



---

## Incorporating Purposeful Assessment Scenarios in L2 Listening Class: A Learner-Oriented Approach

---

Elham Ghorbanpour<sup>1</sup>, Gholam-Reza Abbasian<sup>2\*</sup>, Ahmad Mohseni<sup>3</sup>

<sup>1</sup> Ph.D. Candidate, Department of English Language Teaching, Kish International Branch, Islamic Azad University, Kish Island, Iran

*elhamghorbanpour4@gmail.com*

<sup>2\*</sup> Associate Professor, Department of English Language Teaching, Kish International Branch, Islamic Azad University, Kish Island, Iran; Imam Ali University, Tehran, Iran, (Corresponding author)

*gabbasian@gmail.com*

<sup>3</sup> Associate Professor, Department of English, South Tehran Branch, Islamic Azad University Tehran, Iran

*Amohseny1328@gmail.com*

---

Received: 21 May 2020

Accepted: 23 September 2020

---

### Abstract

This study investigated the effect of purposeful assessment scenarios, Assessment *for* Learning (AFL), Assessment *as* Learning (AAL), and integrated purposeful assessment scenarios on Iranian EFL learners' development of listening comprehension. To this end, 100 learners organized into three experimental and one control group participated in listening ~~tasks~~ the mentioned purposeful assessment forms. Parametric statistical analysis in the form of some one-way ANOVAs revealed that EFL learners exposed to integrated assessment could significantly develop their listening comprehension better than the AFL- and AAL-oriented instruction groups. However, no significant difference between AAL- and AFL-oriented groups' listening comprehension was obtained. Though the findings are not in favor of any of the individual scenarios, they lend support to their feasibility in classroom performance assessment and assert the usefulness and viability of the synthetic approach to assessment through which measurement and instruction are bifurcated.

**Keywords:** AAL Assessment; AFL Assessment; AOL Assessments; Integrated Assessment; Listening Comprehension.

---

\* <sup>2</sup>Corresponding Author's Email:  
*gh\_abbasian@azad.ac.ir*



## INTRODUCTION

When it comes to the assessment of language skills, instructors are concerned with selecting the most practical and applicable approach to evaluate the learner's progress and concentrate on their points of strength and weakness. As an innovative breakthrough in education and in line with the unification approach, AFL, AAL, and AOL emerged when constructivism (Piaget, 1960) attempted to pinpoint the role of assessment in teaching language skills and whether the three above-mentioned assessment approaches could facilitate the learners' development of the language skills.

Assessment *for* Learning (AFL) (Dann, 2014; Earl, 2013; Lam, 2013) conducted during the formation of knowledge offers an alternative perspective to the psychometrics period in schools, while Assessment *of* Learning (AOL) summative in nature is typically conducted at the end of something, for example, a unit, a program, or a course (Bennett, 1984). The feedback is provided in the form of grades, while little advice or direction for improvement is suggested (Ramsden, 2003). AOL is part of institutional accountability processes and quality assurance to ensure the students' performance, via grades, test, certificate, etc. (Guskey & Bailey, 2001). On the other hand, AFL maintains that assessment eventually informs instruction by revealing what is ineffective and effective during the learners' achievement in educational context and is most helpful if used for meaningful change or reform (Penn, 2011). However, formative assessment is more complex than it appears at first sight

(Yorke, 2003). There are different conceptions that can fall into two camps: one group considered formative assessment as primarily involving formally structured tasks; what Bell and Cowie (2001) refer to as planned formative assessment. Another group, who has a more constructivist view, considered formative assessment as mainly informal and ad hoc, similar to Bell and Cowie's interactive formative assessment. Thus, AFis L is mainly aligned as a formative assessment that can be an interactive endeavor, and teachers provide feedback and assistance as part of the assessment in order to scaffold the next steps. Formative assessment is also defined as a kind of "assessment designed to provide direction for improvement and/or adjustment to a program for individual students or for a whole class, that is, quizzes, initial drafts/attempts, homework, and questions during instruction" (O'Connor, 2002, p. 109).

