PRE-SERVICE ELT TEACHERS' PERSPECTIVES OF LANGUAGE ASSESSMENT IN EMERGENCY DISTANCE EDUCATION

Mehmet TUNAZ[*]

Abstract

In the realm of education, and especially in English language teaching (ELT) field, the relationship between teaching strategies and the assessment process is very close. Like instructional strategies, current events like COVID-19 may have an immediate impact on assessment procedures. In order to better understand how traditional and alternative assessment types in emergency distance education relate to in-class instruction, this research sought to interview pre-service language teachers. Students from the ELT department in the 2nd (N: 60) and 4th (N: 66) grades participated in the study. A semi-structured interview was used in this mix-type study to triangulate the quantitative data after the participants completed a Likert-type questionnaire. Four key themes emerged from the independent sample t-test and content analysis: (a) Pre-service teachers for the 2nd and 4th grades differed significantly, and students in the latter grade are more accustomed to using alternative assessment methods; (b) Since 2nd grader students were largely unfamiliar with alternative assessment, there were statistical differences in the perspectives on traditional assessment; (c) 4th grade pre-service teachers favored skill-based assessment more than 2nd grade students; and (d) when it came to the integration of alternative and traditional assessment, there was no discernible difference between the students. In light of the study findings, it was proposed that alternative assessment types be included into the present ELT curriculum at earlier grade levels.

Keywords: Distance education, ELT, Traditional and alternative assessment, Covid-19.

İngilizce Öğretmeni Olacak Öğrencilerin Zorunlu Uzaktan Eğitimde Ölçme Değerlendirme Tutumları

Öz

Eğitim alanında ve özellikle İngiliz dili eğitimi (ELT) alanında, öğretim stratejileri ile değerlendirme süreci arasındaki ilişki çok yakındır. Öğretim stratejileri gibi, COVID-19 gibi güncel olayların da değerlendirme süreçleri üzerinde ani bir etkiye sahip olabilmektedir. Acil uzaktan eğitimde geleneksel ve alternatif değerlendirme türlerinin sınıf içi eğitimle ilişkisinin anlaşılması amacıyla bu araştırma, dil öğretmen adaylarının görüşlerini tespit etmeyi amaçlamıştır. Çalışmaya İngilizce Öğretmenliği bölümü 2. (N:60) ve 4. (N:66) sınıf öğrencileri katılmıştır. Bu karma tip çalışmada, katılımcılar Likert tipi bir anketi tamamladıktan sonra nicel verileri üçgenlemek için katılımcılarla yarı yapılandırılmış bir görüşme yapılmıştır. Bağımsız örneklem t-testi ve içerik analizinden dört ana tema ortaya çıkmıştır: (a) 2. ve 4. sınıflardaki öğretmen adaylarının alternatif değerlendirme yöntemlerini kullanma alışkanlıkları yönünden 4. sınıftaki öğrenciler lehine farklılık gösterdikleri tespit edilmiştir; (b) 2. sınıf öğrencileri alternatif değerlendirmeye büyük ölçüde aşina olmadıklarından, geleneksel değerlendirmeye bakış açılarında istatistiksel olarak daha fazla yatkınlık gösterdikleri tespit edilmiştir; (c) 4. sınıf öğretmen adayları, 2. sınıf öğrencilerine göre beceriye dayalı değerlendirmeyi daha fazla tercih etmişlerdir, ve (d) alternatif ve geleneksel değerlendirmenin birlikte kullanılması söz konusu olduğunda, iki grup arasında anlamlı bir fark çıkmamıştır. Çalışma bulguları ışığında, alternatif değerlendirme türlerinin mevcut ELT müfredatına daha önceki sınıf seviyelerinde dâhil edilmesi önerilmiştir.

Anahtar sözcükler: Uzaktan eğitim, ELT, Geleneksel ve alternatif değerlendirme, Covid-19.

