulary, style, grammar, spelling, punctuation, and meaning accuracy during the assessment process. Similarly, Hughes (2003) points out the importance of assessing various sub-skills, including grammar, vocabulary, mechanics, fluency, and form, when evaluating writing abilities. Taking into account the insights provided by these researchers, it can be concluded that English teachers must possess assessment knowledge to determine assessment criteria effectively and evaluate students' writing skills.

Listening skills assessment is also evaluated as a compelling task for English teachers (Richards, 1983). Richards (1983) categorizes listening subskills into two groups: micro skills and macro skills. These encompass a variety of activities such as tasks with limited responses, multiple-choice tasks, tasks requiring extended communication, tasks involving the recognition of paraphrases, listening to cloze test tasks, sentence repetition tasks, dictation tasks, and tasks requiring communicative responses.

The four skills assessments need to be followed through an effective feedback process provided by language teachers. The feedback process involves the comments and information provided to learners regarding their performance or lacking learning tasks or assessments (Richards & Schmidt, 2011). In addition to teachers' feedback process, self-assessment and peer assessment are considered integral components of the feedback process which develops students' ability to assess their work and that of their peers. Brown (2004) also highlights that the main objective of implementing self-and peer assessment is to assist students in acknowledging the disparity between their present performance and their desired performance, thereby fostering motivation. Therefore, the feedback process might be

considered one of the essential components of the language assessment process. (Cheng & Wang, 2017).

As it has been explained above, beliefs are described as mental representations of reality that encompass meanings, preferences, and beliefs, which in turn help individuals make sense of decisions (William & Thompson, 2007) This definition emphasizes the link between English teachers' assessment practices and their underlying beliefs. Research on assessment has been conducted both in Turkey and around the world. In his study, Öz (2017) investigated the assessment methods used by Turkish EFL teachers in the classroom, comparing factors such as gender, teaching experience, and school type. Another recent study by Ölmezer-Öztürk and Aydın (2019) focused on the knowledge of English teachers employed at universities regarding skill-based language assessment. The findings revealed that the overall language assessment score was moderate, with the lowest score observed in listening assessment knowledge. Similarly, Işık (2021) conducted a study examining English teachers' training in assessment and their assessment practices in the classroom. The results of this study confirmed that teachers' assessment practices are influenced by inadequate pre-service and in-service assessment training, as well as their previous experience in assessment.

The research conducted by Narathakoon et al. (2020) delved into the beliefs of English teachers regarding assessment in the English classroom, as well as their actual practices. The study revealed that various contextual factors, such as educational policies, time constraints, heavy workloads, and a lack of assessment knowledge among teachers, greatly influence their beliefs and practices. Similarly, Wicking (2017) explored English teachers' assessment purposes and discovered that the common goals include determining students' final grades, assessing their progress, and providing feedback. In a quantitative study by Cheng and Wang (2017), which focused on multiple assessments, teachers' beliefs and practices were examined, along with the challenges they experience. The study involved 520 primary school English teachers from Northern Taiwan through a questionnaire and open-ended questions and the results confirmed the effectiveness of formative and summative assessment. Lastly, Abu-Rahmah and Al-Humaidi (2012) proposed that self-assessment as an assessment strategy could enhance student involvement in the assessment process. Hakim (2015) conducted a study to investigate the impact of portfolio-based assessment on students and English teachers, specifically focusing on the wash-back effect. The findings proved that portfolio assessment positively influences the learning process through the wash-back effect.

Methodology

Research Design

As Cresswell (2007) stated, using a mixed-method approach increases the strength of the research conducted. Thus, both qualitative and quantitative research methods were incorporated in the study. The quantitative method which reveals statistical data about the variables was used to have more generalizable results as mentioned by Dawadi et al. (2021). Additionally, the qualitative method was preferred to get information about alternative assumptions and deductions. The qualitative part of the study involved a questionnaire which is used for gathering findings from a practical sample of a broader population (Babbie, 1990). As for the qualitative part of the study, a semi-structured interview which provides the researcher with deductive information about the studied research variables and themes was

preferred (Cresswell, 2007). The following questions involve the research questions of the current study.

- 1) What are the beliefs of Turkish EFL teachers regarding assessment purposes?
- 2) What are the methods, techniques, and frequencies that Turkish EFL teachers typically use for assessment?
- 3) What are the beliefs of Turkish EFL teachers regarding feedback and assessment?
- 4) What are the beliefs and methods of Turkish EFL teachers regarding the four skills assessment involved?
- 5) What are the challenges Turkish EFL teachers encountered during their assessment practices and the factors influencing their beliefs and practices for assessment?

Participants and Data Collection

A total of 257 English language teachers employed by the Turkish Ministry of Education were involved in the research. The participants' profiles differed in gender, age, educational background, undergraduate major, English teaching experience, and level of education as presented in Table 1. The quantitative data was gathered through a website dedicated to English teachers using a purposive sampling method because of COVID-19. The qualitative data was obtained through semi-structured interviews conducted either face-to-face or online meetings with 25 English teachers from the participants of the research.

