



Using Alternative Assessment to Assess Undergraduate EFL Students at Yemeni Universities: Challenges from the Perspectives of Faculty

Marwan Saed Saif Moqbel¹

¹ marwan_s1977@yahoo.com

Abstract

The present study aimed at investigating alternative assessment (AA) within the Yemeni university EFL classrooms in terms of identifying the challenges of using AA to assess undergraduate EFL students from the perspectives of the faculty. Data were collected, through a questionnaire, from (66) lecturers and professors of different ranks who teach EFL at the collegiate level in two Yemeni public universities and one national university. The results of the study indicated that faculty members of English departments at Yemeni universities face some challenges of using AA to assess undergraduate EFL students. These challenges belong to various factors, including the instructional environment, the students, the management of Yemeni universities, the faculty members, or the nature of AA. Statistically significant differences were found in the mean scores according to years of experience variable between respondents with less than 5 years and those with 5-10 years of experience in favor of those with less than 5 years of experience, but not according to the type of university, gender, and specialization variables. Accordingly, some recommendations were given to address the identified challenges, which might lead to a better implementation of AA in Yemeni university EFL classrooms.

Keywords: Alternative Assessment, Challenges, Yemeni Context, Undergraduate EFL Students, Students EFL Faculty Members

How to cite: Moqbel, M. (2020). Using Alternative Assessment to Assess Undergraduate EFL Students at Yemeni Universities: Challenges from the Perspectives of Faculty. *ELS Journal on Interdisciplinary Studies in Humanities*, 3(3), 440-458. **DOI:** <https://doi.org/10.34050/elsjish.v3i3.11077>

1. Introduction

The perspectives on assessment have changed a lot over the last few decades. Nowadays, assessment is no longer viewed as a process of testing students at the end of a unit of study or a course for the purpose of ranking them. Rather, it is viewed as an inseparable part of teaching and learning (Khattari & Sweet, 1996; Nasab, 2015). It is a complementary process of teaching that should be planned and designed in the light of the teaching methods and the learning objectives (Stoyhoff, 2012) in order to measure the level of students learning against the stated objectives and contribute to teaching improvement.

¹Ibb University, Yemen.

This connectedness between assessment and teaching and learning makes any change in teaching and learning and their theories inevitably leads to change in the views on assessment (Aliasin & Amanlu, 2017). Therefore, the recent sociocultural view of learning and the shift towards a constructivist learning paradigm in the last few decades spurred the appearance of new perspectives on assessment, such as assessment as a tool for learning or assessment culture reflected in the uses of alternative assessment (AA) (Gielen, Dochy, Dierick, 2003; Stoyhoff, 2012).

This constructivist approach to assessment, which is formative in nature, strongly emphasizes the integration of assessment and instruction and learning and the provision of constructive feedback to students on their learning, which can help them improve their learning and allow teachers to monitor and adjust teaching continuously in the light of the assessment results (Assessment Reform Group [ARG], 1999; Birenbaum, 2003; Monib, Karimi, & Nijat, 2020; Nasab, 2015; Stoyhoff, 2012).

Besides, the dissatisfaction with traditional assessment (TA) due to its limitations in EFL/ESL classroom - it does not indicate precisely what students can do in English or their ability to use English holistically in real-life situations - has contributed to the appearance of AA, which can alleviate the limitations of TA through providing opportunities for EFL students to use English in real-life situations and showing what they can do in English (Al-Mahrooqi & Denman, 2018; Grabin, 2007). Likewise, the shift in teaching paradigm – from teacher-centered to communicative student-centered teaching approaches and methods has called for a shift from TA to AA.

Moreover, two other factors relating to the nature of language learning and assessment have paved the way for the appearance of AA in EFL/ESL classrooms. The first factor is related to the holistic view of language. According to this view, being competent in a language requires knowledge of the various aspects of that language. Hamayan (1995) claimed that AA can reflect the holistic view of language as its procedures are based on the notion that the various aspects of any language are interrelated. The second factor is related to language performance and competence, which depends to a large extent on the purpose of language use and the context in which it is used. Therefore, it is useful to provide students with assessment opportunities that reflect the intended practices, which can be achieved through AA (Barlow & Coombe, 2000).

In brief, affected by some factors, the perspectives on assessment have changed. Birenbaum and Dochy (1996) and Segers, Dochy, and Cascallar (2003) described this change as a shift from a *testing culture* to an *assessment culture*. This change can also be described as a shift from *assessment of learning*, i.e., assessment to measure learning to *assessment for learning*, i.e., assessment to improve learning (Dochy & McDowell, 1997; Segers et al., 2003). Assessment in these senses is seen as an aid to learning and as a central component of what happens in the classroom and should take place simultaneously with learning activities (Butler, 1997; Khattri & Sweet, 1996; Wikström, 2007).

In language context, AA is defined as “a type of evaluation that directly evaluates learners’ language skills” and shows their ability to use it (Opp-

Beckman & Klinghammer, 2006, p. 105). AA takes various forms, such as performance-based tasks (e.g., projects, role playing, simulations, presentations, discussions/ debates, writing samples), portfolios, self-assessment (learning journals/logs, conferences, and checklists), peer assessment, teacher observations, etc. (Al-Mahrooqi & Denman, 2018; Birenbaum, 2003; Dochy & McDowell, 1997; North Carolina State Department, 1999; Stoyhoff, 2012).

