Humanities and Educational Sciences Journal
ISSN: 2617-5908 (print)
مجلـــــة العلـــــوم التربـــــوية والدراســـــات الإنســـــانيـــة
ISSN: 2709-0302 (online)

Alternative Assessment in English Departments at Yemeni Universities: Attitudes and Practices of Faculty Members(*)

Marwan Saeed Saif Moqbel
Asst. Prof. of Applied Linguistics
Center for Languages & Translation
Ibb University, Yemen

Abdu M. Talib Al-kadi
Center for Languages & Translation
Ibb University, Yemen

*) تاريخ تسلم البحث18/8/2020 تاريخ قبوله للنشر3/9/2020
**) موقع المجلة: http://hesj.org/ojs/index.php/hesj/index

اتجاهات أعضاء هيئة التدريس نحو التقييم البديل وممارستهم له في أقسام اللغة الإنجليزية في الجامعات اليمنية

مستخلص البحث:

هدفت هذه الدراسة إلى معرفة اتجاهات أعضاء هيئة التدريس نحو التقييم البديل ومدى استخدامهم له في أقسام اللغة الإنجليزية بالجامعات اليمنية، وعلاقة ذلك بخمسة متغيرات: نوع الجامعة، والجنس، والتخصص، وسنوات الخبرة، والرتبة الأكاديمية، واعتمدت الدراسة على بيانات تم جمعها من خلال إستبانة لعدد 66 عضو هيئة تدريس تم اختيارهم من جامعتين حكوميتين وجامعة أهلية، وبالرغم أن نتائج الدراسة أظهرت اتجاهات ايجابية لأفراد العينة نحو التقييم البديل إلا أن ممارستهم له كانت أقل من المستوى المقبول، كما أظهرت النتائج وجود فروق دالة إحصائياً في اتجاهاتهم نحو هذا النوع من التقييم وفقاً لمتغير الجنس دون المتغيرات الأخرى: نوع الجامعة، والتخصص، وسنوات الخبرة، والرتبة الأكاديمية، وفيما يتعلق باستخدام أعضاء هيئة التدريس للتقييم البديل، أظهرت النتائج وجود فروق دالة إحصائياً وفقاً لمتغير نوع الجامعة، أما متغيرات الجنس، والتخصص، وسنوات الخبرة، والرتبة الأكاديمية فلم تظهر فيها أي فروق ذات دلالة إحصائية. وبناءً على ذلك تم وضع بعض المقترحات والتوصيات لتعزيز استخدام التقييم البديل لمواكبة أساليب التقييم الحديثة المتبعة في كثير من الجامعات العالمية.

الكلمات المفتاحية: التقييم البديل، نظام التقويم، أعضاء هيئة تدريس اللغة الإنجليزية كلغة أجنبية، الجامعات اليمنية.

Alternative Assessment in English Departments at Yemeni Universities: Attitudes and Practices of Faculty Members

Marwan Saeed Saif Moqbel
Asst. Prof. of Applied Linguistics
Center for Languages & Translation
Ibb University, Yemen

Abdu M. Talib Al-kadi
Center for Languages & Translation
Ibb University, Yemen

Abstract:

The study aimed to uncover faculty members' attitudes towards alternative assessment (AA) and substantiate the differences in their attitudes, if any, based on five variables: university, gender, specialization, experience, and academic rank. It also elucidates the extent this cohort of informants adopts AA in their teaching and ascertains the differences in their uses according to the same variables. The study involved collecting data, through a questionnaire, from a sample of (66) lecturers and professors of different ranks who teach EFL at the collegiate level in Yemen. The study took place in two public universities along with a private one. The results of descriptive statistics and inferential tests (t-test, Pearson correlation, ANOVA) revealed that the respondents were positive towards AA but this was hardly reflected in their teaching. Significant differences in their attitudes towards AA were attributable to gender but not to university, specialization, magnitude of experience, or academic rank. In terms of usage, significant differences pertained to the variable of university and not to gender, specialization, experience, or academic rank. The findings rationalized incorporating AAs in EFL programs to enhance seeding new models of assessment that results in efficient learning and teaching.

Keywords: Alternative Assessment, Evaluation System, EFL Faculty Members, Yemeni Universities.

1. Introduction & Background

It is fair to state that assessment, being one of the rudiments of instruction, impacts how teachers teach and students learn. It goes hand in hand with teaching/learning; they are inseparable and in complementary coexistence (Al-Mahrooqi & Denman, 2018; ARG, 1999; Huerta-Macías, 2002). Hence, any change in one of these three elements requires a corresponding change in the other (Nasab, 2015). In this light, assessment is not an end in itself, neither is it less important than teaching. The traditional assessment (TA) focused mainly on the *product* but not *process* of learning. Assessment is no longer viewed as a process of testing that determines students' performance and ranking at the end of a unit of study or a course; it is rather mooted as an integral part of the teaching/learning process.

The limitations of the TA instigated educators to search for alternatives that alleviate such limitations. For one thing, TA measures students' abilities to recall information but it does not indicate precisely what students can do in the English language or their ability to use English holistically in real-life situations. That is, results of the TA imprecisely denote learners' performance, abilities or progress. It also causes stress and anxiety to students, and barely motivates them. Taken together, such weaknesses boiled down to a sense that TA is obsolete and unreliable measurements of the real performance of learners. Baillie (2004), among others, argued for a form of assessment that is consistent with the course content and learning objectives. Educators all over the World have been trying to find proper alternatives that accord with the emerging learning trends and teaching paradigms. This stimulated a radical shift to what has been termed *alternative assessment* (AA), also called 'performance assessment' or 'direct assessment' (Monib, Karimi, & Nijat, 2020).

Besides limitations of TA, the re-conceptualization of learning over the last few decades is a salient factor that gave way for AA. Believingly, effective learning takes place when learners construct their own knowledge. Learners are assumingly self-dependent and motivated to continually learn and broaden their horizons. They need to develop effective strategies and metacognitive skills to reflect on their own and others' learning (Segers, Dochy, & Cascallar, 2003). Likewise, shifts in teaching paradigms – from teacher dominance to learner centeredness and from whole-class teaching to individualized learning fueled rejection of the TA mode. Thanks to these paradigm

shifts, the role and nature of assessment in learning have radically changed (Tierney, 2006). Assessment is touted now as an integral part of teaching/learning (Stoynoff, 2012), aiming at facilitating student learning and improving the quality of teaching. This has inevitably capitalized on new assessment forms that involve students in the process of assessment (Grabin, 2007). To Wikström (2007), AA came as a response to the changes in education that encourage students to demonstrate their knowledge and skills instead of demonstrating comprehension of the acquired knowledge and skills.