Table 1. Teachers' Background Information (Frequencies, Percentages)

		Frequency	Percent
English Teaching Experience	1-5 years	32	12
	6-10 years	50	19
	11-20 years	134	52
	more than 20 years	41	16
	Total	257	100
The School Level	Primary School	28	10
	Secondary School	162	63
	High School	67	26
	Total	257	100
The School Type	Public School	249	96,9
	Private School	8	3,1
	Total	257	100,0
Age	21-30	39	15,2
	31-40	146	56,8
	41-50	65	25,3
	Above 51	7	2,7
	Total	257	100
Gender	Female	206	80,2
	Male	51	19,8
	Total	257	100
Educational Background	Undergraduate	189	73,5
	Master's Degree	67	26,1

	PhD	1	,4
	Total	257	100,0
Undergraduate Major	English Language Teaching	194	75,5
	English Language and Literature	43	16,7
	Linguistics	8	3,1
	Other	12	4,7
	Total	257	100,0

Instruments

The research involved a questionnaire and a semi-structured interview. The first research tool is the assessment belief questionnaire created by Elshawara et al. (2017). This questionnaire has 37 Likert scale items and 26 questions that require information about the teachers' opinions about assessment, grading, providing feedback, and reporting. The sample statements of the questionnaire involve;

Table 2. Assessment Beliefs Questionnaire (Likert Scale Items)

	SD	D	U	A	SA
Assessment helps teachers in grouping students for instructional purposes (2) *					
Assessment may provide information about students' progress (5) *					
Paper and pencil assessment is the best method in evaluating students' work. (13) *					
Formal assessment provides a good evaluation of students' work. (11) *					
Informal assessment provides a good evaluation of students' work. (12) *					

Note: Note: The number in the brackets with the asterisk (*) represents the item number in the questionnaire. SD= 'Strongly Disagree'; D= 'Disagree'; U= 'Undecided'; A= 'Agree'; SA='Strongly Agree';

Table 3. Assessment Beliefs Questionnaire (Opinion Statement Items)

	A	D
Reading is assessed through; read aloud (1) *		
Writing is assessed through; reflective writing (10) *		
Listening is assessed through; taking notes (24) *		
Speaking is assessed through; oral interview (3) *		

Note: "A" stands for Agree and "D" stands for Disagree.

The semi-structured interview involved 10 open-ended questions that were adapted from Dasthi's (2019) study. All the interview questions were parallel with the research questions. Twenty English teachers employed at different school levels in MoNE were interviewed via phone conversations, online meetings, or in-person meetings. Before conducting the interviews, three teachers were interviewed and their suggestions were used to ensure clarification-related improvements. The interviews lasted about fifteen minutes and they were recorded by the researcher.

The sample interview questions involve;

1. How do you assess your student's performance? Please describe the content of your exams.
2. How often do you assess your students in your class? Why?

3. Which skills do you assess? What kind of assessment is most effective for measuring students' performance in these skills?
4. What kind of feedback do you usually provide to your students? Could you please give me some examples of your feedback?

Data Analysis

The quantitative data were analyzed through descriptive statistics. Using descriptive statistics, the association between different variables in a sample or population was revealed. Additionally, English teachers' responses to the items in the questionnaire about their practices and beliefs were analyzed and descriptive graphs were employed. The qualitative data was manually transcribed by the researcher and content analysis was used. Bengtsson (2016) described the phases of content analysis as de-contextualization, re-contextualization, classification, and compilation.

The researcher utilized these steps during the content analysis. In order to extract meaningful units known as codes from the data during the de-contextualization step, the researcher first transcribed the data. Secondly, the codes and themes were controlled to check if all the aspects of the study were reflected. Themes were identified during the classification phase and assessed as a comprehensive idea of an underlying meaning at an interpretative latent level (Graneheim & Lundman, 2004). Thirdly, the researcher focused on the teachers' words during the compilation step, and the themes were displayed as quotes. Finally, the codes' frequencies were presented in the tables.

Results

The findings are organized according to the research questions. First, the results of the quantitative data are presented, followed by the qualitative data.

Turkish EFL Teachers' Beliefs About Assessment Purposes

The first research question seeks to answer English teachers' beliefs about the purposes of assessment. The data revealed that English teachers mostly based their assessment purposes (Table 4) on informing students about their progress and giving information about their instruction and teaching methods. In addition, English teachers mostly stated their assessment purposes as measuring achievement, providing feedback, and increasing students.