Over the last few decades, the concern of researchers, educators, and teachers about AA in EFL/ESL classroom has increased and a lot of studies were conducted to investigate various aspects of AA, including its impact on EFL students' language learning, challenges of using AA, attitudes towards AA, etc. The relevant literature includes a lot of studies that highlighted the positive impact of AA on EFL students language learning (e.g., Ahangari, Rassekh-Alqol, & Hamed, 2013; Baleghizadeh & Zarghami, 2012; Baniabdelrahman, 2010; Ghaslani, 2015; Kalra, Sundrarajun, Komintarachat, 2017; Macari, 2017; Meihami & Varmaghani, 2013; Monib et al., 2020; Shokraie & Tabrizi, 2016).

AA is characterized by some features that make it a valuable assessment tool for EFL/ESL learners. In EFL/ESL context, AA is seen as an ongoing process that involves assessing the learners' progress and competence in language over time, depending in that on numerous sources of evidence and various non-conventional assessment strategies (Aliasin & Amanlu, 2017; Barlow & Coombe, 2000; Gielen et al., 2003; Hamayan, 1995), such as teacher observations, performance assessment, portfolio assessment, self-assessment, and peer assessment. This allows English teachers to address students' diverse learning styles (Al-Mahrooqi & Denman, 2018; Barlow & Coombe, 2000; North Carolina State Department, 1999) and to focus on any particular language aspect or skill (North Carolina State Department, 1999). In this, AA can show what students can do in English and offer a comprehensive image of their overall language competence (Al-Mahrooqi & Denman, 2018; Chirimbu, 2013; North Carolina State Department, 1999).

Besides, using AA in EFL classrooms, teachers can address realistic tasks that reflect everyday situations within meaningful contexts (Barlow & Coombe, 2000; North Carolina State Department, 1999). In this, teachers can provide EFL students with opportunities to make real use of English for actual purposes (Opp-Beckman & Klinghammer, 2006). Moreover, using AA in EFL classroom can create opportunities of collaboration and interaction among students similar to what is happening among people in the real world (North Carolina State Department, 1999). This, of course, can enhance students' language proficiency. AA can also enhance students involvement in and their responsibility for their own learning and assessment (ARG, 1999; Opp-Beckman and Klinghammer, 2006; Wikström, 2007), which has a positive influence on their language learning.

For these advantages and others of AA, many educators and teachers support the transition to AA. The relevant literature includes a lot of studies that highlighted the positive attitudes of teachers (e.g., Alkharusi, Aldhafri, Alnabhani, & Alkalbani, 2012; Chan, 2008; Ghaicha & Omarkaly, 2018; Gonzales & Aliponga, 2012) or students (e.g., Azarnoosh, 2013; Cornelius & Kinghorn, 2014; Elezovic, 2011; Moqbel, 2018) towards AA. Nonetheless, AA has its own opponents who expressed their reservation about the validity, reliability, and

objectivity of AA methods (El-Koumy, 2009; Huerta-Macías, 2002). The proponents of AA, in contrast, stated that AA is valid in terms of consequences, authenticity, cognitive complexity, significance, and efficiency and that reliability can be achieved through using a variety of assessment methods to collect data about student learning; appropriate rubrics for scoring; and more than a single observer, interviewer or reader (El-Koumy, 2009).

Similarly, Khattri and Sweet (1996) indicated that validity, reliability, fairness, and generalizability can be issues with performance assessment. Nevertheless, Khattri and Sweet argued that these issues can be addressed. While validity can be resolved through matching assessment tasks to curricular areas, reliability can be achieved through using standardized tasks, training scorers, and establishing explicit scoring criteria as Khattri and Sweet argued. Regarding fairness issue, it can be of a particular concern if assessment is used for student certification and sorting as Khattri and Sweet stated. With respect to generalizability issue, it can be addressed by using multiple examples of student work as Khattri and Sweet argued. For Herman et al. (1992), to ensure validity, consistency, and fairness in AA; teachers should establish scoring criteria that should be “well-conceived, explicitly defined, and consistently applied” (p. 45).

Gielen et al. (2003) argued that the new assessment culture, i.e., AA, cannot be evaluated solely on the basis of pre-era criteria. Gielen et al. suggested replacing the *psychometric* criteria used to evaluate the quality of TA with *edumetric* criteria. This can be achieved through addressing four criteria: the validity of assessment tasks (i.e., judging how well assessment matches the content and cognitive specifications of the competency measured), the validity of the assessment scoring (i.e., searching evidence for the appropriateness of the inference from the performance to an observed score), the generalizability of the assessment (i.e., searching for evidence of the appropriateness of the inference from the observed score to a conclusion about expected performance in the construct domain), and the consequential validity of assessment (i.e., investigating if the actual consequences of assessment are also the expected consequences).

Practically, using AA and integrating its methods into classroom are still a big challenge for teachers and students. Reviewing the relevant literature (Abbas, 2012; Al-Ruqeishi & Al-Humaidi, 2016; Butler, 1997; DeLuca, Luu, Sun, & Klinger, 2012; Demir, Cynthia, & Başboğaoğlu, 2018; Grabin, 2007; Janisch, Liu, & Akrofi, 2007; Lombardi, 2008; Metin, 2013; Webb & Jones, 2009; Yang, 2008), the researcher found that there are notable obstacles and challenges impeding teachers' use of AA in their classrooms, including EFL classroom. Some of these obstacles and challenges are related to the nature of AA itself in terms of the difficulty of designing authentic tasks, preparing AA tools such as rubrics, incorporating AA into classroom, implementing AA in large-enrollment courses, administering AA, and the plenty of time required to implement AA or to establish an appropriate culture in the classroom community for AA.