Over the last three decades, AA has been used in various disciplines, including L2 pedagogy. A wealth of prior research has unearthed a great deal of issues related to AA in several contexts (American Federation of Teachers [AFT], 2008; Al-Mahrooqi & Denman, 2018; Burnaz, 2011; Monib et al., 2020). A general outcome ensued from such studies is a call for AA in place of the TA because the former widens the spectrum of assessment by bringing in a variety of activities that make learning more enjoyable and engaging (Abu Rezeq & Abu Taha, 2018; Ghaicha & Omarkaly, 2018; Moqbel, 2018; Öz, 2014; Purwanti, 2015; Shokraie & Tabrizi, 2016).

Despite recommendations of prior research to shift to the new assessment paradigm (i.e., AA), educators in the local EFL context continue using the traditional mode, which is mainly based on paper-and-pencil tests. Such methods of assessment generally give erroneous indicators of students' real potentials and skills (Iqbal & Manarvi, 2011; Ghaicha & Omarkaly, 2018; Yang, 2008). Believingly, ignoring the AA deprives learners from benefits that directly contribute to their learning. Hence, it is necessary to highlight its significance for EFL students and to offer a comprehensive image of their overall competence rather than routinely testing the acquired knowledge at the end of a semester or an academic year. In a bid to emulate the worldwide assessment system within ELT communities, this study explores the extent the EFL faculty members at the university level are familiar with AA in their teaching. Familiarity is discussed in terms of attitudes to and use of AA.

Objectives and Questions

The current study set out with two primary objectives in mind: (a) identifying the attitudes of faculty members of English departments towards AA and (b) illuminating the AAs they actually use in assessing their EFL undergraduates. With these objectives in mind, this attempt addressed the following questions:

1- What are the attitudes of the faculty members of English departments at Yemeni universities towards AA?
2- Are there any statistically significant differences in their attitudes to AA based on university type, gender, specialization, experience, or academic rank?
3- What is the mean score of their use of AA to assess their EFL undergraduates?
4- Are there any statistically significant differences in their use of AA based on university type, gender, specialization, experience, or academic rank?
5- Is there any significant relationship between their attitudes to and use of AA?

Significance of the Study

The significance of the current study lies in its focus on AA, which is a modern trend of assessment. It is profoundly helpful for pedagogues to make a radical change in the way they currently assess their learners – a change that corresponds to the prevalence teaching modes such as learner-centeredness, learner autonomy, and individualized language learning. The study is an addition to the growing body of literature on the topic within the area of EFL assessment in general and the use of AA to assess undergraduate EFL students at Yemeni universities in particular. Add to that surveying the attitudes of the staff towards AA and their use of various methods would enhance the transition from TA to AA with its merits for promoting English language education.

2. Literature Review

2.1. Alternative Assessment (AA)

Alternative assessment is generally defined as the "procedures and techniques which can be used within the context of instruction and can be easily incorporated into the daily activities of the school or classroom" (Hamayan, 1995, p. 213). Huerta-Macías (2002) refers to it as situations in which "students are evaluated on what they integrate and produce rather than on what they are able to recall and reproduce" (p. 339). Other researchers (e.g. Grabin, 2007; Herman, Aschbacher, & Winters, 1992) viewed it as strategies used to verify what students can do or produce. In another relevant conception of AA, Opp-Beckman and Klinghammer (2006) defined AA in language classroom as "a type of evaluation that directly evaluates learners' language skills" and shows their ability to use it (p105). Whatever the

definition, AA is a formative process that aims at enhancing learning rather than ranking students or finding out about the amount of knowledge they acquire (Țeican, 2016).

In other words, it is an assessment *for* learning rather than an assessment *of* learning. Assessment for learning, Little (2009) argued, is based on the belief that the right kind of assessment plays a crucial role in effective teaching and learning. AA requires new instructional and assessment roles for teachers (Herman et al., 1992) based on sharing the responsibility of assessment with students and engage them in self-assessment and/or assessing other's performance (Stoynoff, 2012). Janisch, Liu, and Akrofi (2007) view it from another angle - the constructivist view of learning. It is an instruction-driven process (Dochy & McDowell, 1997) in which students construct their own knowledge. It strongly emphasizes the integration of leaning, instruction and assessment, giving students roles that are more active in assessment practices (Assessment Reform Group [ARG], 1999). Besides enhancing students' responsibility for their learning and encouraging them to study in a more profound way, it involves constructive feedback to students on their progress and achievement.

2.2. Methods of Alternative Assessment

The diversity of teaching/learning process requires various assessment means to collect information about students' abilities and knowledge. This can be achieved through AA, which hosts a range of tools – sometimes called methods; some are simple such as checklists and some are complex such as portfolios. The latter may stretch over a semester or a year. AAs mainly include performance-based assessments (projects, role-play, demonstrations, presentations, interviews, discussions/debates, writing samples, reports, story of text retelling, cloze tests, and open-ended questions), teacher observations, portfolios, self-assessment (journals, learning logs, conferences, and checklists), peer assessment, and games. Teachers can select appropriate methods to assess their students' performance, taking into account students' age, level, ability and interest, aims of assessment, time and resources available for assessment.

AA has gained momentum in the worldwide context and there is increasing research works on the nature of this type of assessment, its principles, criteria, and its methods. Researchers have covered different areas related to AA including its implementation in EFL classrooms or its impacts on EFL learners' language skills. A plethora of studies has delved into AA in general (e.g., Abu Rezeq & Abu Taha,

2018; Chirimbu, 2013; Ghaicha & Omarkaly, 2018; Grabin, 2007; Öz, 2014). Some other studies focused on one technique of AA. For instance, self-assessment (e.g., Alibakhshi & Shahrakipour, 2014; Ghaslani, 2015; Honsa, 2013; Meihami & Varmaghani, 2013; Moheidat & Baniabdelrahman, 2011; Moqbel, 2018; Purwanti, 2015); peer assessment (e.g., Azarnoosh, 2013; Peng, 2009); portfolio assessment (e.g., Charvade, Jahandar, & Khodabandehlou, 2012; Shokraie & Tabrizi, 2016; Tavakoli & Amirian, 2012); performance-based assessment (e.g., El-Koumy, 2009; Yildirim & Orsdemir, 2013); conferencing assessment (e.g., Baleghizadeh & Zarghami, 2012). Some others focused on two techniques of AA, such as self-assessment and peer assessment (e.g., Birjandi & Tamjid, 2012) or portfolios and conferencing (e.g., Moradan & Hedayati, 2011). These studies, among others, highlighted the effectiveness of AA as an assessment tool integrated into L2 pedagogy.