[*] Dr. Öğr. Üyesi, Nevşehir Hacı Bektaş Veli Üniversitesi, Yabancı Diller Yüksekokulu, Mütercim ve Tercümanlık Bölümü (İngilizce), mehmet.tunaz@nevsehir.edu.tr, https://orcid.org/0000-0002-1401-5073

1. Introduction

Assessment is a crucial step in all educational contexts since it serves to determine the effectiveness of the process by focusing on the process itself or the final product. As stated in Brown and Abeywickrama (2010, p.9), "assessment instrument, whether formal tests or informal assessments, serve measure proficiency, place students into one several levels of a course, or diagnose students' strengths and weaknesses according to specific linguistic categories, among other purposes". Therefore, initially determining the purpose of an assessment instrument is significant to construct assessments in a specific context. In accordance with this perspective, the types of assessment deserve a substantial amount of attention in education (Brown & Knight, 2012).

There is no doubt that assessment in foreign language teaching has been influenced by educational revisions applied during COVID-19 pandemic like all the other educational areas. Especially with the compulsory and sudden prevalence of distance education, assessment process has become a significant concern in education. Although alternative assessment types in language teaching were stressed previously parallel to the modern language teaching methodologies (Liddicoat & Scarino, 2013; Darling-Hammond, 2013; Shrum, 2015; Shohamy, 2020), the unexpected pandemic conditions have led to drastic divergence from the previously applied traditional assessments. In this vein, the perspectives of pre-service teachers are of crucial importance in order to evaluate the efficiency of current application of alternative assessment types in language teacher education. Therefore, this research aims to reveal the pre-service ELT teachers' attitudes towards alternative assessment types in comparison with traditional assessment types, specifically in distance education.

2. Review of Literature

2.1. Types of assessment according to their purposes

Literature mainly stresses five types of assessment tests used in foreign language learning and teaching. The types of the tests are mainly named according to the purpose of their use. To start with, commonly focusing on the educational process and conducted at the end of a specific training process, achievement tests are used to reveal whether the objectives of the education process have been acquired or not (Paris et al., 1991). Though the results of achievement tests might be used for diagnostic purposes, they are mainly intended to measure the success of test takers at the end of the training process. Diagnostic tests, on the other hand, attempt to determine possible strengths and weaknesses of the test takers to revise the curriculum (Sun & Suzuki, 2013). The results of the diagnostic tests can also be utilized to track the students' learning in sequence, which might lead to proper courses for the improvement of learning (Mohan, 2016). As the third type of assessment tests, Brown and Abeywickrama (2010) suggest placement test, and state that depending on how to use the test results, achievement and proficiency tests can also serve as placement test. The main purpose of placement tests is to correctly place the learners into a course or level. If there is a need to test the learners' overall

ability rather than discrete abilities in a language, proficiency tests are preferred whereas aptitude tests are used to predict and measure potential capacity or general ability of learning a foreign language.

2.2. Comparison of traditional and alternative assessment

Tests in traditional terms have been the most commonly applied instrument in the assessment process due to their practicality. However, the argument during 1990s started to question the notion that all skills could be measured by traditional tests and as a result, the need for an alternative to traditional assessment emerged (Huerta-Macias, 2002). As indicated in Table 1 below, there are some generalized assumptions about the features of traditional and alternative assessments. However, it is useful to consider that there is no clear distinction between the two types in the process of educational application since institutions, schools or individual teachers might benefit from the combination of them according to their contextual needs.

Table 1. *Traditional and alternative assessment*

Traditional Assessment	Alternative Assessment
One-shot, standardized exams	Continuous, long-term assessment
Timed, multiple-choice format	Untimed, free-response format
Decontextualized test items	Contextualized communicative tasks
Scores sufficient for feedback	Individualized feedback and washback
Norm-referenced scores	Criterion-referenced scores
Focus on discrete answers	Open-ended, creative answers
Summative	Formative
Oriented to product	Oriented to process
Non-interactive performance	Interactive performance
Fosters extrinsic motivation	Fosters intrinsic motivation

(Adapted from Brown and Abeywickrama, 2010, p.17)

Along with the features of assessment types stated in the table, alternative assessment is also mentioned as "authentic" assessment in literature since alternative assessment provides an accurate evaluation of what the learners have learnt by focusing on the application of the acquired knowledge (Palm, 2014). Moreover, alternative assessment types are naturally categorized as formative assessments because they are conducted through process evaluation of the learners' understanding, learners' academic progress and learners' needs.