Obstacles or challenges that may impede the implementation of AA in EFL classroom may also be related to students. EFL students are used to TA and may be unprepared well for AA methods. This may make them feel uncomfortable to be assessed through AA. Besides, EFL students may not have the necessary

skills or language proficiency required to carry out AA tasks. The discrepancy of level among students in a single class can be an obstacle to implementing AA methods as some researchers reported. Moreover, EFL teachers may not be motivated to use AA in their classrooms due to their lack of faith in AA; lack of time, knowledge, and competence to use AA; lack of good models on how to use AA techniques in different situations; and their concerns about the subjectivity of AA methods as some researchers stated.

Furthermore, the classroom environment can be an obstacle to using AA in EFL classroom in terms of lack of administrative support and resources required for teachers to adopt AA methods and in terms of the size of classroom, the large number of students, and the existing classroom culture which supports the use of TA activities rather than AA tasks. Besides, the system of evaluation, being preoccupied with test scores, can be an obstacle to using AA in EFL classroom.

Considering the issues raised against AA and the challenges or obstacles to using AA in classroom, EFL teachers may think that it is better to keep on using TA to assess EFL students' language learning and competence rather than using AA in order to avoid facing the obstacles and difficulties associated with using this type of assessment. In this regard, the National Research Council [NRC] (2001) argued that any assessment operates within constraints that can limit its ability to provide useful information about student learning. Thus, as AA methods can provide precise information about student language performance and competence, constraints and challenges of using this type of assessment methods should be eliminated and removed, which is possible, in order to the full potential of such effective type of assessments is to be realized.

In agreement with this view of NRC (2001), the researcher thinks that AA is such an effective tool, a fact which has been highlighted by many previous studies, that should be used to assess undergraduate EFL students at Yemeni universities and whatever obstacles or challenges identified in this regard, they should be eliminated and removed, which may lead to a better assessment process in EFL classrooms at Yemeni universities.

The current study aims at identifying the challenges of using AA to assess undergraduate EFL students at Yemeni universities from the perspectives of the faculty, giving some recommendations with a view to overcoming these challenges, which may lead to a better implementation of AA. The study also addresses the following questions:

1. What are the challenges of using AA to assess undergraduate EFL students at Yemeni universities from the perspectives of faculty?
2. Are there any statistically significant differences in their responses according to the variables of *type of university* (public university & national university) and *gender*?
3. Are there any statistically significant differences in their responses according to the variables of *specialization* (linguistics, literature, and applied linguistics) and *years of experience* (less than 5 years, 5 to 10 years, more than 10 years)?

The significance of the current study emerges from its focus on a modern trend of assessment, namely AA. The study adds to the growing body of literature on the process of assessing EFL students in general and assessing undergraduate EFL students at Yemeni universities in particular. It tries to identify the challenges of using this type of assessment to assess undergraduate EFL students at Yemeni universities, offering some recommendations to overcome these challenges. This may motivate Yemeni universities to rethink about the process of assessing EFL students, which depends mainly on TA and to work on overcoming any challenges or dealing with any concerns in this regard to ensure that faculty members can use AA in their EFL classrooms.

2. Methodology

The current study is descriptive in nature. It aims at identifying the challenges of using AA methods to assess undergraduate EFL students at Yemeni universities from the perspectives of faculty. It is a quantitative and qualitative study. The quantitative part is reflected in the responses to the (21) three-point Likert items of the questionnaire, while the qualitative part is reflected in the responses to the request to add any other challenges or concerns of using AA to assess undergraduate students of English departments at Yemeni universities that were not mentioned in the questionnaire.

2.1. Participants

The respondents of the present study are (66) lecturers and professors of different ranks from English departments of two Yemeni public universities, namely Ibb University and Aden University and one national university, namely University of Sciences and Technology. Table (1) shows the sample of the study and their distribution according to the study variables.

Table 1. Number of respondents according to the study variables

Variable		No.	Variable		No.
University	Public	48	Gender	Male	49
	National	18		Female	17
Specialization	Literature	17	Years of Experience	Less than 5 years	24
	Applied linguistics	37		5-10 years	28
	Linguistics	12		More than 10 years	14

2.2. Instrument

Before developing the study instrument to collect the data for the current study, the researcher first conducted an exploratory survey on a group of faculty members of English departments from different universities asking them about the possibility of using AA methods to assess undergraduate EFL students at Yemeni universities and the possible challenges that may be faced. In the light of the results of the exploratory survey and literature review, the researcher developed the main instrument of the study in the form of a questionnaire which included (21) three-point Likert items. Besides, the researcher requested the respondents to add any other challenges in this regard. To estimate the reliability

coefficient of the study instrument and to measure its reliability, the researcher used Cronbach's alpha coefficient. The value of alpha coefficient achieved was (0.85), which indicates that the instrument was highly reliable.

2.3. Procedures

First, the researcher conducted an exploratory survey on a group of faculty members of English departments from different universities asking them about the possibility to use AA methods to assess undergraduate EFL students at Yemeni universities and the challenges of this process. Then, in the light of the survey and literature review, a questionnaire of (21) three-point Likert items was prepared and administered to the targeted sample during the academic year 2019-2020. Before administering the study instrument, it was verified for its content and face validity through giving it to four senior colleagues from Ibb University and two more experts for other universities. Different ways were used to reach the respondents. Hard copies of the questionnaire were distributed to the faculty members of English departments at Ibb University and University of Sciences and Technology – Ibb and Taiz Branches. Yet, soft copies of the questionnaire were sent to a few respondents via email or WhatsApp. For faculty members of English departments at Aden University and University of Sciences and Technology – Sana'a, a questionnaire using *Google Forms* was created and the link was sent to them to fill in the questionnaire. Finally, the data collected via the questionnaire were computed and analyzed using the Statistical Package of Social Sciences (SPSS) for Windows (version 25.0) to answer the study questions.