To avoid the doubts and confusions about the processes of formative assessment, Carless, coining the term learning-oriented assessment (LOA) in 2006, looked at the issues from different lenses. LOA evolved from both summative assessment and formative assessment, and its framework involves three integrated components, assessment *for* learning, assessment *as* learning, and assessment *of* learning. In Carless and his colleagues' (2006) framework, a fundamental component of assessment is represented by the term "assessment tasks as learning tasks" (p. 9). When assessment tasks involve the desired learning outcomes, learners can

experienced deep learning by progressing towards these outcomes (Carless, 2007).

The second component of LOA predisposes that student involvement in assessment is crucial. Involving the students can help them develop a better understanding of learning goals and engage them more actively with standards and criteria (Carless, 2007). Thus, self-evaluation skills, peer feedback (Liu & Carless, 2006), or peer assessment (Falchikov, 2005) can be considered as parts of the assessment processes. The combination of learning and teaching tied with the significance given to self in psychology led to the notion of Assessment as Learning (AAL) (Earl, 2013). In AAL, students are engaged as critical and active assessors, who can comprehend the information, associate it with previous knowledge, and gain the target ability. These are the regulatory processes in metacognition which occur when learners monitor their learning process and use the feedback to make adjustments, changes, and adaptations in their knowledge (Earl, 2013). Thus, AAL is "the ultimate goal, where students are their own best assessors" (Earl, 2003, p. 47). Lee (2017) asserts that AAL tenets draw on several theoretical viewpoints as well as theories of autonomy, motivation, self-regulation and metacognition.

Thirdly, assessment can lead to learning if students receive proper feedback which can be used to "feed-forward into future work" so as to support current and future student learning (Carless, 2007, p. 59). AFL includes "all those activities undertaken by teachers, and/or by their students, which provide information to be

used as feedback to modify the teaching and learning activities in which they are engaged" (Dunn & Mulvenon, 2009, p. 10). Wehlburg (2007) described the use of data in such a way as "closing the feedback loop". The teachers also use the information gained from those tests to review their teaching. They also identify areas of improvement or weaknesses in instruction to improve their practice in the classroom and the quality of pedagogy. As Gibbs and Simpson (2005) point out, feedback in itself may not enhance learning, unless learners engage with it and act upon it. In addition to teachers, peers can also provide feedback (Falchikov, 2001).

Similarly, Earl (2003) pointed out the values of AAL, AFL and called for the integration of all these approaches in a right balance. Earl (2003) believes that the instructors can utilize classroom-based assessment to promote learners' achievement under the influence of assessment reforms. In the same vein, some scholars believe in the integration of these two complementary approaches. For instance, as Black and Wiliam (2009) state, AFL tends to be public as it is about formative interactions taking place in classrooms, while AAL is private as the students engage in reflection (Yancey, 1998). Moreover, AFL relies on using several sources of feedback to inform the overall learning and teaching process; whereas, AAL mostly focuses on the internal feedback of learners. Moreover, Carless (2011) argued that the integration of AAL and AFL would be more helpful. Other researchers also believe that "AFL/AAL can co-exist" (Lee, 2016, p. 271).

Additionally, Lam (2008) not only combined the four language skills into one portfolio project but also utilized an integration of assessment approaches. Then, he suggested a renewed assessment cube, in which AAL and AFL are complementing each other formatively.

However, assessment professionals have different views about the integration of assessment approaches. Likewise, other researchers (e.g. Wiliam & Thompson, 2008) doubt that the simultaneous implementation of assessment approaches can fulfill their particular goals. Taking these issues into account on one hand, and the paucity of research in terms of implementing AFL, AAL, and integration of them as to developing listening ability on the other, the rationale behind this study can be justified. So, the study can be rationalized on the grounds of theoretical, pedagogical and methodological claims, particularly in the Iranian pedagogically non-democratic and psychometric-dominated EFL education.