2.3 Principles of Effective Alternative Assessment

Apart from the common principles of validity, reliability and objectivity, which are fundamental cornerstones of TA, are not key standards of AA. Grabin (2007) referred to eight different principles, which can contribute to the effectiveness of AA. The first principle relates to the *purpose of assessment*. Assessment should aim at improving learning, which requires teachers to provide students with opportunities to practice tasks and activities of various areas of performance, which can help students acquire a variety of important skills and improve their learning. The second three principles relate to the *nature of AA tasks and activities* and the *instructions* given: authenticity, feasibility, and clear instructions. AA should be based on authentic and real-world tasks and activities in terms of getting students to deal with meaningful situations similar to those of the real life. This can enable students to apply what they learn to real-life situations. In order to help students to achieve assessment tasks and activities successfully, teachers should provide students with clear instructions, criteria, and guidelines. Feasibility in assessment requires teachers to consider students' knowledge, level, ability, and the available time and resources when creating tasks and activities. Assessment methods should be practical and its cost should be acceptable as well.

Two other principles, reported by Grabin (2007), relate to the *nature of assessment*: continuity and connectivity with the other elements of the teaching and the learning processes. The principle of

continuity that AA is based on involves dealing with assessment as a continuous process rather than an intermittent process. This requires teachers to assess students and monitor their progress and learning over time rather than just at certain times. The principle of connectivity involves connecting assessment with the curriculum, the learning outcomes, and daily instruction and the assessment methods should be related to what teachers are teaching in the classroom.

The last two principles, referred to by Grabin (2007), involve *students*: centeredness on students and fairness. As AA resulted from the shift from teacher-centered to student-centered teaching approaches and methods, it is student-centered. Thus, AAs should work to facilitate and evaluate students' learning; and assessment tasks and activities should be built around topics and issues of interest to students. Fairness in assessment involves considering students' individual differences in terms of familiarity, level, interest, and motivation. It also involves having all students equal access to resources. Taking these principles into account, EFL teachers can make assessment more effective and successful, achieving the real purpose of assessment in EFL classrooms, i.e., measuring students' English language skills and knowledge and what they can do with it.

Comparing these principles with the principles of assessment reported in North Carolina State Department (1999), which are based on the assumption that the purpose of L2 instruction is to prepare students to use language with cultural understanding and knowledge in real-life contexts, the researchers found a lot of similarity. Like the principles of effective AA reported by Grabin (2007), the principles reported in North Carolina State Department (1999) are related to various aspects: the purpose of assessment, the nature of assessment, the nature of assessment tasks and activities, and the students.

The purpose of assessment should be clear as the first principle states. Regarding the nature of assessment, the principles state that assessment should be linked with curricular practices that are based on second/foreign language theory and with the goals and objectives of the course. Besides, assessment should be developmentally appropriate and conducted regularly and frequently, employing a broad range of assessment methods over time. Moreover, assessment should be (a) formative, occurring with instruction and learning, and (b) summative, assessing the degree to which final objectives have been met. Principles concerning the nature of assessment tasks and activities indicate that assessment tasks and activities should be

authentic and contextualized, including practical contexts and culturally appropriate situations. Tasks and activities should also be curriculum-embedded and part of the teaching and the learning processes.

Regarding students, the principles emphasize that assessment should be student-centered. Assessment should also encourage students to reflect on their own learning and progress, allowing them to demonstrate their ability to function in a variety of tasks and taking into account some factors when assessing students' language, including students' different learning styles, their level, and any other characteristics affecting their performance.

2.4. Advantages of Alternative Assessment

Because assessment influences students' motivation and their approach to learning (Baillie, 2004; Yin, 2006) and helps teachers refine their teaching (AFT, 2008), it involves timely and constructive feedback to students on their progress, which leads to active learning (Dochy & McDowell, 1997; Segers et al., 2003; Wikström, 2007). It also helps students recognize how to improve (ARG, 1999; Little, 2009; Spratt, Pulverness, & William, 2011), influences their intrinsic motivation positively (ARG, 1999; Spratt et al., 2011), and allows teachers to monitor and modify instruction continuously in light of the results of assessment (Nasab, 2015). Besides, the close alignment of instruction and assessment within authentic contexts helps teachers to identify what students know and can do (Janisch et al., 2007) and positively affects teaching by making it more related to real life experiences (Dochy & McDowell, 1997). Additionally, AA strategies are more interesting, meaningful, authentic, challenging and less threatening for students than TA (Dochy & McDowell, 1997; Struyf, Vandenberghe & Lens, 2001). To Yin (2006), designing and selecting appropriate assessment tools can contribute to the process of teaching as well as help students to develop lifelong learning and assessment skills.

In L2 classroom, AA has many advantages. Being a valuable language-learning tool (Cornelius & Kinghorn, 2014), it looks at learning as an integrative process and allows integrating various aspects of learning, which can enhance students' language proficiency (Hamayan, 1995). Besides, AA is an ongoing process that makes judgments on student's progress, competence, and proficiency in language over time based on multiple assessment tools. Such tools are non-conventional (Hamayan, 1995; Nasab, 2015; Tannenbaum, 1996),

adaptable for different situations (Tannenbaum, 1996) and can address students' different learning styles (North Carolina State Department, 1999). As such, AA informs teachers about what students can do in English. It also offers a comprehensive image of their overall language competence (Chirimbu, 2013; North Carolina State Department, 1999; Tannenbaum, 1996). To reiterate, using AA in English classroom makes students more involved in their own evaluation and increases their responsibility for their own learning (Opp-Beckman & Klinghammer, 2006; Wikström, 2007).

In a similar vein, Opp-Beckman and Klinghammer (2006) contend that AA is an objective-based process that can help teachers evaluate student performance using realistic tasks that reflect everyday situations within realistic and meaningful contexts. Towards this end, a set of criteria are used to describe the desired achievement, guide assessment and help students complete the assessment activities (Hamayan, 1995). AA enables teachers to provide their EFL students with opportunities to make real uses of the English. According to North Carolina State Department (1999), AA allows teacher to focus on any particular language aspect or skill by aligning assessment with instruction.

2.5. Attitude towards Alternative Assessment

Several studies have showed that EFL students have positive attitudes towards AA. For instance, Peng (2009) revealed that both high and low-intermediate students reacted positively to peer assessment and that their attitudes became more positive after experiencing this type of assessment. Similarly, Azarnoosh (2013) indicated that practicing peer assessment resulted in changing students' attitudes to a positive perception on it. Likewise, Cornelius and Kinghorn (2014) reported that first year Japanese university EFL students had positive attitudes towards self and peer assessment. In line with these studies, Elezovic (2011) considered university students' acceptance of AA in foreign language classroom. The findings revealed that the bulk of students endorsed AA techniques in foreign language classroom. Likewise, Burnaz (2011) maintained that Turkish university EFL students, in general, prefer portfolio assessment to TA because they believe that TA involves time pressure and leads to memorization and does not measure English speaking skills effectively.