Table 4. Frequencies, Percentages, Means and Standard Deviations of the Assessment Beliefs about the Instructional Assessment Purposes (n =257, Overall Mean =3.81 and SD=1.21)

No	Item	SD	D	U	A	SA	Mean	SD
		F(P)	F(P)	F(P)	F(P)	F(P)		
1	Assessment... may provide information about students' progress (5)*	23 (%8.9)	11 (%4.3)	17 (%6.6)	101 (%39.3)	105 (%40.9)	3.99	1.21
2	helps me to focus on teaching (1) *	27 (%10.5)	19 (%7.4)	25 (%9.7)	118 (%45.9)	68 (%26.5)	3.70	1.23

3	helps me to group students for instructional purposes (2) *	25 (%9.7)	20 (%7.8)	23 (%9.0)	136 (%52.9)	53 (%20.6)	3.67	1.17
4	can diagnose strengths and weaknesses in teaching (4) *	28 (%10.9)	9 (%3.5)	21 (%8.2)	110 (%42.8)	89 (%34.6)	3.87	1.24

Note: Note: The number in the brackets with the asterisk (*) represents the item number in the questionnaire. SD= 'Strongly Disagree'; D= 'Disagree'; U= 'Undecided'; A= 'Agree'; SA= 'Strongly Agree';

Turkish EFL Teachers' Beliefs and Practices about Methods and Techniques of Assessment

The third research question seeks English teachers' beliefs about methods and techniques of assessment. Although they were both highly preferred as assessment types (Table 5), informal assessment was more favoured by the teachers compared to formal assessment. Additionally, teachers were against the idea of the use of paper-pencil assessment. On the other hand, quantitative data revealed that teachers used written exams and quizzes which usually involved tests, completion tasks, and skill-based questions. Additionally, the grading methods of teachers (Table 6) were usually based on students' classroom participation, making presentations and projects, portfolios and homework. As the performers of the assessment, teachers stated that they usually used internet sources and cooperated with their colleagues and written sources respectively while preparing their assessment tools. Moreover, teachers believe that the feedback provided needs to involve giving verbal explanations along with using digital assessment tools. As for the feedback and process, teachers used both self-assessment and peer assessment equally in their classes. Lastly, the formative assessment had a higher level of agreement compared to summative assessment and teachers' assessment frequency varied according to the school types they worked at.

Table 5. Frequencies, Percentages, Means and Standard Deviations of the Assessment Beliefs about Assessment Formats (n =257, Overall Mean = 3.49, SD=1.18)

No	Item	SD F(P)	D F(P)	U F(P)	A F(P)	SA F(P)	Mean	SD
1	Formal assessment provides a good evaluation of students' work. (11) *	27 (%10.5)	38 (%14.8)	36 (%14.0)	111 (%43.2)	45 (%17.5)	3.42	1.24
2	Assessment questions should reflect real life language use. (15) *	20 (%7.8)	17 (%6.6)	22 (%8.6)	110 (%42.8)	88 (%34.2)	3.89	1.18
3	Informal assessment provides a good evaluation of students' work. (12) *	18 (%7)	14 (%5.5)	24 (%9.3)	129 (%50.2)	72 (%28.0)	3.87	1.10

4	Paper and pencil assessment is the best method in evaluating students' work. (13) *	37 (%14.4)	84 (%32.7)	60 (%23.4)	50 (%19.5)	26 (%10.1)	2.78	1.21
---	---	---------------	---------------	---------------	---------------	---------------	------	------

Note: The number in the brackets with the asterisk (*) represents the item number in the questionnaire. SD= 'Strongly Disagree'; D= 'Disagree'; U= 'Undecided'; A= 'Agree'; SA='Strongly Agree'; F='Frequency'; P = 'Percentage'

Table 6. Turkish EFL Teachers' Assessment Practices

Themes	Codes	Teachers' Opinions	f	%
Expression on using written exams and quizzes	-written exams	1,2,3,4,5,9,10,11,12,13,15,16,17,18,19,21,22,23	18	72
	-end of unit quizzes	1,2,3,4,7,23	6	24
Expression of the preferred content on exams	-four skill included	12,2	2	8
	-matching, multiple choice tests, fill in the blank parts, and reading a text	9,13,21	3	12
Expression on the content of grading	-term project	1,2	2	8
	-performance work	12,15,17,21,22	5	20
	-homework	6,8,14,15,22	5	20
Expression on the types of tasks and activities used	-classroom participation	2,8,13,14,15,16,17,22	8	32
	-role play activities	6, 21	2	8
	-playing games	6,8,21	3	12
	-portfolios	20	1	4
	--presentations	6,11,21,22	4	16
	-self assessment	6	1	4
	-preparing poster	7	1	4
Expression on the type of feedback provided	-commenting on students work	6	1	4
	-checking grammar& spelling mistakes	7	1	4
	-using online assessment tools	2,4	2	8
	-giving minus and plus	5	1	4

Turkish EFL Teachers' Beliefs and Practices about Feedback and Grading

In the second research question, the beliefs and practices of English teachers about grading and feedback are explored. Teachers don't have an agreement on the idea that students' grades should only depend on their exams (Table 7). On the other hand, EFL teachers agree on informing students about the assessment criteria before the exam and giving feedback in a short period. Additionally, teachers disagreed on the superiority of criterion-based assessment over norm-based assessment and the feedback types used involved positive and informal

feedback. The preparation of rubrics is primarily based on teachers' individual work or collaborative work with the other teachers.