3. Data Analysis

3.1 Study Limitations

The scope of this study is limited in terms of the aspects investigated. It is based on the self-report of the respondents as expressed in their responses to the (21) three-point Likert items of the questionnaire and the request to add extra challenges of using AA not mentioned in the questionnaire. Besides, the respondents of this study were limited to faculty members from English language departments of two Yemeni public Yemeni universities (one from the north and one from the south of the country) and one national university that has branches in several governorates. Thus, caution should be considered when making generalizations from the results to other contexts.

3.2 Technique of Data Analysis

The data collected via the questionnaire were computed and analyzed using the Statistical Package of Social Sciences (SPSS) for Windows (version 25.0). The researcher calculated the descriptive statistics (means, standard deviations, and percentages) for each item of the instrument. Besides, the researcher used *t*-test to identify if there were statistically significant differences in the faculty' responses to the challenges of using AA according to the variables of type of university and gender. In addition, the researcher used the one-way analysis of variance (ANOVA) to identify if there were statistically significant differences in the faculty' responses to the challenges of using AA according to the variables of

specialization and years of experience. The significance level in this study was set at $P < 0.05$.

For statistical analysis, the challenges investigated were categorized into three levels (high, medium, and low) according to the mean scores they received. This categorization is calculated through identifying the difference between the high value in the Likert scale (i.e., 3.00) and the low value (i.e., 1.00) and then dividing the figure by the number of levels (i.e., 3) to get (0.66). This figure is used to create the three distinct levels used to describe the means that the challenges under investigation received:

- From 1.00 to 1.66 indicates a low-value mean.
- From 1.67 to 2.33 indicates a medium-value mean.
- From 2.34 to 3.00 indicates a high-value mean.

4. Findings & Discussion

RQ #1: What are the challenges of using AA to assess undergraduate EFL students at Yemeni universities from the perspectives of faculty?

To discuss the results of the first question and to identify the challenges of using AA to assess undergraduate EFL students, the researcher calculated the means, the standard deviations, and the percentages for each statement of the questionnaire and arranged them in descending order as shown in Table (2). The researcher also set a benchmark to judge the challenges under investigation. As the scale used to collect data was a three-point Likert scale, the minimum standard limit set for considering a challenge under investigation a real challenge was specified by getting (2) out of (3) degrees with a percentage of (66.66 %) out of the whole (100 %).

Table 2. Descriptive statistics of the questionnaire's items sorted in descending order

#	Rank	Item	<i>M</i>	<i>SD</i>	%	Estimation
17	1	The number of students in classroom is large.	2.73	0.67	90.90	high
20	2	There is a lack of administrative support to implement AA.	2.61	0.72	86.86	high
14	3	AA methods require a lot of time for preparation.	2.52	0.81	83.83	high
19	4	The instructional environment in Yemeni universities is not supportive to use AA.	2.48	0.79	82.82	high
2	5	Students lack the skills necessary to excel on AA methods.	2.45	0.73	81.81	high
3	6	Students lack language proficiency required to carry out AA.	2.45	0.75	81.81	high
7	7	Faculty members desire to assess certain aspects of students learning, such as grammar, vocab. etc. rather than holistically, which cannot be done through AA.	2.44	0.68	81.31	high
8	8	Faculty members are concerned with objectivity of AA.	2.39	0.65	79.79	high

21	9	The requirements and resources required to implement AA are scant.	2.38	0.80	79.29	high
6	10	Faculty's lack of time and heavy workload prevent them from integrating AA into their classrooms.	2.29	0.91	76.26	medium
9	11	Faculty members are afraid of changing the traditional methods of assessment.	2.26	0.83	75.25	medium
18	12	The system of evaluation in Yemeni universities is fixed by the management.	2.26	0.86	75.25	medium
15	13	Implementing AA limits the amount of curriculum that teachers can cover.	2.17	0.80	72.22	medium
4	14	Faculty members lack competence of integrating AA into their classrooms.	2.12	0.87	70.70	medium
1	15	Students may resist to be assessed through AA.	2.11	0.83	70.20	medium
5	16	Faculty members disbelieve the efficiency and importance of AA.	1.95	0.75	65.14	medium
16	17	It is difficult to grade students with AA.	1.83	0.90	61.10	medium
11	18	AA methods are difficult to prepare.	1.79	0.94	59.59	medium
13	19	AA methods are difficult to score.	1.77	0.86	59.08	medium
10	20	It is difficult to incorporate AA methods into classroom activities.	1.70	0.82	56.56	medium
12	21	AA methods are difficult to administer.	1.70	0.82	56.56	medium

Table (2) shows that the averages of the items of the questionnaire ranged from (1.70) to (2.73) with corresponding percentages ranged from (56.56%) to (90.90%). Table (2) also shows that out of the (21) challenges investigated, (15) challenges reached the minimum standard limit (i.e., 2.00 degrees out of 3.00) set for determining the challenges of using AA to assess undergraduate EFL students at Yemeni universities. Out of these (15) challenges, the respondents rated (9) challenges as high-level challenges and (6) as medium-level challenges.

As shown in Table (2), the (9) challenges of using AA to assess undergraduate EFL students at Yemeni universities that reached the minimum standard limit (i.e., 2.00 out of 3.00) and were rated as *high-level* challenges by the respondents are related to the large number of students in classroom ($M = 2.73$, percentage = 90.90%), the lack of administrative support to implement AA ($M = 2.61$, percentage = 86.86%), the long time required to prepare for AA ($M = 2.52$, percentage = 83.83%), the instructional environment in Yemeni universities which is not supportive to use AA ($M = 2.48$, percentage = 82.82%), students' lack of the skills necessary to excel on AA methods ($M = 2.45$, percentage = 81.81%), students' lack of the language proficiency required to carry out AA ($M = 2.45$, percentage = 81.81%), faculty members' desire to assess certain aspects of students learning, such as grammar, vocab. etc. rather than holistically ($M = 2.44$, percentage = 81.31%), faculty members' concerns about the objectivity of AA ($M = 2.39$, percentage = 79.79%), and the unavailability of the requirements and resources needed to use AA ($M = 2.38$, percentage = 79.29%).