Attitudes of teachers towards AA were the thrust of some other studies. For instance, Wikström (2007) showed that teachers consider

AA as an important element in teaching that has a positive impact on students' learning and view it as helpful for students to produce relevant and meaningful learning experiences. Similarly, Iqbal and Manarvi (2011) showed an agreement among university teachers on the benefits of AA and on using it in the Pakistani universities. Likewise, Ghaicha and Omarkaly (2018), revealed that EFL teachers in the Moroccan EFL public schools have positive attitudes toward AA although AA was not fully reflected in their actual assessment practices. In contrast, some studies displayed a resistance to AA. Teachers recruited in such studies generally rejected implementing AA (see Watt, 2005) and some teachers had a medium level of attitudes towards it (see Al-Nouh, Taqi, & Abdul-Kareem, 2014).

The study at hand, however, strengthens evidence on AA by considering the alignments and contrast of findings of prior research. It mainly examines the topic in a new relevant context to shed light on its significance and to bring to the foreground insightful ideas to promote AA for it spawns new opportunities of holistic assessment that covers a wide range of students' skills, abilities and knowledge. In other words, AA provides a more reliable assessment measurement to help pedagogues improve English pedagogy at the university level in Yemen.

3. Methodology

The current study is descriptive-analytical in nature. It describes the attitudes of faculty members of English departments at Yemeni universities towards AA and their use of its assessment methods to assess their under graduate EFL students. It employed the qualitative approach through a questionnaire composed of (37) three-point Likert items to collect the data required to answer the study questions.

3.1. Participants

The respondents of the present study were (66) faculty members of English departments from two public Yemeni universities, namely Ibb University and Aden University and from 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 Type	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
Rank	Lecturer	31	Assoc. Prof.		5
	Asst. Prof.	28	Prof.		2

3.2. Instrument

The questionnaire encompassed two sections developed in the light of literature review and the comments of the referees of the study instrument. The first section consisted of 18 three-point Likert items aiming at identifying the attitudes of the cohort of respondents in Table (1). The second section (19 three-point Likert items) aimed at identifying the faculty members' use of AA to assess their undergraduate EFL students. Cronbach's alpha coefficient was employed to estimate the reliability coefficient of the study instrument and to measure its reliability. The values of alpha coefficient were (0.76) for the entire instrument, (0.77) for the first section of the instrument, and (0.70) for the second section, indicating acceptable levels of reliability.

3.3. Procedures

Reviewing the related literature review, the researchers developed a questionnaire of 37 items on a three-point Likert scale. It 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 by four senior colleagues from Ibb University and two more experts (one from Sana'a University and one from Aden University). 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. A questionnaire using *Google Forms* was created and the link was sent to the respondents from Aden University and University of Sciences and Technology (Sana'a) to fill it in. Finally, the data collected were computed and analyzed using the Statistical Package of

Social Sciences (SPSS) for Windows (version 25.0) to answer the study questions.

4. Data Analysis

The data collected via the questionnaire were computed and analyzed using the SPSS. The descriptive statistics (means, standard deviations, and percentages) for each item were calculated. Besides, the researchers used *t*-test to identify significant differences in the respondents' attitudes towards and use of AA according to type of university and gender variables. Additionally, the ANOVA was used to detect significant differences in the respondents' attitudes towards and use of AA according to the variables specialization, multitude of experience, and academic rank. The significance level in this study was set at $P<0.05$. Moreover, Pearson correlation coefficient was computed to identify any significant relationship between respondents' attitudes towards and use of AAs.

For statistical analysis, the respondents' attitudes towards and use of AAs were categorized into three levels: high, medium, and low. This categorization was calculated by 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 of respondents' estimation of their use of AAs and their attitudes towards them.

- 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.

5. Findings & Discussion

RQ #1: What are the attitudes of the faculty members of English departments at Yemeni universities towards AA?

To answer this question, the means, standard deviations, and percentages of each item of the relevant section were calculated and arranged in a descending order as displayed in Table 2. It also shows the overall values.

Table 2: ***Descriptive Statistics of the Items Sorted in a Descending Order***

#	Rank	I think that …	*M*	*SD*	%	Estimation
18	1	AA is generally not useful.	2.97	0.17	98. 98	high
7	2	Using AA in the classroom provides opportunities for students to effectively cooperate and interact.	2.95	0.21	98.47	high
2	3	AA enhances students' learning motivation.	2.91	0.29	96.96	high
3	4	AA enhances positive learning atmosphere.	2.88	0.41	95.95	high
11	6	AA indicates what students can do with English in real-life situations.	2.86	0.35	95.44	high
14	7	AA helps students improve their English language skills.	2.86	0.46	95.44	high
1	8	AA techniques are more constructive than those of TA.	2.85	0.44	94.94	high
8	9	AA helps students monitor their own progress and performance.	2.83	0.38	94.43	high
6	5	AA offers a broad spectrum of assessment possibilities to address students' different learning styles.	2.82	0.39	93.93	high
15	11	AA methods are effective in measuring language skills.	2.80	0.44	93.42	high
13	10	AA offers a comprehensive evaluation of students' linguistic competence.	2.79	0.43	92.92	high
16	12	AA should be used to assess English students learning in Yemeni universities.	2.76	0.50	91.91	high
12	13	AA offers a comprehensive evaluation of students' communicative competence.	2.71	0.55	90.39	high
9	14	AA helps teachers monitor their students' progress and performance.	2.68	0.61	89.38	high
17	15	I prefer using TA to AA.	2.67	0.62	88.88	high
4	16	AA makes students more aware of course objectives.	2.65	0.48	88.37	high
5	17	AA makes students more aware of assessment criteria.	2.58	0.56	85.85	high
10	18	AA methods are less threatening to students than TA.	2.53	0.66	84.33	high
		Total	**2.78**	**0.21**	**92.77**	**High**

Note: Scoring of the negative items, namely Item No. (17) and Item No. (18), was reversed.

As data in Table (2) shows, the average of items of the relevant section in the questionnaire ranged from (2.53) to (2.97) with corresponding percentages from (84.33%) to (98.98%). This indicates a high level of agreement on all items cited in this regard. The total average of the items is (2.78) out of (3:00) and its corresponding percentage is (92.77%), showing a high level of positive attitudes towards AA. Based on these results, the respondents tend to believe in efficiency of AA in EFL learning. That they have positive attitudes towards AA provides an excellent opportunity to develop AA skills and knowledge and enable them to shift from the traditional to the modern paradigm. Actually, teachers' beliefs about and attitudes to particular assessment methods can play an important role in adopting those methods (Yang, 2008).

The finding surfaced from this part of analysis is in harmony with some previous studies that revealed teachers' positive attitudes towards AA in general (Chan, 2008; Iqbal & Manarvi, 2011; Alkharusi, Aldhafri, Alnabhani, & Alkalbani, 2012; Gonzales & Aliponga, 2012; Ghaicha & Omarkaly, 2018). Some other studies reported a positive attitude of teachers towards a particular means of AA. For instance, Tangdhanakanond and Wongwanich (2012) advocated that teachers of all majors in Thailand, including language teachers, had positive attitudes towards the use of portfolio.