Table 7. The Feedback Types English Teachers Provide to their Students

Themes	Codes	Opinions	f	%
Informal Feedback	Positive chatting with students	8,11,14,16,17,20,21,23	8	32
Formal Feedback	Exams, written comments	7,13,21,24	4	16
Formative Feedback	detailed explanation on learning deficiencies	19	1	4
Summative Feedback	worksheets, quizzes exams at the end of the units	1,21,25	3	12
Student peer feedback	Comments on their friends' assignment	6	1	4
Positive Feedback	Sayings like "Well done, good example", applaud, rewards, play games, giving pluses	1,2,5,10,11,14,15,18,20,21,22,25	12	48
Negative Feedback	Show their mistakes	8, 13, 24	3	12

Turkish EFL Teachers' Beliefs and Practices about Four Skills Assessment

The fourth research question addresses beliefs and practices related to the assessment of four skills. The four skills are evaluated in terms of traditional and alternative assessment methods.

For reading skills (Table 8), while the most frequently used methods involve true-false items, tests, and matching items as traditional methods, the alternative methods involve read-aloud, self-evaluation, portfolios, and peer assessment.

Table 8. Frequencies, Percentages, Means and Standard Deviations of the Assessment Beliefs about Alternative Types of Reading Skill Assessment (n = 257; Overall Mean =0.22)

No	Item	D F(P)	A F(P)	Mea n	SD
	Reading is assessed through;				
1	read aloud (1) *	48 (%18.7)	209 (%81.3)	0.81	0.39
2	self-assessment (26) *	127 (%49.4)	130 (%50.6)	0.51	0.50
3	student portfolios (12) *	140 (%54.5)	117 (%45.5)	0.46	0,50
4	peer assessment (25) *	142 (%55.3)	115 (%44.7)	0.45	0.50
5	taking notes (24) *	159 (%61.9)	98 (%38.1)	0.38	0.49
6	role play (18) *	174 (%67.7)	83 (%32.3)	0.32	0.47
7	reflective writing (10) *	188 (%73.2)	69 (%26.8)	0.27	0.44
8	oral presentation (16) *	194 (%75.5)	63 (%24.5)	0.25	0.43

9	retelling a story after listening to a passage (23) *	193 (%75.1)	64 (%24.9)	0.25	0.43
10	oral summaries following a listening passage (22) *	193 (%75.1)	64 (%24.9)	0.25	0.43
11	oral discussion (19) *	198 (%77)	59 (%23)	0.23	0.42
12	essay writing (11) *	199 (%77.4)	58 (%22.6)	0.23	0.42
13	oral interview (3) *	212 (%82.5)	45 (%17.5)	0.18	0.38

Note: The number in the brackets with the asterisk (*) represents the item number in the questionnaire. D= 'Disagree' A= 'Agree'; F= 'Frequency'; P = 'Percentage'

For writing skills (Table 9), while the most commonly used methods are summarizing, sentence completion items, task edition, and recognition of errors, the alternative methods involve reflective writing, writing an essay, portfolios, and note taking.

Table 9. Frequencies, Percentages, Means and Standard Deviations of the Assessment Beliefs about Alternative Types of Writing Skill Assessment (n = 257; Overall Mean =0.2)

No	Item	D F(P)	A F(P)	Mean	SD
Writing is assessed through;					
1	reflective writing (10) *	80 (%31.1)	177 (%68.9)	0.69	0.46
2	essay writing (11) *	86 (%33.5)	171 (%66.5)	0.67	0.47
3	student portfolios (12) *	90 (%35)	167 (%65)	0.65	0.48
4	taking notes (24) *	98 (%38.1)	159 (%61.9)	0.62	0.49
5	self-assessment (26) *	109 (%42.4)	148 (%57.6)	0.58	0.50
6	peer assessment (25) *	122 (%47.5)	135 (%52.5)	0.53	0.50
7	read aloud (1) *	191 (%74.3)	66 (%25.7)	0.26	0.44
8	retelling a story after listening to a passage (23) *	193 (%75.1)	64 (%24.9)	0.25	0.43
9	oral presentation (16) *	198 (%77)	59 (%23)	0.23	0.42
10	oral summaries following a listening passage (22) *	200 (%77.8)	57 (%22.2)	0.22	0.42
11	role play (18) *	200 (%77.8)	57 (%22.2)	0.22	0.42
12	oral discussion (19) *	208 (%80.9)	49 (%19.1)	0.19	0.39
13	oral interview (3) *	225 (%87.5)	32 (%12.5)	0.12	0.33

Note: The number in the brackets with the asterisk (*) represents the item number in the questionnaire. D= 'Disagree' A= 'Agree'; F= 'Frequency'; P = 'Percentage'

For listening skills (Table 10), half percent of teachers reported using true-false items, matching activities, and tests as traditional methods. As for the alternative methods, note-taking, oral summaries, and story narration were also preferred by less than half a percent of teachers.