1. *Is there any significant difference between the effects of the purposeful AFL- oriented instruction and AAL-oriented instruction in developing Iranian EFL learners' listening comprehension ability?*
2. *Is there any significant difference between the effects of the integrated purposeful assessment scenarios (AFL+AAL) and AFL-oriented instruction in developing Iranian EFL learners' listening comprehension ability?*
3. *Is there any significant difference between the effects of the integrated*

*purposeful assessment scenarios (AFL+AAL) and AAL- oriented instruction in developing Iranian EFL learners' listening comprehension ability?*

4. *Is there any significant difference between the effects of the integrated purposeful assessment scenarios (AFL+AAL) and AAL-oriented instruction in developing Iranian EFL learners' listening comprehension ability?*

## LITERATURE REVIEW

Assessment is commonly the process of collecting information as to the learners' knowledge gained in light of a particular syllabus (Luongo-Orlando, 2003). According to Drummond (2003), in the assessment process, teachers collect and interpret evidence of students' learning and apply that information to make decisions. A similar definition about assessment is suggested by Pinter (2006) which refers to the data analysis that is utilized by the teacher as the evidence of the students' progress and performance. Therefore, it is obvious that assessment is used as a media to evaluate the students' achievement. Of course, the success of any assessment, according to Shaaban (2001), depends on the selection of proper assessment tools and procedures as well as the appropriate interpretation of the students' performance.

According to Earl (2003), three widely recognized assessment approaches (i.e. AFL, AOL, and AAL) underlying the current classroom assessment practice reveal various focuses of learning conceptions. AOL and

grading have a long history in education and parents and the public have widely accepted them (Earl, 2013). Actually, a significant amount of classroom time is dedicated for assessing the students' learning. Brookhart (2007) notes that teachers dedicate 30 percent to 50 percent of classroom time to assessment related activities. AOL tasks can show the students' competence

AOL is used to ensure the students' performance via grades, test, or certificate (Guskey & Bailey, 2001). They are also utilized to ensure the students' experience and the education quality (Kuh, Kinzie, Schuh, & Whitt, 2005). In fact, AOL is generally involved with school accountability, and according to Brown (2008), it submits that assessments can be "used to account for a teacher's, a school's, or a system's use of society's resources" (p. 18). AOL aims to verify the quality of the target educational program (Maki, 2012). Actually, to meet these requirements, Steele and Lutz (1995) suggested that assessment instruments should be consciously applied by the teachers or other responsible parties involved in the curriculum. In contrast, AFL is closely related with learning-oriented formative practices. William (2009) noted that the first priority of this kind of assessment is serving the purpose of fostering students' learning. This type of assessment, which emphasizes the assessment potential to support learning, has gained considerable attention in educational settings over the last decades (Earl & Timperley, 2014). Some of the main alternatives in assessment proposed in this era are reflections,

conferences, journals, portfolios self-assessment, and peer assessment. AFL notes that assessment eventually informs teaching by clarifying the effective and ineffective practices during the learners' achievement in academic context and is most beneficial if utilized for meaningful reform or change (Penn, 2011). On the other hand, AAL "reinforces and extends the role of formative assessment for learning" (Lee & Mak, 2014, p. 66). Dann (2002) claimed that AAL is "a process through which pupil involvement in assessment can feature as part of learning", and the learner is considered to be a "critical connector" between the learning process and assessment (p. 153). The assessment, as stated by Earl and Katz (2006), is emphasized as a process of metacognition in AAL practice. AAL motivates learners to monitor their learning and emphasizes the significance of enhancing students' ability over time to be their own assessors.. Brown (2008) identifies AAL as any assessment which reveals the students' performance and enables the educators to accept the skilled students and take some measures to help the less-skilled ones. Gibbs (2006) states that AAL demands learners to reflect on how they learn to further improve their own learning.