On the other extreme, some studies came up with divergent findings. Contrary to the findings of this study, Watt (2005), revealed that teachers generally did not favor implementing AAs. Similarly, Metin (2011), surveying primary school Turkish teachers' attitudes towards performance assessment, indicated that the general attitudes of teachers towards performance assessment and each subscale were at medium level. Likewise, Al-Nouh et al. (2014) revealed that EFL primary school teachers' attitudes towards AA were at a medium level.

RQ #2: Are there statistically significant differences in the faculty members' attitudes to AA based on university type, gender, specialization, experience, or academic rank?

With respect to this question, the *t*-test for independent samples analysis was used to explicate any significant differences in the respondents' attitudes towards AA as for the variables of university and gender. As well, the ANOVA was used to statistically determine any significant differences in the respondents' attitudes towards AA. The results of *t*-tests and ANOVA are outlined in the following tables.

Table 3: ***T-test Results for Respondents'Attitudes towards AA based on 'University Type' & 'Gender'***

Variable		N	*M*	*SD*	*df*	*t*	*Sig.*(2-tailed)
University Type	Public	48	2.76	0.20	64	1.334	0.187
	National	18	2.84	0.23			
Gender	Male	49	2.75	0.22	48.10	2.865	0.006
	Female	17	2.88	0.13			

Although no statistically significant difference was found [t (64) = 1.334, p = 0.187] in the attitudes towards AA between respondents from the *public universities* (M = 2.76, SD = 0.20) and those from the *national university* (M = 2.84, SD = 0.23) at the (0.05) level of significance as Table (3) shows, it was noted that respondents from the national university showed more positive attitudes towards AA than those from the public universities.

With respect to gender, however, statistically significant difference was found [t (48.10) = 2.865, p = 0.006] between *male* (M = 2.75, SD= 0.22) and *female* (M = 2.88, SD= 0.13) at the (0.05) level of significance as Table (3) shows. This implies that gender can be a factor in the attitudes of faculty members towards AA.

In contrast to this finding, Metin (2011) reported no significant differences in relation to gender in the teachers' attitudes towards performance assessment. Similarly, Alkharusi et al. (2012) reported no statistically significant differences in the attitudes of teachers towards educational assessment, including AA with respect to gender. Likewise, Gonzales and Aliponga (2012) found no significant difference in classroom assessment preferences, including assessment for learning, between male and female language teachers.

Table 4: ***ANOVA Results for Respondents'Attitudes towards AA Based on 'Specialization'***

Source of Variance	Sum of Squares	*df*	Mean Square	*F*	*Sig.*
Between Groups	49.979	2	24.989	1.801	0.174
Within Groups	874.279	63	13.877		
Total	924.258	65			

Table (4) shows that no statistically significant differences were found in the respondents' attitudes towards AA at the (0.05) level of significance according to the *specialization* variable. The F-value was (1.801), indicating no significant differences at α=0.05 since the p-

value >0.05 (p = 0.174). This implies that faculty members with different specializations have similar attitudes towards AA.

However, statistically significant difference in attitudes towards AA may be found among faculty members from different departments (social sciences, management sciences, environmental sciences, applied sciences etc.) as Iqbal and Manarvi (2011) indicated or among teachers teaching different subjects (English language, Arabic language, social sciences, etc.) as Alkharusi et al. (2012) found.

Table 5: ***ANOVA Results for Respondents'Attitudes towards AA Based on 'Years of Experience'***

Source of Variance	Sum of Squares	*Df*	Mean Square	*F*	*Sig.*
Between Groups	4.156	2	2.078	0.142	0.868
Within Groups	920.101	63	14.605		
Total	924.258	65			

Similarly, no statistically significant differences were found in the respondents' attitudes towards AA with respect to *years of experience* variable at the (0.05) level of significance as Table (5) shows. The *F*-value was (0.142), indicating no significant differences at α = 0.05 since the *p*-value>0.05 (p=0.868). This implies that faculty members with different years of EFL teaching experience do not differ significantly in their attitudes towards AA.

In line with this finding, Gonzales and Aliponga (2012) reported no significant difference in classroom assessment preferences between respondents in terms of years of language teaching. In contrast to this finding, some previous studies such as Watt (2005), Iqbal and Manarvi (2011), Metin (2011), and Al-Nouh et al. (2014) revealed that young teachers and those with less years of experience have more positive attitudes towards AA than older teachers or those with more years of experience. Unlike these studies, Chan (2008) reported a statistically significant relationship between beliefs about multiple assessment and EFL teaching experience mainly in favour of those with more years of teaching experience.

Table 6: ***ANOVA Results for Respondents'Attitudes towards AA Based on 'Academic Rank'***

Source of Variance	Sum of Squares	*df*	Mean Square	*F*	*Sig.*
Between Groups	9.900	3	3.300	0.224	0.880
Within Groups	914.358	62	14.748		
Total	924.258	65			

Table (6) reveals that no statistically significant differences were found in the respondents' attitudes towards AA based on their *academic ranks* at the (0.05) level of significance as shown in Table (6). The *F*-value was (0.224), indicating no significant differences at α = 0.05 since the *p*-value>0.05 (*p*=0.880). This implies that the academic ranks of faculty do not influence their assessment preferences. As respondents of all academic ranks showed positive attitudes towards AA, it can be stated that teachers' assessment preferences depend mainly on the effectiveness of assessment, rather than on their academic ranks.

This finding is in consonance with Gonzales and Aliponga (2012) who found no significant difference in assessment preferences with respect to educational degree (bachelor, master, and doctorate) in relation to assessment for learning techniques although those with higher educational degree scored higher means than those with lower educational degree in all assessment preferences, including assessment for learning. In relation to assessment to learning, which is not the focus of this study, Gonzales and Aliponga found significant difference in language teachers' assessment preferences based on the educational degree.

RQ #3: What is the mean score of faculty members' use of AA to assess their EFL undergraduates?

To answer the third question, the means, the standard deviations, and the percentages for each items of the section related to the use of AAs were obtained and arranged in a descending order. These values were also calculated for the whole section as shown in Table (7). A benchmark for the acceptable degree of using AAs was also maintained. As the scale used to collect data was a three-point Likert scale, the minimum standard limit set for acceptable use of AAs under investigation was specified by getting (2) out of (3) degrees with a percentage of (66.66%) out of the whole (100%). Although the respondents showed positive views on AA, the results of the present study revealed an overall medium level of AA use to assess undergraduate EFL students at Yemeni university. In other words, the positive attitudes towards AA were not fully reflected in their actual practices of AAs.