Table 10. Frequencies, Percentages, Means and Standard Deviations of the Assessment Beliefs about Alternative Types of Listening Skill Assessment (n = 257; Overall Mean =0,40)

No	Item	D F(P)	A F(P)	Mean	SD
Listening is assessed through;					

1	taking notes (24) *	136 (%52.9)	121 (%47.1)	0,47	0.50
2	oral summaries following a listening passage (22) *	137 (%53.3)	120 (%46.7)	0,47	0.50
3	retelling a story after listening to a passage (23) *	141 (%54.9)	116 (%45.1)	0,45	0.50
4	role play (18) *	147 (%57.2)	110 (%42.8)	0,43	0.50
5	self-assessment (26) *	150 (%58.4)	107 (%41.6)	0,42	0.49
6	oral presentation (16) *	155 (%60.3)	102 (%39.7)	0,40	0.49
7	oral discussion (19) *	155 (%60.3)	102 (%39.7)	0,40	0.49
8	peer assessment (25) *	157 (%61.1)	100 (%38.9)	0,39	0.49
9	oral interview (3) *	162 (%63)	95 (%37)	0,37	0.48
10	student portfolios (12) *	164 (%63.8)	93 (%36.2)	0,36	0.48
11	read aloud (1) *	181 (%70.4)	76 (%29.6)	0.30	0.46
12	reflective writing (10) *	210 (%81.7)	47 (%18.3)	0.18	0.39
13	essay writing (11) *	226 (%87.9)	31 (%12.1)	0.12	0.33

Note: The number in the brackets with the asterisk (*) represents the item number in the questionnaire. D= 'Disagree' A= 'Agree'; F= 'Frequency'; P = 'Percentage'

For speaking skills (Table 11), teachers had a low level of agreement on using task descriptions and information transfer as traditional methods. On the other hand, oral interviews, role play, and oral presentations were usually favored as alternative methods. Additionally, in light of the qualitative data, the most common methods used in four skills were revealed. For speaking skills oral presentation; for reading skills close tests and sentence completion; for writing skills doing homework; for listening skills listening texts are usually preferred by teachers. Moreover, grammar is assessed through sentence completion items, tests, and written exams while vocabulary is assessed with written exams and quizzes. When the frequency of assessment among the four skills is evaluated it is seen that reading skill is the most assessed skill followed by writing skills, speaking skills, and listening skills respectively. The four skills assessed are also followed by grammar and vocabulary knowledge assessment with a low percent of use.

Table 11. Frequencies, Percentages, Means and Standard Deviations of the Assessment Beliefs about Alternative Types of Speaking Skill Assessment (n = 257; Overall Mean =0,53)

No	Item	D F(P)	A F(P)	Mean	SD
	Speaking is assessed through;				
1	oral interview (3) *	56 (%21.8)	201 (%78.2)	0.78	0.41
2	role play (18) *	73 (%28.4)	184 (%71.6)	0.72	0.45

3	oral presentation (16) *	86 (%33.5)	171 (%66.5)	0.67	0.47
4	oral discussion (19) *	90 (%35)	167 (%65)	0.65	0.48
5	public speaking (20) *	104 (%40.5)	153 (%59.5)	0.60	0.49
6	retelling a story after listening to a passage (23) *	125 (%48.6)	132 (%51.4)	0.51	0.50
7	oral summaries following a listening passage (22) *	131 (%51)	126 (%49)	0.49	0.50
8	self-assessment (26) *	155 (%60.3)	102 (%39.7)	0.40	0.49
9	peer assessment (25) *	155 (%60.3)	102 (%39.7)	0.40	0.49
10	read aloud (1) *	164 (%63.8)	93 (%36.2)	0.36	0.48
11	student portfolios (12) *	171 (%66.5)	86 (%33.5)	0.33	0.47
12	taking notes (24) *	195 (%75.9)	62 (%24.1)	0.24	0.43
13	reflective writing (10) *	228 (%88.7)	29 (%11.3)	0.11	0.32
14	essay writing (11) *	231 (%89.9)	26 (%10.1)	0.10	0.30

Note: The number in the brackets with the asterisk (*) represents the item number in the questionnaire. D= 'Disagree' A= 'Agree'; F= 'Frequency'; P = 'Percentage'

The Challenges Turkish EFL Teachers Experience During Their Assessment Practices in the Classroom

The fifth research question addresses the challenges English teachers face during the assessment process in the classroom. The challenges are evaluated as student-related challenges and external challenges. The student-based challenges involve students' lacking motivation, cheating in the exams, low proficiency levels, discipline problems, being unready for classes, material deficiency, and attendance problems. When the frequency of these problems is investigated it is seen that the most problematic ones are students' lacking motivation, low proficiency level, being unready for classes, and discipline problems respectively.