Table (2) also shows that the (6) challenges of using AA to assess undergraduate EFL students at Yemeni universities that reached the minimum standard limit (i.e., 2.00 out of 3.00) and were rated as *medium-level* challenges

by the respondents are related to the faculty's lack of time and their heavy workload ($M = 2.29$, percentage = 76.26%), faculty members' fear of changing the traditional methods of assessment ($M = 2.26$, percentage = 75.25%), the system of evaluation in Yemeni universities which is fixed by the management ($M = 2.26$, percentage = 75.25%), the thought that implementing AA limits the amount of curriculum that teachers can cover ($M = 2.17$, percentage = 72.222%), faculty's lack of competence to integrate AA into their classrooms ($M = 2.12$, percentage = 70.70%), and the probability that students may resist to be assessed through AA ($M = 2.11$, percentage = 70.20%).

Besides, Table (2) reveals that six challenges of using AA to assess undergraduate EFL students at Yemeni universities investigated in the current study did not reach the minimum standard limit (i.e., 2.00 out of 3.00) and were rated as *medium-level* challenges by the respondents. These challenges are related to faculty members' disbelief in the efficiency and importance of AA ($M = 1.95$, percentage = 65.14%), the difficulty to grade students with AA ($M = 1.83$, percentage = 61.10%), the difficulty to prepare AA tasks and activities ($M = 1.79$, percentage = 59.59%), the difficulty to score AA tasks ($M = 1.77$, percentage = 59.08%), the difficulty to incorporate AA methods into classroom activities ($M = 1.70$, percentage = 56.56%), and the difficulty to administer AA methods ($M = 1.70$, percentage = 56.56%).

Regardless of the level and number of challenges facing teachers when using AA methods to assess their students learning, the findings of the current study are consistent with and support those of some previous studies. The findings of the current study are consistent with the findings of Butler (1997) who recognized large class sizes, lack of student motivation, teachers' lack of enthusiasm, the lack of time as obstacles that would make it difficult for university English teachers in Japan to use AA techniques in classroom. Similar to one finding of this study, Watt (2005) identified that the most common reason why teachers do not use AA methods is that they regard them as too subjective.

In consistency with the findings of this study, Janisch et al. (2007) referred to the lack of administrative support, concerns about student motivation, lack of resources, and limited English proficiency of students as obstacles to implementing AA in the classroom.

Similar results were obtained by Lombardi (2008), who found that faculty members are worried to use AA as preparing AA tasks will increase their workload and that AA is challenging to grade in a consistent manner and more subjective and requires a lot of time to manage. Similarly, Yang (2008) reported a number of obstacles to using AA in front of Taiwanese EFL primary school teachers, including difficulty of implementation, time constraints, difficulties with classroom management, doubts about the objectivity of grading, and heavy workload.

The findings of this study are also consistent with those of Abbas (2012), who reported the lack of technologies and a supportive classroom environment, the lack of time needed to implement AA methods, students being used to TA and are not expected to take responsibility for assessment, and the difficulty of

incorporating AA into classroom as the obstacles that face Iraqi English instructors in using the different methods of AA in their classrooms.

In support of the findings of the present study, DeLuca et al. (2012) identified some practical barriers to integrating assessment for learning into the classroom, including time, class size, and resources required for teachers to adopt assessment for learning practices. Similarly, Metin (2013) identified some difficulties in front of teachers to implement performance tasks, including crowded classroom, lack of time, insufficient learning environment and technological opportunity, low level of students, and difficulty to prepare assessment tools and to assess performance tasks objectively or to grade students with them.

Moreover, the finding of this study are consistent with those of some recent studies, such as Al-Ruqeishi and Al-Humaidi (2016), who identified teaching and administrative loads, lack of time, students' inability to handle such assessments or students' limited English proficiency to carry out AA tasks as the most imposing challenges to teachers to integrate AA into their classrooms in the Omani context; Demir et al. (2018), who indicated that insufficient amount of time, overcrowded class size, testing pressure, lack of support and knowledge, and difference in levels of students as barriers to implementing AA methods in classroom among Turkish teachers; and Ghaicha and Omarkaly (2018), who revealed that Moroccan EFL teachers face different types of challenges, including mainly time constraints, class size, and lack of training that impede them to use various AA strategies.

In brief, there are some challenges and barriers to using AA to assess the learning of students. These challenges belong to various factors, including the teacher, students, the nature of AA, the instructional environment, the system of evaluation, and the nature of the some courses. Therefore, the researcher, in agreement with Webb and Jones (2009), stresses the importance of some conditions to coexist or be developed in order to AA be implemented successfully. These conditions are related to students' beliefs, skills and language proficiency; teachers' beliefs and repertoire to develop AA practices; the system of evaluation; and the instructional environment.

RQ #2: Are there any statistically significant differences in the faculty' responses according to the variables of *type of university* (public university & national university) and *gender*?

To answer the second question in this study, the researcher used *t*-test for independent samples analysis to identify if there were statistically significant differences in the respondents' responses to the challenges of using AA under investigation according to the variables of type of university and gender.