2.3. Distance education and alternative assessment

Though previously investigated with a multidimensional perspective in the literature, distance education drew the attention of the researchers due to the compulsory switch from traditional to online education during the pandemic period (COVID-19). Therefore, studies focusing on the field of assessment in distance education are of great importance to sustain the quality of the educational activities. Dikli (2003) discussed the advantages and disadvantages of assessment types in distance

education, and suggested that there should be a balanced approach between traditional and alternative assessment. The importance of issues such as content, context and audience in the process of deciding assessment strategy were also stressed in this study.

Sapriati and Zuhairi (2010) conducted research to address the application of computer-based testing in online education, or distance education as stated in the original research. They started a pilot project to develop a system and program to use computer-based testing method. By using the computers in networked configuration, they gradually improved the distance testing and test item bank at their examination center.

Oliver (2015) investigated the efficiency of technology driven assessment tools to replace the traditional, paper-based assessment types in the field of theology. In the study of Oliver (2015), it was stressed that instructors need to integrate technology effectively into the educational processes such as teaching, learning and assessment. As a result of the document analysis, ten documents were assessed as possible alternatives to traditional assessment. This research suggested what alternative assessment methods could be possibly used as an alternative by evaluating technology-based multimedia documents. The research concluded that alternative assessment instruments be integrated into higher education.

Another research conducted by Ahmad (2019) focused on the alternative assessment in distance education applied during the spread of coronavirus disease in Indonesian context. According to this study, emergency application of distance education due to the COVID-19 disease resulted in various problems in education, especially in the process of student learning assessment. Therefore, Ahmad (2019) investigated the alternative forms of assessment and evaluation which can be suggested as a reference for instructors to assess the learning outcomes of distance education in Indonesia. As a result, this study proposed several forms of assessment suitable for distance education such as online-based assessment, portfolio assessment, and self-assessment.

In a recent research, Tasci (2021) investigated the English Language Teaching (ELT) students' perspective on emergency distance education during Covid-19 pandemic. By using mix-type research model, Tasci (2021) utilized a combination of descriptive analysis, T-tests, and one-way ANOVA to analyze the perspectives of ELT students regarding suitability, effectiveness, teachability, and study habits of learners in distance education. This research revealed that emergency distance education had some disadvantages such as technological issues, mental problems, lack of interaction, material problems, and lack of experience. On the other hand, flexibility of time and place, efficiency of time and money, and opportunity to improve digital literacy skills were listed as the advantages of emergency distance education as the perspectives of students indicated. The study concluded that distance education should be applied as an additional part of traditional education rather than replacing it.

The research in distance education has gained a dramatic pace during the Covid-19 pandemic due to the prevalence of emergency distance education. However, there is still a lack of research about assessment types in distance education, especially in Turkish context. The scarcity of similar research in the literature is the distinctive feature of the current study. By focusing on the perspectives of pre-service language teachers, this research is expected to provide an insight into the issues occurred during the compulsory switch from traditional to alternative assessment types. In this vein, this study aimed to find out the attitudes and perspectives of the final-year ELT students who are studying at a Turkish state university. Therefore, this research is planned to find possible solutions for the following research questions:

1. Are there notable differences between pre-service language teachers' perspectives on assessment types in the second and fourth grades?
2. Do pre-service language teachers' opinions of "traditional and alternative assessments" in face-to-face instruction and online learning differ significantly?