As for the external factors, technological deficiencies in classrooms, family interference, and lack of enough class hour time are identified as the most problematic issues. On the other hand, when teachers' and students' responses to these challenges are checked, it is observed that teachers struggle to increase students' motivation, change their assessment methods, and try to prevent cheating in the exams as a response to students-based challenges. As for the external challenges, teachers keep contact with the families, change their assessment methods, and use peer work activities for time deficiency. The qualitative data also provided information about the underlying factors behind English teachers' beliefs and practices through the assessment process. These factors involve the school types, education policy, students' language proficiency, technological limitations in the classroom, and their assessment knowledge.

Discussion and Conclusion

The discussion and conclusion part involve the discussion of the five research questions. Each research question is discussed based on both quantitative and qualitative data.

Turkish EFL Teachers' Purposes of Assessment

As the previous research supports, the purpose of assessment is to inform teachers about the student's progress, accomplishments, and limitations. (Harris & Brown, 2014; Munoz et.al.; Yetkin, 2015). Yetkin's (2015) research demonstrated that pre-service English teachers usually use classroom assessment to monitor their students' academic development and evaluate their

teaching methods. Similarly, Harris and Brown (2014) underlined the critical role of assessment in the teaching process and the value of teachers in identifying and resolving potential problems. The quantitative results of the current study point out that English teachers' purpose of assessment involves giving information about teachers' instruction and informing about students' progress in line with the previous studies (Dasthi, 2019; Munoz et al., 2012; Yetkin, 2015).

Turkish EFL Teachers' Assessment Methods and Techniques

The study results confirm that English teachers typically used various types of assessment methods such as formal, informal, summative, and formative assessment methods to encourage students' success. Previous researchers focused on different methods (Hakim, 2015; Öz, 2017). Hakim (2015) acknowledged that teachers must recognize both summative and formative assessment methods and use them in their classes effectively. Also, informal assessment was claimed to be more effective at evaluating students' progress, this would indicate that the teacher's priority was to spread the assessment process over a larger scope, including students' learning needs and deficiencies, rather than focusing on exams alone (Öz, 2017). On the other hand, in Brown's study, teachers preferred formative assessment tools to encourage students to participate in assessing their learning progress using alternative assessment methods. Ultimately, the methods and practices of teachers differ, which could lead to the belief that the methods' being efficacious is more important than the number or variety of them.

Along with the assessment types, the content of the assessment tools proves that English teachers employ multiple methods in the assessment process, including written exams, questions, projects, performance work, assignments, portfolios, self-assessments, and providing different types of feedback. The exam content includes multiple-choice questions based on the text, fill-in-the-blank questions, and reading texts. Öz (2017) also recorded that fill-in-the-blank questions, multiple choice questions, true-false questions, matching questions, and short answer questions were the most commonly used assessment methods by used by Turkish EFL teachers. In light of the previous research, it is observed that the methods EFL teachers use for the assessment process are changeable. (Öz, 2017; Ölmezer & Öztürk, 2019; Yetkin 2015).

Along with the assessment types, the content of the assessment tools proves that English teachers employed multiple methods in the assessment process, including written exams, questions, projects, performance work, assignments, portfolios, self-assessments, and providing types of feedback. The exam content includes multiple-choice questions based on the text, fill-in-the-blank questions, and reading texts. Öz (2017) also recorded that fill-in-the-blank questions, multiple choice questions, true-false questions, matching questions, and short answer questions were the most commonly used assessment methods by used by Turkish EFL teachers. In light of the previous research, it is seen that the methods teachers use for the assessment process are changeable as there are numerous assessment methods (Öz, 2017; Ölmezer & Öztürk, 2019; Yetkin,2015).

Turkish EFL Teachers' Assessment Methods and Techniques

The study results regarding the assessment beliefs and practices of English teachers on giving feedback and grading students', getting immediate feedback after the exams, and preparing a

marking scheme before giving their exams (Brown, 2004; Elshawara et al., 2017). Brown (2014) suggested that students' classroom participation, attendance, and improvement should be assessed through feedback which might increase the student's motivation and success. In parallel, the study reveals that teachers usually comment on the students' work after utilizing different types of assessment either as a group or individually, using terms like "Thank you, very good, okay, your idea is correct, my students can answer" which is a sign of positive feedback. Conversely, some teachers preferred negative feedback to emphasize students' errors (Al-Humaidi & Rahmah, 2012).

However, Harris and Brown (2014) suggested that the assessment process should no longer be only considered a testing process by teachers, but rather a necessary skill for improving students' self-regulation. Therefore, it can be seen that although self-assessment and peer assessment are pointed out as being very important, they have low performance in practice. While English teachers tend to use multiple types of feedback including positive feedback, informal feedback, and formal feedback, self-assessment and peer feedback are seen to be rarely used. This may be attributed to the teachers' negative beliefs towards these methods or their lack of knowledge on how to utilize them in the classroom effectively. As for the grading system along with the feedback process, English teachers believe that the final evaluation should include assignments, classroom performance, and written tests. However, English teachers' ideas in the current study differed regarding utilizing rubrics or establishing standards for evaluation. While approximately half of the teachers mentioned using self-created rubrics or rules, the other half preferred using The MoNE's rubrics and collaborative work. In addition, Nunan (2007) remarked on the necessary characteristics of a rubric as objectivity, covering all of the requirements and having a communicative basis for the assessment process in his study.