Table 3. Results of *t*-test for the variables 'type of university' & 'gender'

Variable	N	M	SD	df	t	Sig. (2-tailed)	
University	Public	48	2.25	0.37	64	1.311	0.195
	National	18	2.10	0.48			
Gender	Male	49	2.22	0.37	64	0.385	0.702

Female 17 2.18 0.51

Table (3) shows that no statistically significant difference was found in the mean scores [$t(64) = 1.311, p = 0.195$] between respondents from the *public universities* ($M = 2.25, SD = 0.37$) and respondents from the *national university* ($M = 2.10, SD = 0.48$) at the (0.05) level of significance although the respondents from the national university rated the challenges of using AA under investigation lower than those from the public universities.

Similarly, no statistically significant difference was found in the mean scores [$t(64) = 0.385, p = 0.702$] between *male respondents* ($M = 2.22, SD = 0.37$) and *female respondents* ($M = 2.18, SD = 0.51$) at the (0.05) level of significance as Table (3) shows. This indicates that the variable of *gender* does not have any effect on the respondent's estimation of the challenges of using AA to assess undergraduate EFL students at Yemeni universities.

RQ #3: Are there any statistically significant differences in the faculty' responses according to the variables of *specialization* (linguistics, literature, and applied linguistics) and *years of experience* (less than 5 years, 5 to 10 years, more than 10 years)?

To answer the third question in this study, the researcher used one-way analysis of variance (ANOVA) to identify if there were statistically significant differences in the respondents' responses to the challenges of using AA under investigation according to the variables of *specialization* and *years of experience*.

Table 4. Results of one-way ANOVA analysis for the variable 'specialization'

Source of Variance	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21.420	2	10.710	0.144	0.866
Within Groups	4674.338	63	74.196		
Total	4695.758	65			

Table (4) shows no statistically significant differences were found in the mean scores of the respondents' responses according to the variable of *specialization* at the (0.05) level of significance. The *F*-value was (0.144), indicating no significant differences at $\alpha = 0.05$ since the *p*-value > 0.05 ($p = 0.866$). This means that the faculty members of different specialization have the same estimation of the challenges of using AA to assess undergraduate EFL students at Yemeni universities.

Table 5. Results of one-way ANOVA analysis for the variable 'years of experience'

Source of Variance	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	689.918	2	344.959	5.425	.007
Within Groups	4005.839	63	63.585		
Total	4695.758	65			

However, statistically significant differences were found in mean scores of the respondents' responses according to *years of experience* at the (0.05) level of significance as Table (5) shows. The *F*-value was (5.425), indicating significant

differences at $\alpha = 0.05$ since the p -value < 0.05 ($p = 0.007$). This means that the level of challenges facing faculty members of English departments at Yemeni universities using AA to assess their undergraduate EFL students varies according to their experience.

To determine the sources of these differences, the researcher used the *Tukey HSD* test. The results of the *Tukey HSD* test are shown in Table (6).

Table 6: Results of Tukey HSD test for the differences in the averages according to the variable 'years of experience'

(I) Experience	(J) Experience	<i>M</i>	Mean Difference (I-J)	Std. Error	Sig.
Less than 5 years	5-10 years	2.40	0.34779*	0.106	0.005
	More than 10		0.19813	0.128	0.274
5-10 years	Less than 5 years	2.05	-0.34779*	0.106	0.005
	More than 10		-0.14966-	0.124	0.455
More than 10	Less than 5 years	2.20	-0.19813-	0.128	0.274
	5-10 years		0.14966	0.124	0.455

Table (6) shows that the statistically significant differences were found in the respondents' responses between faculty members with years of experience less than five years ($M = 2.40$) and those with 5-10 years of experience ($M = 2.05$) in favour of those with less experience. This means that the faculty members of English departments at Yemeni universities with less experience face higher level of challenges when using AA to assess undergraduate EFL students due to their little teaching experience. This point was highlighted by Mertler and Campell (2005), who argued that teaching experience may have an effect on assessment competency.

Responding to the request at the end of the questionnaire regarding adding any other challenges or concerns about using AA to assess undergraduate EFL students at Yemeni universities that were not mentioned in the questionnaire, the respondents referred to various challenges that can be summarized in the following points:

- Faculty members are not encouraged to integrate AA methods into their classrooms.
- Faculty members are not trained on how to prepare AA tasks.
- Faculty members may have no desire to use AA to assess their students.
- The weak input of the faculties makes faculty members concentrate on the theoretical parts of the curriculum and on assessing students through TA rather than on the practical parts or skills and on using AA to assess their students learning.
- The nature of some courses may hinder faculty members to use AA methods.
- The current teaching paradigm is not in harmony with AA methods.
- Classrooms are not equipped to accommodate AA methods.
- Most AA methods allow cooperation among students, which, for some faculty

members, may run counter to the notion of assessment.

- Student's gender: Either male or female students do not like to perform in front of the other ones.
- Time limitation for covering the course is not enough to spend some of it in evaluating students using AA methods.
- Using some AA methods, teachers may not be able to control cheating among students.

Looking deeply at the additional challenges or concerns about using AA to assess undergraduate EFL students at Yemeni universities suggested by the respondents, the researcher found that they belong to the same sources of challenges mentioned in the questionnaire: the students, the faculty members, the instructional environment, the nature of AA, and the nature of the some courses.

5. Conclusion & Recommendations

The current study highlighted, in alignment with the literature, that there are some challenges of using AA to assess students of English language at Yemeni universities from the perspectives of faculty. These challenges may impede and discourage faculty in the departments of English language at Yemeni universities to use AA to assess their students language skills and knowledge. As the results of the study indicated, these challenges are related to the instructional environment, the students, the management in Yemeni universities, the faculty members, or the nature of AA.