AAL demands the students to reflect on their 'how of learning' to reach further improvement in their own learning (Gibbs, 2006). Language scholars view self-assessment as an important aspect of language learning suggesting that the ability to evaluate one's own work is the principal goal in education (Boud & Falchikov, 1989). As Archer (2010) specifically states, "Self-monitoring is the

ability to respond to situations shaped by one's own capability at the moment in that set of circumstances, rather than being governed by an overall perception of ability" (as cited in Evans, 2013, p. 87). To do a self-assessment procedure, Sadler (1983, as cited in Boud, Lawson, Thompson, & Simpson, 2011) stated that a student needs the ability to "close the gap" (p. 4). This provides an opportunity for students to investigate their own performance for future success.

In AAL, students are engaged in the assessment process, which can enable them to become responsible for their learning behaviors, leading to their self-reflection and self-monitoring (Archer, 2010). Learners are their own assessors, which results in a learning-oriented procedure through which the learners are expected to manipulate the learning environment (Lam, 2015). AAL originates from the fact that learner self-assessment encourages permanent learning. Evans (2013) highlights this point by reflecting Boud and Lawson's (2011) work. They believe that self-assessment skills demand lifelong acquisition. In fact, the students themselves can be gradually involved in the self-assessment process (Sendziuk, 2010). Moreover, students can do the revisions on their own based on the feedback during the course (Graziano-King, 2007). AAL tries to follow the learners' progress during the semester, and frequently evaluate them for the purpose of better learning (Gibbs, 2006).

Several studies have acknowledged the significance of different kinds of assessment in the pedagogy of teaching English as a second

language. For example, Li (2018) examined the washback and validity of self-assessment, as a specific form of AAL, in interpreter and translator education, whereby he carried out four self-assessments of a group of translators' performance. The findings showed that the students' self-assessment correlated positively with their instructor's assessment; the assessment accuracy of the students enhanced over time with regular repetition; and self-assessment promoted positive learning attitudes among students. Along the same line, Xiao and Yang (2019) investigated how formative assessment could support secondary students' self-regulated learning in English language learning. The researchers conducted classroom observations and interviews with two teachers and 16 students in a foreign language secondary school. The findings indicated that under the guidance of their teachers, the students engaged in formative assessment in a proactive manner and appeared to be emerging as self-regulated learners.

The theoretical underpinnings of LOA were used by Almalki (2019) to investigate Saudi EFL learners' perceptions and their progress in their English speaking proficiency. The findings indicated that LOA played a significant role in enhancing the students' abilities for self and peer assessment, their overall English language skills, as well as their critical thinking. Nevertheless, the empirical studies relevant to the integrated assessment-oriented instruction can indicate the kind of challenges and possible benefits likely to be promoted by such integration. For example, Sadeghi and Rahmati (2017) explored the

validity of arguments regarding assessment integration strategies, tensions, and the potential of an integrated assessment model in improving the students' writing skill. The participants were all high-school students with an age range of 15–18. An integrated assessment as, for, and of learning model was used with a group of students. Moreover, an assessment *for* and *of* (non-integrated) model was experimented with other participants as the comparison group. Afterwards, the candidates' writing performance estimated by Cambridge Assessment in terms of overall band descriptions was converted into numerical indices. It was concluded that "an integrated assessment model tailored to contextual specifications can contribute both theoretically and practically to teaching and assessing writing" (p. 50).

Another pertinent study was conducted by Lee and Coniam (2013) who aimed to investigate the factors that could inhibit or facilitate the integration of AFL in a predominantly AOL setting and explored the influence of this implementation on students' motivation. The results suggested that teachers' interest, commitment, knowledge about the principles of AFL, and their previous experience of using AFL facilitated its implementation. However, some factors were found to inhibit the AFL implementation, including students' focus on summative scores, their lack of involvement with instructor's feedback, the need to prepare students for external exams, time constraint, as well as school policy.

In an experimental investigation, Zarei and Yasami (2016) examined the impact of formative assessment and remedial instruction on Iranian EFL students' listening comprehension. An ANCOVA was conducted not only for intergroup and intragroup performance comparisons. Data analysis indicated that formative assessment and remedial instruction had a considerable impact on the listening comprehension of EFL students. Bayat, Jamshidipour, and Hashemi (2017) also examined the influence of using formative assessment on EFL students' listening efficacy and anxiety. The groups were treated differently in terms of developing listening ability. At the end, it was realized that the formative assessment affected the students' listening ability and efficacy.