Turkish EFL Teachers' Four Skills Assessment Methods

In the study, reading was determined to be the most commonly assessed skill in four skills. The reading skills are followed by writing, speaking, listening, and writing skills. According to Ölmezer-Öztürk and Aydın (2019), this may be because reading skills are highly valued, teachers have more expertise and access to more resources, and students find it easier to assess reading skills. On the other hand, the four skills assessment was evaluated in terms of traditional and alternative assessment types. Brown and Abewickrama (2010) distinguished assessment types as traditional or alternative. The current study revealed that English teacher's practices and beliefs regarding the four skills assessment methods in terms of traditional and alternative methods were different. While this study identified true-false tasks, multiple-choice tasks, and matching items as the most commonly used traditional assessment methods, similar studies have found that read-aloud, self-assessments, and student portfolios are the most popular traditional assessment methods used for reading skills (Brown & Abewickrama, 2010; Elshawara et al., 2017).

When the assessment methods were investigated regarding listening skills, the most common traditional assessment methods involved true, false, and multiple choice items, taking notes, narratives, and recounting a story were highly preferred as an alternative method. For speaking skills, information transfer, error recognition, and dictation were typically employed as traditional methods. On the other hand, oral interviews, role plays, and spoken presentations were used by English teachers as alternative methods as in Elshawara et al.'s

(2017) study. As a result, it might be concluded that English teachers determine their assessment methods according to their assessment beliefs and practices.

The Challenges and The Influencing Factors on Turkish EFL Teachers during Their Assessment Practices

The difficulties English teachers encountered during the assessment process were students' low motivation, students' low proficiency, students' not being ready for classes, or technological issues in classes, all of which are considered to be challenges for them as in similar research (Brown & Abeywickrama, 2010). As a response to the mentioned challenges, English teachers struggle to endeavour to increase students' motivation, control cheating, alter the way they assess student performance, communicate with parents regarding behavioural issues, and collaborate with other teachers. As a result, it may be possible to deduce that the English teachers' sources of difficulty are diverse due to the different teaching environments and the struggle to achieve their goals as teachers. On the other hand, The study results showed that English teachers' assessments of students and their practices are also influenced by factors such as the type of school and teachers' lack of knowledge regarding assessment.

The current study results revealed that English students expressed their lack of knowledge regarding the assessment process and their need for additional training on assessment. This outcome was backed by the previous research on teacher's assessment knowledge (Ölmezer-Öztürk & Aydın, 2019; Yetkin, 2015). The findings positively indicated that teachers were aware of their inadequacy and wanted to improve their performance on assessment knowledge. In addition to stating their training need on assessment, about one-third of the English teachers mentioned that MoNE's policy regarding assessment was a significant part of their practice of assessment and criticized the educational policy for not taking advantage of the various methods and practices of assessment in their classroom in the current study. However, the reform of the curriculum assessment strategy takes into account the Common European Framework of Reference for Languages: Learning, Teaching, and Assessment (CEFR), which promotes a variety of assessment methods, including self-assessment, formal assessment, questioning, written and oral examinations, projects and assignments completed during the academic year. Despite the suggestion of taking advantage of different assessment methods, teachers perceive educational policy and central exams as barriers to their assessment practices. In contrast, Paker and Höl (2012) pointed out that despite curriculum changes, the assessment of oral performance has been compromised and remains a concern for learners which is also confirmed by the current study results. In addition, Good and Brophy (2000) also pointed out that English teachers, as observers of students' classroom performance, policymakers do not prioritize their assessment concerns and decisions. As a result, it may be deduced that as English teachers' assessment practices are mostly shaped by the curriculum and the placement exams, there needs to be conformity between language teachers' and policymakers' expectations to ensure students' learning and success.

Implications

As explained before, the study aimed to investigate Turkish EFL teachers' assessment beliefs, practices, and challenges. There have been a variety of research studies in the literature. However, these studies have mostly been conducted with pre-service EFL teachers in Turkey. Thus, this study might be valuable in revealing insights into Turkish EFL teachers employed at the Turkish Ministry of Education. That's why, increasing the number of parallel research

might help English teachers gain awareness about their assessment practices and improve their assessment skills. On the other hand, the results of such research might also reveal information about English teachers' possible training needs on assessment practices and contribute to teachers' professional development. As it is known, this study benefited from a questionnaire a research tool. In addition to the questionnaire, a scale might be developed about the assessment practices of Turkish FLE teachers to ensure a higher level of reliability. The current study is limited to 257 English teachers. The number might be increased in future studies. Moreover, different research designs such as experimental studies might be conducted to observe the outcomes of the applied assessment methods. Additionally, classroom observations might be involved to have a closer view regarding teachers' assessment practices. Lastly, English teachers might be supported through in-service training on assessment practices.