In the light of the findings of this study, which highlighted the challenges of using AA to assess EFL students at Yemeni universities, the researcher suggests the following recommendations to overcome the challenges identified:

- Reducing the size of classes to facilitate integrating AA into EFL classrooms at university level.
- Considering the time and effort required from faculty members to integrate AA into their classrooms when determining load of faculty members.
- Providing administrative support for faculty members and their assistants to encourage and enable them to use AA to assess their EFL students.
- Encouraging faculty members to integrate various forms of AA into their EFL classrooms.
- Creating supportive environments through providing EFL classrooms with the necessary resources and equipment that can facilitate using AA.
- Broadening faculty members' knowledge about AA and its underlying theory.
- Training faculty members on creating AA tasks and activities and on integrating various forms of AA into their EFL classrooms.
- Providing good models on how to use AA methods in different situations to encourage EFL teachers to adopt AA in their classrooms.

- To ensure shift to AA practices, there should be first a shift on the part of the beliefs of teachers, students, university management, and the policy of assessment at university level.
- Training students on the skills required on their part to implement AA in EFL classrooms.
- Using appropriate AA methods which are appropriate for the students level of language proficiency.
- Eliminating any obstacles and barriers to using AA and reinforcing factors that can enhance the use of AA in EFL classroom.
- Concerns and resistance to culture change regarding assessment practices should not be underestimated.

The researcher suggests conducting further research on AA in EFL classroom in Yemen at school or university levels. Researchers can investigate the facilities required to implement AA in EFL classroom effectively or challenges to using AA in EFL classroom at school level. Researchers can also investigate the faculty members' desire to be trained on AA methods.

References

- Abbas, Z. (2012). Difficulties in using methods of alternative assessment in teaching from Iraqi instructors points of view. *AL-Fatih Journal*, 48, 23-45.
- Ahangari, S., Rassekh-Alqol, B., & Hamed, L. (2013). The effect of peer assessment on oral presentation in an EFL context. *International Journal of Applied Linguistics & English Literature*, 2(3), 45-53. doi:10.7575/aiac.ijalel.v.2n.3p.45
- Aliasin, S. & Amanlu, M. (2017). The effect of alternative assessment techniques on EFL learners' reading comprehension ability and self-efficacy in reading: The case of Iranian junior high school students. *Linguistics and Literature Studies*, 5(3): 160-168. doi: 10.13189/lls.2017.050302
- Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2012). Educational assessment attitudes, competence, knowledge, and practices: An exploratory study of Muscat teachers in the Sultanate of Oman. *Journal of Education and Learning*, 1(2), 217-232.
- Al-Mahrooqi, R. & Denman, C. (2018). Alternative assessment. In J. I. Liantas (Ed), *The TESOL Encyclopedia of English Language Teaching* (pp. 1-6). Hoboken, USA: Wiley & Sons, Inc. doi: 10.1002/9781118784235.eelt0325
- Al-Ruqeishi, M. & Al-Humaidi, S. (2016). Alternative assessment as perceived by EFL teachers. *The IUP Journal of English Studies*, 11(3), 88-101.
- Assessment Reform Group (ARG). (1999). *Assessment for learning: Beyond the black box*. Cambridge: University of Cambridge, School of Education.
- Azarnoosh, M. (2013). Peer assessment in an EFL context: Attitudes and friendship bias. *Language Testing in Asia*, 3(11). doi:10.1186/2229-0443-3-11

- Baleghizadeh, S. & Zarghami, Z. (2012). The impact of conferencing assessment on EFL students' grammar learning. *Profile Issues in Teachers' Professional Development*, 14(2), 131-144.
- Baniabdelrahman, A. A. (2010). The effect of the use of self-assessment on EF students' performance in reading comprehension in English. *TESL-EJ*, 14(2), 1-22.
- Barlow, L. & Coombe, C. (2000). Alternative assessment acquisition in the United Arab Emirates. Retrieved from <http://www.eric.ed.gov/>
- Birenbaum, M. (2003). New insights into learning and teaching and their implications for assessment. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimizing New Modes of Assessment: In Search of Qualities and Standards* (pp. 13–36). Dordrecht: Kluwer Academic Publishers.
- Birenbaum, M., & Dochy, F. (1996). *Alternatives in assessment of achievement, learning processes and prior knowledge*. Boston: Kluwer Academic.
- Butler, P. (1997). What do we do now? A survey of current assessment practices at KGU. In P. Butler (Ed.), *Issues in Alternative Assessment: The Japanese Perspective* (pp. 96-118). Nishinomiya: Language Center, Kwansai Gakuin University.
- Chan, Y-C. (2008). Elementary school EFL teachers' beliefs and practices of multiple assessments. *Reflections on English Language Teaching*, 7(1), 37-62.
- Chirimbu, S. (2013). Using alternative assessment methods in foreign language teaching. Case study: Alternative assessment of business English for university students. *Scientific Bulletin of the Politehnica University of Timișoara Transactions on Modern Languages*, 12(1-2), 91-98.
- Cornelius, S. & Kinghorn, O. (2014). Student attitudes towards self and peer assessment in Japanese university first year EFL classes. Retrieved from <https://www.kansai-u.ac.jp>
- DeLuca, C., Luu, K., Sun, Y., & Klinger, D. (2012). Assessment for learning in the classroom: Barriers to implementation and possibilities for teacher professional learning. *Assessment Matters*, 4, 5-29.
- Demir, M., Cynthia, A., & Başboğaoğlu, U. (2018). Comparative investigation of alternative assessment methods used in Turkey and United States elementary 4th grade mathematics curriculum. *International Journal of Educational Administration and Policy Studies*, 10(7), 72-82. Doi: 10.5897/IJEAPS2018.0561
- Dochy, F., & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), 279–298.
- Elezovic, S. (2011). University students' attitudes towards alternative assessment in FLT. *Proceedings of the 1st FLTAL*, 1058-1067. Retrieved from <https://core.ac.uk/reader/>