3. **Methods**

3.1. Research Model

This research utilized a mix-method approach to investigate the assessment perspectives of pre-service teachers. Mix-method research is an examination that mixes qualitative and quantitative methods. Researchers that use this approach assert that combining quantitative and qualitative methodologies gives more comprehensive insights than either methodology alone (Fraenkel et al., 2012). Therefore, this research included both quantitative data obtained from questionnaire and qualitative data from semi-structured interview sessions. This research model was preferred since it provided a broader understanding of pre-service language teachers' perspectives. By using this model, the data which could not be exposed via the questionnaire were discussed and elaborated with the participants during interviews.

3.2. Participants and Setting

All the participants of this research had had distance education through Zoom applications at least for the last three semesters when this study was conducted. Since the 4th and 2nd year students have experienced both advantages and disadvantages of emergency distance education, the current study was conducted with 4th year (number: 66) and 2nd year (number: 60) pre-service ELT teachers to reveal their perspectives of traditional and alternative assessment types in distance education especially applied during Covid-19 pandemic. Due to the demographic features of the department, the 4th year participants consisted of 48 female and 18 male students, and their ages ranged from 22 to 32. Similarly, the 2nd year participants included 40 female and 20 male students. In the participant selection process, convenience sampling method was employed since these participants were easily accessible for the researcher. Convenience sampling is often used because it is simple, inexpensive, and subjects are easily accessible.

Researchers may get information they need without a sample frame by using convenience sampling (Dornyei, 2007).

3.3. Data Collection Instruments and Procedure

As the current study employed a mix-method research procedure, both qualitative and quantitative data were collected through varying means. Initially, an adapted version of "Authentic Assessment Inventory for Goal Setting" developed by O'Malley and Pierce (1992) was sent to the all participants online in order to obtain data about two sets of assessment; traditional assessment (true-false test, multiple choice test, open-ended tests, and short answer test) and alternative assessment (electronic portfolios, rubrics, research, and reflection papers). The items of the scale had 5 possible degrees ranging from strongly agree (5) to strongly disagree (1). The contrasting items were converted before the reliability analysis. For reliability and validity concerns, content of the Likert-type questionnaire was revised in parallel with application of the instrument at different time intervals. Moreover, test-retest method was initially utilized to ensure the reliability of the instrument as suggested by Fraenkel et al. (2012), and the correlation between the applications of the instrument at different time intervals was found as .84. For the application of the questionnaire to collect the research data, SPSS version 23 indicated that Cronbach-Alpha of the adapted scale was .882 for alternative assessment types, .887 for traditional assessment types and .895 for the whole scale, which is accepted as highly reliable according to Fraenkel et al. (2012).

Following the questionnaire, the participants were involved into online semi-structured interviews to answer the following questions:

1. Did you have difficulty in getting used to online education and online assessment?
2. What kind of assessment do you prefer? Why?
3. What are the advantages and disadvantages of alternative assessment types?
4. Which one do you think assess the learners more objectively: alternative or traditional? Why?
5. What kind of assessment do you prefer in distance education and in-class education? Why?

3.4. Data Analysis

The data in this research required both qualitative and quantitative data analysis. Firstly, the data from the questionnaire was analyzed via SPSS version 23 software, and the tables of descriptive analysis and independent sample T-test were obtained. The items were converted to same direction before the analysis procedure to prevent possible statistical miscalculations. Then, the data collected through semi-structured interviews were subjected to content analysis by using thematic coding. The themes that emerged at this point were connected to the research questions and survey results in order to triangulate the data and guarantee its credibility. Additionally, the findings were supported by direct quotes from the participants for the same reason.

3.5. Research Ethics

We affirm that we adhere to the standards of unbiased reporting. The study was approved and registered by Ethical Board of Nevsehir Haci Bektas Veli University at the meeting held on December 28th, 2021, with the corresponding decision number of 2021.12.438.