Table 7: ***Descriptive Statistics of the Use of AAs in a Descending Order***

#	Rank	Alternative Assessment Methods	*M*	*SD*	*%*	Estimation
15	1	Open-ended questions	2.65	0.59	88.37	high
5	2	Writing samples	2.55	0.59	84.84	high
19	3	Discussions /debates	2.38	0.52	79.29	high
2	4	Presentations	2.23	0.60	74.24	medium
12	5	Observations	2.14	0.68	71.21	medium
3	6	Role playing	2.11	0.70	70.20	medium
14	7	Story or text retelling	1.98	0.59	66.15	medium
8	8	Cloze tests	1.97	0.68	65.65	medium
11	9	Peer assessment	1.95	0.64	65.14	medium
4	10	Demonstrations	1.89	0.47	63.12	medium
10	11	Self-assessment	1.88	0.67	62.62	medium
16	12	Checklists	1.80	0.50	60.09	medium
1	13	Projects	1.73	0.54	57.57	medium
18	14	Reports	1.67	0.54	55.55	medium
17	15	Games	1.58	0.50	52.52	low
6	16	Conferences/Interviews	1.50	0.50	50.00	low
9	17	Portfolios	1.45	0.56	48.48	low
13	18	Anecdotal records	1.39	0.49	46.46	low
7	19	Journals & learning logs	1.30	0.46	43.43	low
		Total	**1.90**	**0.23**	**63.42**	**Medium**

As data in the Table (7) indicates, the averages of using AAs ranged from (1.30) to (2.65) with corresponding percentages ranged from (43.43%) to (88.37%). The total average of using AAs under investigation was (1.90) out of (3:00) and its corresponding percentage was (63.42%), a value below the minimum standard limit set for acceptable use (i.e., 2.00 out of 3.00). At the level of each method, Table (7) shows that six methods out of nineteen (31.58% of the total number of methods under investigation) reached the minimum standard limit set for acceptable use. The 'open-ended questions' ranked first (*M*=2.65, 88.37%), followed by 'writing samples' (*M*=2.55, 84.84%). The 'discussions/debates' ranked third (*M*=2.38, 79.29%), followed by 'presentations' (*M*=2.23, 74.24%). In the fifth rank came 'observations' (*M*=2.14, 71.21%), followed by 'role playing' (*M*=2.11, 70.20%).

Table (7) also shows that the respondents reported a medium-level use for (8) AAs (42.11% of the total number of AAs under investigation). However, the use of these AAs did not reach the minimum standard limit set for acceptable use. These methods are 'story or text retelling' (*M*=1.98, 66.15%), followed by 'cloze tests' (*M*=1.97, 65.65%), followed by 'peer assessment' (*M*=1.95, 65.14%),

followed by ‘demonstrations’ (*M*=1.89, 63.12%), followed by ‘self-assessment’ (*M*=1.88, 62.62%), followed by ‘checklists’ (*M*=1.80, 60.09%), followed by ‘projects’ (*M*=1.73, 57.57%), followed by ‘reports’ (*M*=1.67, 55.55%). Out of the (19) AAs under investigation, the respondents reported a low-level use of (5) methods (26.31% of the total number of AAs under investigation). These methods are ‘games’ (*M*=1.58, percentage=52.52%), followed by ‘conferences/interviews’ (*M*=1.50, percentage=50.00%), followed by ‘portfolios’ (*M*=1.45, percentage=48.48%), followed by ‘anecdotal records’ (*M*=1.39, 46.46%). The last rank included ‘journals and learning logs’ (*M* =1.30, 43.43%).

With these results, it can be concluded that the faculty members in focus use some AAs to assess their students, which is a very encouraging result. It suggests that introducing AA to their teaching would not be a big challenge. A flashback on the literature shows that these results are similar to those of some previous studies, which reported medium-level of practicing AA, albeit differences at the level of practicing some AA forms. For instance, Yang (2012) indicated that EFL teachers at tertiary level adopt a variety of classroom assessment tasks, but differed in the frequency of each task used. While informal questioning and observations were the most commonly used AA tasks, oral exams, conferences, role-playing, and presentations were the less common AA tasks. Self-assessment, portfolios, creative writing, journals, projects, and peer assessment were the least common AA tasks. Similarly, Alkharusi et al. (2012) reported a medium-level of AA practices that the teachers adopted. Likewise, Öz (2014) found that Turkish EFL teachers’ preferences and practices of assessment for learning methods in EFL classroom are at medium-level for oral exam, group work, projects, portfolios, performance assessment, and presentations and at low-level for rubric, self-assessment, peer assessment, and observation forms.

Relatively similar to such studies, Abu Rezeq and Abu Taha (2018) revealed that the use of AA strategies among EFL teachers in Gaza was generally moderate. The authors found that performance-based assessment (presentations, demonstrations, role-playing), group work, observations, questions and answers, checklists, and rating scales strategies are practiced at moderate level, while debates, conferences, interviews, self-assessment, peer assessment, diaries, portfolios, free tasks, projects, learning log, and anecdotal records strategies are practiced at low level. However, the results of this study

are in contrast to the findings of Chan (2008), who reported that EFL teachers from Northern Taiwan used more AA techniques (classroom observations, portfolio assessments, peer assessment, self-assessment) than traditional tests. Likewise, the results of Iqbal and Manarvi (2011) revealed that most faculty members at Pakistani universities practiced AA techniques in some way or the other.

RQ #4: Are there statistically significant differences in the faculty members' use of AA based on university type, gender, specialization, experience, or academic rank?

With respect to this question, the *t*-test for independent samples analysis was used to identify significant differences in the respondents' use of AA based on university and gender, and ANOVA to statistically identify significant differences in relation to the other three variables.

Table 8: ***T-test Results for Respondents' Use of AAs Based on 'University Type' & 'Gender'***

Variable		N	*M*	*SD*	*df*	*t*	*Sig.*(2-tailed)
University Type	Public	48	1.84	0.21	64	3.714	0.000
	National	18	2.06	0.20			
Gender	Male	49	1.88	0.24	64	1.334	0.187
	Female	17	1.97	0.20			

Statistically significant difference was found [t (64) =3.714, p= 0.000] in the use of AAs between respondents from the *public universities* (M=1.84, SD=0.21) and respondents from the *national university* (M=2.06, SD=0.20) at the (0.05) level of significance in favor of those from the national university as Table (8) shows. This finding can be justified by the relative availability of the requirements needed for AA implementation in terms of reasonable size of classes, availability of some resources and facilities, evaluation system, and encouragement for faculty members to adopt some AAs.

With respect to gender, no statistically significant difference was found [t (64)=1.334, p= 0.187] in the use of AAs between *males* and *females* at the (0.05) level of significance although females reported higher level of AA use (M=1.97, SD=0.20) than males (M=1.88, SD=0.24) as shown in Table (8). This finding is consistent with Abu Rezeq and Abu Taha (2018), who found no statistically significant differences in using AA strategies among English language teachers in relation to gender. However, it is in contrast to Alkharusi et al. (2012), who reported statistically significant differences on teachers' use of

AA according to gender, male teachers used AA more frequently than female teachers.