- El-Koumy, A. S. (2009). The effect of classroom performance assessment on EFL students' basic and inferential reading skills. Retrieved from <https://files.eric.ed.gov/>
- Ghaicha, A. & Omarkaly, E. (2018). Alternative assessment in the Moroccan EFL classrooms teachers' conceptions and practices. *Higher Education of Social Science*, 14(1), 56-68. Doi: <http://dx.doi.org/10.3968/10161>
- Ghaslani, R. (2015). The effect of self-assessment on Iranian EFL learners' reading comprehension skill. *Journal of Academic and Applied Studies*, 5(3), 1-9.
- Gielen, S., Dochy, F. & Dierick, S. (2003). Evaluating the consequential validity of new modes of assessment: The influence of assessment on learning, including pre-, post-, and true assessment effects. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimizing New Modes of Assessment: In Search of Qualities and Standards* (pp. 37-54). Netherlands: Kluwer Academic Publishers.
- Gonzales, R., & Aliponga, T. (2012). Classroom assessment preferences of Japanese language teachers in the Philippines and English language teachers in Japan. *MEXTESOL Journal*, 36(1), 1-19.
- Grabin, L. A. (2007). *Alternative assessment in the teaching of English as a foreign language in Israel* (Doctoral Dissertation). University of South Africa, South Africa.
- Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics*, 15, 212-226.
- Herman, J., Aschbacher, P., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Huerta-Macías, A. (2002). Alternative assessment: Response to commonly asked questions. In J. Richards & W. Renandya (Eds.), *Methodology in Language Teaching: An Anthology of Current Practice* (pp. 338-343). Cambridge: Cambridge University Press.
- Janisch, C., Liu, X., & Akrofi, A. (2007). Implementing alternative assessment: Opportunities and obstacles. *The Educational Forum*, 71(3), 221-230.
- Kalra, R., Sundrarajun, C., & Komintarachat, H. (2017). Using Portfolio as an Alternative Assessment Tool to Enhance Thai EFL Students' Writing skill. *Arab World English Journal (AWEJ)*, 8 (4), 292-302. doi: <https://dx.doi.org/10.24093/awej/vol8no4.20>
- Khattari, N. & Sweet, D. (1996). Assessment reform: Promises and challenges. In M. Kane & R. Mitchell (Eds.), *Implementing Performance Assessment Promises, Problems, and Challenges* (pp. 1-21). Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.
- Lombardi, M. M. (2008). Making the grade: The role of assessment in authentic learning. *EDUCAUSE Learning Initiative*. Retrieved from <https://www.researchgate.net/>

- Macari, I. O. (2017). Graduates' self-assessment of their oral presentations of group projects: An EFL case study in Romania. *Issues in Educational Research*, 27(3), 2017 453.
- Meihami, H. & Varmaghani, Z. (2013). The implementation of self- assessment in EFL writing classroom: An experimental study. *International Letters of Social and Humanistic Sciences*, 9(2013), 39-48.
- Mertler, C. & Campbell, C. (2005). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the assessment literacy Inventory*. A paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada (11-15 April, 2005).
- Metin, M. (2013). Teachers' difficulties in preparation and implementation of performance task. *Educational Sciences: Theory & Practice*, 13(3). 1664-1673. doi: 10.12738/estp.2013.3.1452.
- Monib, W., Karimi, A. & Nijat, N. (2020). Effects of alternative assessment in EFL classroom: A systematic review. *American International Journal of Education and Linguistics Research*, 3(2), 7-18.
- Moqbel, M. (2018). Self–assessment in EFL grammar classroom: A study of EFL learners at the Centre for Languages and Translation, Ibb University. *International Journal for Research in Education*, 42(2), 288-324.
- Nasab, F. (2015). Alternative versus traditional assessment. *Journal of Applied Linguistics and Language Research*, 2(6), 165-178. Retrieved from <http://www.jallr.com/>
- National Research Council (NRC). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press. <https://doi.org/10.17226/10019>.
- North Carolina State Department of Public Instruction. (1999). *Assessment, articulation, and accountability, 1999: A foreign language project*. Raleigh, NC: Author.
- Opp-Beckman, L. & Klinghammer, S. J. (2006). *Shaping the way we teach English: Successful practices around the world (Instructor's manual)*. Washington, DC: University of Oregon.
- Segers, M., Dochy, F., & Cascallar, E. (2003). The era of assessment engineering: Changing perspectives on teaching and learning and the role of new modes of assessment. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimizing New Modes of Assessment: In Search of Qualities and Standards* (pp. 1–12). Dordrecht: Kluwer Academic Publishers.
- Shokraie, S. & Tabrizi, A. (2016). The effect of portfolio assessment on EFL learners' L2 writing performance. *Journal of Applied Linguistics and Language Research*, 3(5), 321-331.
- Stoynoff, S. (2012). Looking backward and forward at classroom-based language assessment. *ELT Journal*, 66(4), 523-532. doi:10.1093/elt/ccs041

- Watt, H. (2005). Attitudes to the use of alternative assessment methods in Mathematics: A study with secondary mathematics teachers in Sydney, Australia. *Educational Studies in Mathematics*, 58, 21-44.
- Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy and Practice*, 16(2), 165-184.
- Wikström, N. (2007). *Alternative assessment in primary years of international baccalaureate education* (Master's thesis). The Stockholm Institute of Education, Sweden.
- Yang, T. (2008). Factors affecting EFL teachers' use of multiple classroom assessment practices with young language learners. *English Teaching & Learning*, 32(4), 85-123.