4. Findings

The answer given by 126 pre-service language teachers to the questionnaire indicated that there is a significant difference between 2nd and 4th grade students in terms of their perspective of using traditional and alternative assessment in distance/online education. Table 3 indicates that 2nd grade students favored the use traditional assessment types, specifically true-false, multiple choice, open-ended tests, and short answer tests (Mean: 3.35, N: 60) whereas 4th grade students gave more credit to alternative assessment like electronic portfolios, rubrics, research, and reflection papers (Mean: 3.81, N: 66). As for the application of assessment depending on the language skills being taught and integration of assessment types, both of the groups were found to have closer mean values as seen in the Table 2 below.

Table 2. *Descriptive statistics of 2nd and 4th grade pre-service teachers' perception*

	Grade	N	Mean	Std. Deviation	Std. Error Mean
Traditional Assessment	2nd grade	60	3,3500	1,35077	,17438
	4th grade	66	2,5455	,99509	,12249
Alternative Assessment	2nd grade	60	1,9833	1,14228	,14747
	4th grade	66	3,8182	,92668	,11407
Skill-based Assessment	2nd grade	60	3,0000	1,05766	,13654
	4th grade	66	4,2273	,90801	,11177
Integration of Assessment Types	2nd grade	60	4,0000	1,00844	,13019
	4th grade	66	4,3485	1,08829	,13396

After getting the descriptive statistics, an independent sample T-test was conducted to see if 2nd grade and 4th grade groups differed on four themes: traditional assessment in distance education, alternative assessment in distance education, skill-based assessment, and integration of assessment types. As shown in Table 3, Levene's test for equality of variances is not significant at p<.05 level for skill-based assessment and integrated assessment, which means that "equal variances assumed" lines for these data sets were reported for *t*-value and *p*-value (Larson-Hall, 2015). Table 4 contains the results of independent-samples t-test.

Table 3. *The results of independent sample t-test*

Variable	95% CI	Mean Difference	N1/N2	t-value	p-value	Effect size (Cohen's *d*)
Traditional assessment	.39, 1.22	.80	60/66	3.83	.001	.68
Alternative assessment	-2.2, -1.47	-1.83	60/66	-9.94	.01	1.76
Skill-based assessment	-1.57, -.88	-1.22	60/66	-7	.08	1.28
Integrated assessment	-.72, .02	-.35	60/66	-1.85	.065	.32

Regarding traditional assessment types (2nd grade mean= 3.35, sd= 1.35, N= 60; 4th grade mean= 2.54, sd= .99, N= 66), the %95 CI for the difference in means is .39, 1.22 ($t = 3.83$, $p=$ .001). The null hypothesis that there is no difference in the mean scores of 2nd and 4th grade per-service teachers in terms of their perspective of traditional assessment could be rejected since there was a relatively high effect size for the difference between the groups ($d=$ .68).

For alternative assessment types (2nd grade mean= 1.98, sd= 1.14, N= 60; 4th grade mean= 3.82, sd= .93, N= 66), the %95 CI for the difference in means is -2.2, -1.47 ($t = -9.94$, $p=$ .01). The null hypothesis that there is no difference in the mean scores of 2nd and 4th grade per-service teachers in terms of their perspective of alternative assessment could be rejected since there was a very high effect size for the difference between the groups ($d=$ 1.76).

As for the skill-based assessment types (2nd grade mean= 3, sd= 1.05, N= 60; 4th grade mean= 4.23, sd= .91, N= 66), the %95 CI for the difference in means is -1.57, -.88 ($t = -7$, $p=$ .08). The null hypothesis that there is no difference in the mean scores of 2nd and 4th grade per-service teachers in terms of their perspective of skill based assessment types could be rejected since there was a very high effect size for the difference between the groups ($d=$ 1.28).

As for the integration of assessment types, (2nd grade mean= 4, sd= 1, N= 60; 4th grade mean= 4.35, sd= 1.08, N= 66) the %95 CI for the difference in means is -.72, -.02 ($t = -1.85$, $p=$ .07). For this category, the null hypothesis that there is no difference in the mean scores of 2nd and 4th grade per-service teachers in terms of their perspective of the integration of assessment types could "not" be rejected for two reasons. Firstly, there was a low effect size for the difference between the groups ($d=$ .32), and secondly, the CI of integration of assessment types passes through zero. This finding indicates that 2nd and 4th grade students had similar perspective only in this category.