Table 9: ***ANOVA Results for Respondents' Use of AAs Based on 'Specialization'***

Source of Variance	Sum of Squares	*df*	Mean Square	*F*	*Sig.*
Between Groups	54.666	2	27.333	1.475	0.237
Within Groups	1167.819	63	18.537		
Total	1222.485	65			

With regard to specialization, no statistically significant differences were found in the respondents' use of AAs at the (0.05) level of significance as shown in Table (9). The *F*-value was (1.475), indicating no significant differences at α = 0.05 since the *p*-value > 0.05 (p = 0.237). This implies that faculty members with different specializations have no different attitudes towards AA.

Table 10: ***ANOVA Results for Respondents' Use of AAs Based on 'Years of Experience'***

Source of Variance	Sum of Squares	*df*	Mean Square	*F*	*Sig.*
Between Groups	58.455	2	29.2281	1.582	0.214
Within Groups	1164.030	63	18.477		
Total	1222.485	65			

Unexpectedly, no statistically significant differences were found in the respondents' use of AA at the (0.05) level of significance according to the multitude of teaching experience as Table (10) shows. The *F*-value was (1.582), indicating no significant differences at α=0.05 since the *p*-value>0.05 (p=0.214). This can be justified by the absence of factors, such as training and/or practice that can differentiate between teachers based on experience in favour of those with more years of experience.

This finding is in agreement with Alkharusi et al. (2012), who indicated that teaching experience did not correlate significantly with teachers' use of AA. However, Alkharusi et al. found that there was a statistically significant positive relationship between teaching experience and teachers' use of TA. Similarly, Gonzales and Aliponga (2012) found no differences in classroom assessment practices with regard to years of teaching although it was found that language teachers with only one to three years of teaching reported lowest scores in all assessment practices. Likewise, Abu Rezeq and Abu Taha (2018), found no statistically significant differences in the degree of using AA strategies among English language teachers based on years

of experience. In contrast, Chan (2008) found statistically significant difference in EFL teachers' practices of multiple assessments, including AA, according to the variable of experience in favor of the more experienced teachers.

Table 11: ***ANOVA Results for Respondents' Use of AAs Based on 'Academic Rank'***

Source of Variance	Sum of Squares	*df*	Mean Square	*F*	*Sig.*
Between Groups	14.047	3	4.682	0.240	0.868
Within Groups	1208.438	62	19.491		
Total	1222.485	65			

Similarly, no statistically significant differences were found in the respondents' use of AAs based on the academic rank at the (0.05) level of significance as shown in Table (11). The *F*-value was (0.240), indicating no significant differences at α=0.05 since the *p*-value>0.05 (p=0.868). This can be justified by the absence of factors, such as training and/or experience in practice that usually make the practice of teachers with higher academic rank better than that of those with lower academic rank.

RQ #5: Is there any significant relationship between their attitudes to and use of AA?

To answer this question, Pearson correlation coefficient was computed to identifying if the attitudes of English faculty members at Yemeni universities towards AA were related to their use of AAs.

Table 12: ***Pearson Correlation Coefficient Between Respondent's Attitudes towards and Use of AA***

	M		Attitudes	Use
Attitudes	2.78	Pearson Correlation	1	.181
		Sig. (2-tailed)		.147
		N	66	66
Use	1.90	Pearson Correlation	.181	1
		Sig. (2-tailed)	.147	
		N	66	66

Table (12) shows that the relationship between the respondents' attitudes towards and use of AA was not significantat α = 0.05 since the *p*-value > 0.05 (p = 0.147). The Pearson correlation coefficient for the relationship between respondents' attitudes towards and use of AA was (0.18), indicating a low positive relationship. This suggests that faculty members' positive attitudes towards AA may not significantly

determine how frequently they will use AAs to assess their EFL students. In agreement with Tierney (2006) who identified two types of sources that influence teachers' assessment practices: external sources (educational policy and professional development) and an internal source (teachers' beliefs and conceptions), the researcher thinks that the disparity between faculty members' attitudes towards and their practices of AA can be attributed to some reasons and factors, including educational assessment system at Yemeni universities, faculty members' assessment competence, size of classes, lack of administrative support, lack of resources and facilities required for AA implementation, etc.

In line with this finding, Alkharusi et al. (2012) found that teachers have positive attitudes towards educational assessment and perceive themselves as competent in educational assessment, yet Alkharusi et al found that their competence and attitudes were not reflected in their assessment practices. Alkharusi et al. reported that the heavy teaching load might have a negative impact on teachers' assessment practices. In contrast, Chan (2008) reported a positively significant relationship between EFL teachers' beliefs and practices of multiple assessments, including AA. Similarly, Yang (2008) revealed a substantial relationship between teachers' beliefs about the pedagogical benefits of AA and their practices of this type of assessment.

In brief, teachers' beliefs, attitudes, and experience relating to EFL students' assessment are usually reflected in their assessment practices (Quilter & Gallini, 2000); however, this does not usually happen due to some reasons. In our context, such reasons should be identified in order to understand what influences the alignment of the faculty members' positive attitudes towards AA and their assessment practices in the classroom. Yemeni universities should play effective role in removing the reasons that impede using AAs to assess EFL students and provide the necessary support for the faculty members in this regard with a view of realizing the potential benefits of AA.

6. Conclusion and Recommendations

Taken the findings together, it is obvious that AA is largely unripe in the context under scrutiny. Although faculty members in the target universities have positive attitudes towards AA and they use some of its means to assess their undergraduate students, their efforts have not come to fruition. All what they do is insufficient to reap the potential benefits of AA. Hence, the following recommendations are put

forward to encourage EFL faculty members to adopt AA in their assessment practices.

- Training EFL faculty members on AA and how to integrate it into their classes.
- Developing standards to be used as a guide for evaluating faculty members' assessment practices.
- Developing an ongoing professional program to strengthen educational assessment knowledge and skills of the faculty members at Yemeni universities, including knowledge and skills of AA and its underlying theory.
- Encouraging faculty members at English departments to diversify means of assessment to gauge their learners' performance.
- Providing necessary resources to ensure a better implementation of AA.
- Training students on how to depart allegiance to the TA and imbibe concepts and skills required for AA in EFL programs.
- Making the necessary changes in the system of evaluation of Yemeni universities to adopt AA in lieu of the TA forms.

Limitations and further research

The findings of this endeavor are limited to a relatively small number of respondents whose responses were on a self-report three-point Likert scale. They might have rated more uses of AAs than what they actually did in the current study if they used open-ended instruments. Perhaps, observational data from both faculty and students would have yielded better insight into the use of alternative methods to assess undergraduate EFL students. Since the topic of AA is immature in the local context and, coupled with these limitations, it could be food for thought that inspires other researchers to conduct further research. Future researchers may take up solo AAs in EFL classroom. They may also survey the perceptions of EFL teachers and/or students, illuminate the obstacles of using AA, explore the impact of implementing AA on EFL students, and the degree of using AA in EFL classroom from the point of view of students, and the like.