Additionally, as Alderson and Bachman (2005) argue, assessing listening skills is one of the least developed, least understood, and yet one of the most significant areas of language assessment and testing. Actually, the review of the related literature yielded that the present study is the first attempt that specifically compares learners in all three approaches to assessment in the area of listening comprehension, particularly in Iranian EFL setting. Further, having reviewed the studies related to the integration of assessment approaches, it can be concluded that sometimes in different contexts, the findings seem rather contradictory. For instance, several researchers (Brookhart, 2001; Carless, 2011; Lam, 2013; Sadeghi & Rahmati, 2017) suggested the integration of assessment scenarios arguing that the use of assessment in a variety of integrated

ways could promote the students' achievement. However, other researchers, (e.g. Harlen, 2012; Lam & Lee, 2010; Lee & Coniam, 2013; Taras, 2005) agree that when these approaches are integrated, the focus is greatly on summative assessment rather than doing the assessment in the learning process by involving the learners

## METHODS

### Participants

The participants were 100 pre-intermediate Iranian EFL language learners classified as three experimental and one control groups each of which composed of 25 EFL learners.

### Instrumentation

The instruments employed in the investigation are explained below:

#### Oxford Placement Test (OPT)

OPT was used to select a homogeneous sample of the participants.

#### Researcher-Made Diagnostic and Achievement Listening Tests

The researcher-made diagnostic listening test was developed based on the contents Touchstone Level 1 (McCarthy, McCarten, & Sandiford, 2004) and *Basic Tactics for Listening* (3rd edition) (Richards & Trew, 2010) which were covered in the course

syllabus. The test was administered before the treatment sessions for the diagnostic purpose of the listening comprehension ability. The achievement listening test was a parallel test (post- test) to the Diagnostic test which similarly underwent piloting and then reliability estimation. The reliability of the pre- test was measured through a pilot study. Reliability coefficient was found to be .79 (using KR-21 formula) as to be an acceptable value in terms of consistency of scores as highlighted in Farhady, Jafarpour, and Birjandi (1994).

### Procedure

The study resembling a continuum of instruction and assessment moving from more teacher-moderated to gradually LOA trend, was conducted in three phases including 1) pre-treatment phase (focusing on the pilot study, the administration of OPT and the pre-test); 2) while-treatment phase (including the complete explanation of the treatment sessions of the three purposeful assessment scenarios, and the AOL-oriented instruction as the control group), and 3) post-treatment phase (entailing the information regarding the post-test). Each phase is taken into account below:

#### Pre-Treatment Phase

As a preparatory stage, this phase started with a pilot process in order to make sure of the development of sound instruments. Following some coordination meetings with the Head of the Institute, the participants were also briefed



on the purpose of the study. Out of 110 language learners 100 subjects, whose OPT scores lied between one SD below the mean and identified as pre-intermediate, were apt to take part in this research. They were randomly divided into three experimental groups and one control group (i.e., AOL-oriented instruction). After that all groups took the listening ability pre-test to check their initial ability prior to exposure to purposeful assessment scenarios.

### **While-Treatment Phase**

All four groups were exposed to specific assessment scenario-based listening instruction based on the materials from Touchstone Level 1 (McCarthy, McCarten, & Sandiford, 2004) and Basic Tactics for Listening (3rd edition) (Richards, 2013). Moreover, they all underwent 12 two-hour sessions of assessment scenario – based instruction of listening ability development detailed as follows:

### **Purposeful AFL-oriented instruction**

As an iterative assessment, AFL was conducted through requiring the learners to carry out the listening activities, and do some listening projects as homework assignments and make them ready for the upcoming sessions. They were interviewed for diagnostic purposes and probing their perceptions towards AFL, whereby the teacher was looking into the learners' developmental progress in their listening and how the existing gaps in the listening instruction could be solved. The teacher carried out a LOA in the form of

teaching and feedback looping process of listening activities. Repeated oral and written quizzes in the forms of questioning, conversations, and learning logs were held in order to test their listening development and provide immediate descriptive feedback on their responses. In the light of AFL they WERE motivated and commitment to learning which brought changes in the classroom culture by making visible what students believed to be true, and used that information to help students move forward in manageable, efficient, and respectful ways. In order to make AFL more systematic, the teacher benefitted from record-keeping for individual which was used a source of individualized descriptive feedback that would help further their learning.