As the second data collection instrument, the semi-structured interviews provided significant results to support the statistical data. Comparison of the color-coded content analysis of 2nd and 4th grade students' interviews revealed that most of the students in each group had difficulty in getting used to online assessment and online education. One of the participants (S1) stated:

> *"I have been using technology for so long, but even so, it was very challenging for me to be assessed online. I felt that there was no teacher near me. It made me feel insecure."*

Upon being asked which assessment types they prefer, most of the participants at 2nd grade stated that they are not familiar with the alternative assessment types such as e-portfolio, research, rubric, and reflection paper. However, they were familiar with the traditional assessment types since they were exposed to them in many occasions. One of the participants (S2) uttered that:

> *"I think we are all familiar with true-false test, multiple choice test, open-ended tests, and short answer test. I also believe that these types are not outdated yet, because they are used very commonly by our teachers. However, during the online education, these types did not work well because it was very easy to cheat during these kinds of exams."*

The participants of 2nd grade commonly stressed the practical aspects of the traditional assessment types in terms of application and grading processes whereas 4th grade students mostly focused on the positive features of the alternative assessment types such as fostering creativity, increasing learner autonomy, and leading to real learning rather than memorization. However, as for the objectivity of the assessment types, both 2nd and 4th grade students agreed that traditional assessment types are more objective in grading the questions, and alternative assessment types might lead to teacher subjectivity. As for this issue, one participant at 2nd grade (S3) stated that;

> *"During online lessons, I wrote lots of reflection papers but I think our teacher graded my paper randomly because there was no correct answer for this kind of assessment, and it is very subjective."*

Similarly, another participant (S4) compared the assessment in online and face-to-face to education by giving specific examples from their own lessons. Especially in online lessons, alternative assessment types were commonly preferred by 4th grade students as this participant (S4) explained:

> *"As a prospective language teacher, I am aware that using rubric or portfolio might require extra effort of the teacher, but still using such kind of alternative assessment types might also motivate students to be more organized and autonomous. Moreover, during the online lessons I felt more motivated to conduct research rather than answering some multiple-choice questions."*

The content analysis of the interview revealed some main themes about traditional and alternative assessment types as summarized below in Table 4.

Table 4. *Common themes revealed in content analysis*

2nd grade students	4th grade students
- Multiple choice and true/false tests are the most practical way of assessment for both in-class and distance education. (45 students) - Open-ended tests and short answer test are more suitable for face-to-face education. (39 students) - In general, traditional assessment types are more suitable in distance education since they are easy to score. (36 students) - Alternative assessment should be taught to the students because most of them are not familiar with this type. (30 students)	- E-portfolio, research, and reflection paper sound more academic than traditional assessment types. (51 students) - Rubric increases the objectivity of assessment process. (47 students) - Multiple choice and true/false tests are not reliable in online education though they are more practical. (38 students) - Different assessment types might be useful for different purposes so students need to integrate them. (35 students)

5. Discussion, Conclusion and Suggestions

The current study focused on how alternative and traditional assessment examples are internalized by pre-service language teachers and whether there is a difference between their perception resulting from distance and in-class education. Stated in findings, there were meaningful differences between the study and control groups in three categories as the effect sizes were found to be high. Students' familiarity with alternative assessment types were observed to increase parallel to their educational improvement. Indeed, the results indicated that 2nd and 4th grade students had different perspectives of traditional, alternative and skill-based assessment whereas both there was no difference between the groups in terms of the integration of assessment types according to the curricular needs. In this vein, integration of the assessment types refers to the balance between the two assessment types and this finding complies with the Dikli (2003) who discussed the advantages and disadvantages of assessment types and stressed the importance of a balanced approach in distance education.