References

Abu Rezeq, K. A & Abu Taha, M. A. (2018). The level of using alternative assessment strategies among English language teachers in the preparatory stage in government schools in Gaza. *Journal of Al-Quds Open University for Educational &*

Psychological Research & Studies, 24, 1-18. doi: 10.5281/zenodo.1344989

Alibakhshi, G. & Shahrakipour, H. (2014). The effect of self-assessment on EFL learners' receptive skills. *Journal Pendidikan Malaysia, 39*(1), 9-17.

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. Liontas (Ed), *The TESOL encyclopedia of English language teaching* (pp. 1-6). Wiley & Sons. DOI: 10.1002/9781118784235.eelt0325

Al-Nouh, N., Taqi, H. &. Abdul-Kareem, M. (2014). EFL primary school teachers' attitudes, knowledge and skills in alternative assessment. *International Education Studies, 7*(5), 68-84.

American Federation of Teachers (AFT). (2008). *The appropriate use of student assessments*. Washington, DC: Author. Retrieved from http://www.aft.org/

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

Baillie, C. (2004). Introduction: Reflecting on effective learning and teaching in engineering education. In C. Baillie & I. Moore (Eds.), *Effective learning and teaching in engineering* (pp. 1- 6). London and New York: Routledge Falmer.

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.

Birjandi, P. & Tamjid, N. (2012). The role of self-, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment & Evaluation in Higher Education, 37*(5), 513-533. doi: 10.1080/02602938.2010.549204

Burnaz, Y. E. (2011). *Perceptions of EFL learners towards portfolios as a method of alternative assessment: A case study at a Turkish state university* (Unpublished master's thesis). Middle East Technical University, Turkey.

Charvade, M., Jahandar, S., & Khodabandehlou, M. (2012). The impact of portfolio assessment on EFL learners' reading comprehension ability. *English Language Teaching, 5*(7), 129-139. doi: 10.5539/elt.v5n7p129

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 Timişoara 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

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.

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). Retrieved from https://www.semanticscholar.org/

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.

Honsa, Jr., S. (2013). Self-assessment in EFL writing: A study of intermediate EFL students at a Thai University. *Voices in Asia Journal, 1*(1), 34-57.

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.

Iqbal, A. & Manarvi, I. (2011). Teachers' attitudes and perceptions for alternative assessment techniques: a case study of Pakistani universities. *Int. J. Teaching and Case Studies, 3*(2/3/4), 131-164.

Janisch, C., Liu, X., & Akrofi, A. (2007). Implementing alternative assessment: Opportunities and obstacles. *The Educational Forum, 71*(3), 221-230.

Little, D. (2009). The European language portfolio: Where pedagogy and assessment meet. A paper presented in *the 8th International Seminar on the European Language Portfolio*. Graz, AT: Council of Europe. Retrieved from https://www.researchgate.net/

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.

Metin, M. (2011). The examinations of teachers' attitude towards performance assessment with respect to the different variables. *Energy Education Science and Technology Part B: Social and Educational Studies, 3*(3), 269-284.

Moheidat, A. S. & Baniabdelrahman, A. A. (2011). The impact of Omani twelfth-grade students' self-assessment on their performance in reading in English. *Asian EFL Journal, 13*(1), 48-84.

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.

Moradan, A. & Hedayati, S. (2011). The impact of portfolios and conferencing on Iranian EFL learners' writing skill. Journal of English Language Teaching and Learning, 8, 115-140.

Nasab, F. (2015). Alternative versus traditional assessment. *Journal of Applied Linguistics and Language Research, 2*(6), 165-178. Retrieved from http://www.jallr.com/

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.

Öz, H. (2014). Turkish teachers' practices of assessment for learning in the English as a foreign language classroom. *Journal of Language Teaching and Research, 5*(4), 775-785.

Peng, J-C. (2009). *Peer assessment of oral presentation in an EFL context*. Retrieved from ProQuest Dissertations and Theses database. (UMI Number: 3380148).

Purwanti, T. T. (2015). The implementation of self-assessment in writing class: A case study at Stba Lia Jakarta. *TEFLIN Journal*, *26*(1), 97-116. doi: 10.15639/teflinjournal.v26i1/97-116

Quilter, S. M., & Gallini, J. K. (2000). Teachers' assessment literacy and attitudes. *The Teacher Educator, 36*(2) 115-131. doi: 10.1080/08878730009555257

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.

Spratt, M., Pulverness, A., & William, M. (2011). *The teaching knowledge test (TKT) - glossary*. Cambridge: Cambridge University Press.

Stoynoff, S. (2012). Looking backward and forward at classroom-based language assessment. *ELT Journal, 66*(4), 523-532. doi:10.1093/elt/ccs041

Struyf, E., Vandenberghe, R., & Lens, W. (2001). The evaluation practice of teachers as a learning opportunity for students. *Studies in Educational Evaluation, 27*(3), 215–238.

Tangdhanakanond, K., & Wongwanich, S. (2012). Teacher attitude and needs assessment concerning the use of student portfolio assessment in Thailand's educational reform process. *International Journal of Psychology*, *10*, 71-88.

Tannenbaum, J-E. (1996). Practical ideas on alternative assessment for ESL students. *ERIC Digest*: Center for Applied Linguistics.

Tavakoli, M. & Amirian, Z. (2012). The effect of portfolio assessment on the development of meta d is course awareness in EFL learners' writing in the academic context. *Sheikhbahaee EFL Journal*, *1*(1), 1-21.

Țeican, B. E. (2016). Outcomes of alternative assessment in adult language training. *Challenges of the Knowledge Society*, *6*, 961-965.

Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education, 13*(3), 239-264. doi: 10.1080/09695940601035387

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.

Wikström, N. (2007). *Alternative assessment in primary years of international baccalaureate education* (Master's thesis, The Stockholm Institute of Education, Sweden). Retrieved from http://su.diva-portal.org/

Yang, T-L. (2008). Factors affecting EFL teachers' use of multiple classroom assessment practices with young language learners. *English Teaching & Learning*, *32*(4), 85-123.

Yang, T-L. (2012). Tertiary EFL teachers' classroom assessment practices. *The International Journal of Learning, 18*(3), 1-8.

Yildirim, R. & Orsdemir, E. (2013). Performance tasks as alternative assessment for young EFL learners: Does practice match the curriculum proposal? *International Online Journal of Educational Sciences*, *5*(3), 562-574.

Yin, F. (2006, November). Applying methods of formative and summative assessment to problem-based learning in computer courses. *The China Papers*, 42-45.