### **Purposeful AAL-oriented instruction**

Though a bit similar to AFL, AAL was conducted a bit differently given to its nature and purpose. The major purpose behind AAL is to understand the way learners learn; to explore the learners' metacognition and help them become aware of it. They were expected to write diary journals and express their points of view regarding each session, which enabled the teacher to explore and provide introspection of the learners' metacognition by allowing them to assess their classmates' comprehension and challenge their understanding of the listening activities they were required to carry out in-class and as homework assignment. They were also encouraged to write portfolios and checklists concerning the listening activities and any comments they found necessary to be

used by the teacher for the future classes.

### **Integrated purposeful assessment scenarios**

In another scenario, a bit more teacher-dominated AFL and learner-involved AAL scenarios were integrated to improve their listening comprehension abilities. Depending on the time of the sessions, the teacher attempted to do the assessment himself and put it on the shoulders of more active learners in order to benefit from integrated assessment. However, caution was made by the teacher in order not to bombard the learners with varieties of assessment tools at the same time. In fact, formative assessment procedure was run in this group as the learners' development of the listening comprehension ability was focused regardless of their final listening score. The main purpose of assessment was to provide a variety of procedures in order to help learners be the evaluator and be assessed during the listening comprehension tasks. The former involved writing diaries and portfolios by the learners to be the teacher's assistant in doing the assessment, while the latter was concerned with the teacher's initiatives in playing an important role in evaluating the learners' listening comprehension progress during the course by interviewing them and analyzing the points of strengths and weaknesses. In fact, the learners were moderated to utilize the integration of AFL and AAS.

### **Purposeful AOL-oriented instruction (the control group)**



In the light of AOL, attempts were made to engage learners in the listening activities and provide interactive listening class but it was conducted in the form of more product and score oriented approach based on which the assessment was fulfilled by the teacher-made listening test taken by the participants at the end of the term for the purpose of summative reports. In fact, the process of instruction was conventional and the assessment was mainly done at the end, though process-based and final data were used to make evaluation.

### **Post-Treatment Phase**

After the 13 treatment sessions, all participants took the Researcher-Made Achievement Listening Test to investigate the effect of different purposeful assessment scenarios on their listening comprehension.

### **Design**

This study was conducted based on quasi-experimental design as the participants were selected non-randomly and based on convenience sampling due to the availability, purposefulness and homogeneity factors as justified in Dörnyei, 2007; and Mackay & Gass, 2005). Though their original selection was done non-randomly, they were randomly divided into four instructional groups.

### **Data Analysis**

In order to analyze the data, quantitative methodologies was considered to answer the

research questions of the study. In addition, so the three research questions were analyzed through the statistical measure using one-way analysis of variance (one-way ANOVA) to inferentially find the significant difference among the three experimental groups and one control group.

## RESULTS

First, the data were checked in terms of normal distribution assumptions in order to acknowledge the proper statistical approach. The Shapiro-Wilk test showed p-values of the

listening comprehension pre- and post-tests of AFL as .153 and .058, those of AAL group as .180 and .071, those of the integrated group as .066 and .113, and those of AOL group as .121 and .060, respectively; inferring all more than .05, which meets the normality assumption and thereby, as argued by Hatch and Lazaraton (1991) justifies the use of parametric tests.

### Addressing the First question

In a bid to address the first question, both descriptive and inferential statistical analyses were carried out as reported in tables 1 and 2.