References

- Abu-rahmah, M., & Al-humaidi, S. (2012). Enhancing ELT through self-assessment. *ELT Research Journal*, 1(1), 31-52.
- Babbie, E. (1990). *Survey research methods* (2nd ed.). Belmont, CA: Wadsworth.
- Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *Nursing Plus Open*, 2. 10.1016/j.npls.2016.01.001
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education. 324 pp. ISBN 0-13-098834-0
- Brown, H.D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practice*. Pearson Longman.
- Cheng, L., & Wang, X. (2017). Grading, feedback and reporting in ESL/EFL classrooms. *Language Assessment Quarterly*, 4(1), 85-107.
- Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Dashti, S. (2019). *Efl Teachers' Beliefs and Practices about Classroom Assessment: A Multiple Case Study in the Context of Kuwait* [Unpublished doctoral dissertation]. York University.
- Dawadi, S., Shrestha, S., & Giri, R. (2021). Mixed-Methods Research: A Discussion on Its Types, Challenges, and Criticisms. *Journal of Practical Studies in Education*, 2, 25-36. <https://doi.org/10.46809/jpse.v2i2.20>
- Elshawara, N. R. M., Abdullah, A. N., & Md. Rashid, S. (2017). Malaysian Instructors' Assessment Beliefs in Tertiary ESL Classroom. *International Journal of Education & Literacy Studies*, 5(2), 135-141.
- Ezir, E. (2013). *Assessment and Testing in ELT: The Difference between Assessment and Testing*. *Jurnal SAINTIKOM*, 12 (1), 37-42.
- Good, T., & Brophy, J. (2000). School effects. In: Wittrock, M.C.(Eds.), *Handbook of Research on Teaching*. (p. 570-602). New York, Macmillan.

- Graneheim, U.H., Lundman B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today*. 24 (2), 105-12. doi: 10.1016/j.nedt.2003.10.001. PMID: 14769454.
- Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education & Literacy Studies*, 3(4), 42-48.
- Harris, L. R., and Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: case studies into teachers' implementation. *Teach. Teach. Educ.* 36, 101–111. doi: 10.1016/j.tate.2013.07.008
- Hughes, A. (2003) *Testing for Language Teachers (2nd Edition)*. Cambridge.
<https://doi.org/10.1017/CBO9780511732980>
- Işık, A. (2021). Exploring How ELT Teachers Perceive and Practice English Language Assessment. *Journal of Language and Education*, 7(1), 109-126.
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.
<https://doi:10.1017/CBO9780511733017>
- Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' Beliefs about Assessment in an EFL context in Colombia. *Profile Issues in Teachers' Professional Development*, 14 (1), 143-158.
- Narathakoon, A., Sapsirin, S., & Subphadoongchone, P. (2020). Beliefs and Classroom Assessment Practices of English Teachers in Primary Schools in Thailand. *International Journal of Instruction*, 13(3), 137-156.
- Nunan, D. (2007). Standards-based approaches to the evaluation of ESL instruction. In: Cummins, J., Davison, C. (Eds.) *International handbook of English language teaching. Springer International handbooks of education*. (pp.421-438). Springer, Boston, MA.
https://doi:10.1007/978-0-387-46301-8_31
- Ölmezer-Öztürk, E. & Aydın, B. (2019). Investigating Language Assessment Knowledge of EFL Teachers. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* , 34 (3) , 602-620.
- Öz, S. (2017). Turkish EFL instructors' in-class language assessment literacy: perceptions and practices. *ELT Journal*, 6 (1).
- Paker, T. & Höl, D. (2012). İngilizce Hazırlık Öğrencilerinin ve İngilizce Okutmanlarının İletişimsel Konuşma Sınavına İlişkin Tutum Ve Algıları. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 32 (32) , 13-24.
- Richards, J. (1983). Listening comprehension: Approach, design and procedure. *TESOL Quarterly*, 17(2). pp. 219-240.
- Richards, J.C., & Schmidt, R.W. (2011). *Longman Dictionary of Language Teaching and Applied Linguistics* (4th ed.). Routledge. <https://doi.org/10.4324/9781315833835>
- Weigle, S. (2002). *Assessing Writing*. Cambridge: Cambridge University Press.
- Wicking, P. (2017). The Assessment Beliefs and Practices of English Teachers in Japanese Universities. *JLTA Journal*, 20, 76-89.

- Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: what will it take to make it work? In C.A. Dwyer (Eds.), *The future of assessment: Shaping teaching and learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Yetkin, C. (2015). *An investigation on ELT teacher candidates' assessment literacy* (Unpublished Master's Thesis). Çağ Üniversitesi Sosyal Bilimler Enstitüsü, Mersin.