As the examples of authentic assessment, this study included electronic portfolios, rubrics, research, and reflection papers as an alternative to traditional true-false test, multiple choice test, open-ended tests, and short answer test. Similarly, Oliver (2015) investigated the influence of using technology-based multimedia documents as alternative assessment instruments. The results of the current study match up with Oliver (2015) since the findings revealed the importance of the integration of alternative assessment instruments. Similarly, the results of the current study also comply with Ahmad (2019) who suggested several forms of alternative assessment suitable for distance education such as online-based assessment, portfolio assessment, and self-assessment. However, alternative assessment forms in this study were electronic portfolios, rubrics, research, and reflection papers, and mostly different from the ones suggested by Ahmad (2019).

To sum up, distance education has become an inevitable component of education in many fields, and assessment forms are to be modified according to this sudden and unprecedented era. In this vein, pre-service teachers are of great importance since they need to improve their assessment skills both in in-class and distance education. As Sapriati and Zuhairi (2010) concluded, the integration of technology into progress of alternative assessment forms might be suggested to increase the assessment literacy of

pre-service teachers. Moreover, as revealed in this research, pre-service language teachers become more familiar with the alternative assessment forms as their grade increases.

In this light of the results of this research, it can be suggested that alternative assessment might be introduced to the pre-service teachers at the earlier stages of their education. Alternative assessment types are relatively more recent and student-centered than traditional ones, and are getting more and more prevalent in educational contexts. Therefore, the integration and utilization of alternative assessment types in the current curriculum of ELT departments alongside the traditional ones would be especially useful to increase students' awareness. Moreover, especially in distance education, student teachers should be guided to improve their skills in alternative assessment types, which would also increase the quality of the assessment practices of pre-service language teachers after their graduation.

References

Ahmad, I. F. (2020). Alternative assessment in distance learning in emergencies spread of coronavirus disease (Covid-19) in Indonesia. *Jurnal Pedagogik*, *7*(01), 195-222.

Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (Vol. 10). Pearson Education.

Brown, S., & Knight, P. (2012). *Assessing learners in higher education*. Routledge.

Darling-Hammond, L. (2013). Performance-based assessment and educational equity. In H. R. Hollins (Ed.), *Transforming curriculum for a culturally diverse society* (pp. 245-272). Routledge.

Dikli, S. (2003). Assessment at a distance: Traditional vs. alternative assessments. *Turkish Online Journal of Educational Technology-TOJET, 2*(3), 13-19.

Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*. McGraw-hill.

Huerta-Macias, A. (2002). Alternative assessment: Responses to commonly asked questions. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 338-343). Cambridge University Press.

Larson-Hall, J. (2015). *A guide to doing statistics in second language research using SPSS and R*. Routledge.

Liddicoat, A. J., & Scarino, A. (2013). *Intercultural language teaching and learning*. John Wiley & Sons.

Mohan, R. (2016). *Measurement, evaluation and assessment in education*. PHI Learning Private Limited.

Oliver, E. (2015). *Alternative assessment for effective open distance education.* [Unpublished doctoral dissertation]. University of South Africa.

O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners: Practical approaches for teachers.* Longman.

Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment, Research, and Evaluation*, *13*(1), 4.

Paris, S. G., Lawton, T. A., Turner, J. C., & Roth, J. L. (1991). A developmental perspective on standardized achievement testing. *Educational Researcher, 20*(5), 12-20.

Sapriati, A., & Zuhairi, A. (2010). Using computer-based testing as alternative assessment method of student learning in distance education. *Turkish Online Journal of Distance Education, 11*(2), 161-169.

Shohamy, E. (2020). *The power of tests: A critical perspective on the uses of language tests*. Routledge.

Shrum, J. L. (2015). *Teacher's handbook, contextualized language instruction*. Cengage Learning.

Sun, Y., & Suzuki, M. (2013). Diagnostic assessment for improving teaching practice. *International Journal of Information and Education Technology, 3*(6), 607.

Tasci, S. Evaluation of Emergency Distance Language Education: perspectives of ELT students. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi*, *11*(1), 286-300.