**Table 1**

*Descriptive Statistics for the Pre-Test of Three Groups*

	Groups	N	Mean	Std. Deviation	Std. Error Mean
Pre-test	AFL	25	28.0000	4.1120	1.13210
	AAL	25	26.5000	5.2920	1.99538
	AOL	25	27.0000	5.0101	1.01005

Table 1 reveals that AFL enjoys the highest mean score (M=28.00), while AAL possesses the least (M=26.50) and AOL occupying the

middle position (M=27.00). For the comparison, pre-test, (one-way ANOVA) was run in Table 2.

**Table 2**

*One-way ANOVA Statistics for the Pre-Test of Three Groups*

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	133.236	2	71.219	2.813	.081
Within Groups	1233.213	72	24.501		

Total	1366.449	74
-------	----------	----

---

Since AFL possesses the highest mean score (M=31.00), while AOL possesses the lowest one (28.00). It seems that there could be similarity between AAL and AFL groups, while

difference can be observed between AAL and AFL groups with AOL group. To compare the mean scores, one-way ANOVA was run in Table 3.

**Table 3**

*One-way ANOVA Statistics for the Post-Test of Three Groups*

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	421.118	2	233.761	7.113	.003
Within Groups	1341.361	72	27.813		
Total	1762.479	74			

Table 3, shows a significant difference among post-test of three groups (F<sub>2, 72</sub> = 7.11, p = .002). Thus, the three groups were different.

To locate the difference Games-Howell multiple comparison statistics were run among the three groups (Table 4).

**Table 4**

*Games-Howell Multiple Comparison Statistics for the Post-Test of Three Groups*

(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
AFL	AAL	0.50000	1.01271	.129	.9803	9.0691
	AOL	3.00000*	1.43170	.001	3.7918	9.3421
AAL	AFL	-0.50000	1.01271	.129	-9.0691	-.9803
	AOL	2.50000*	0.9915	.002	2.8497	7.6619
AOL	AFL	-3.00000*	1.43170	.001	-3.7918	-9.3421
	AAL	-2.50000*	0.9915	.002	-2.8497	-7.6619

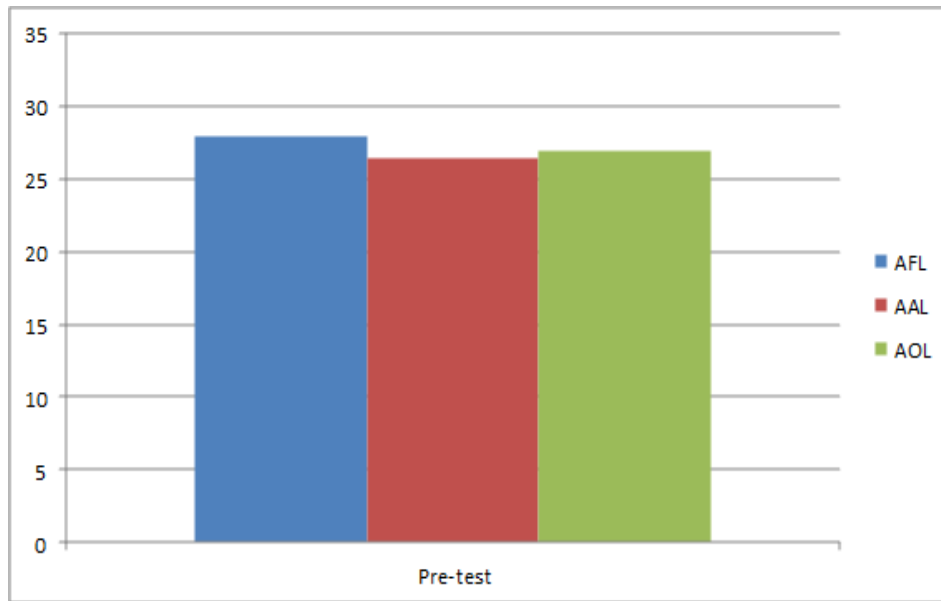
\*. The mean difference is significant at the 0.05 level.

Table 4 shows the significant difference

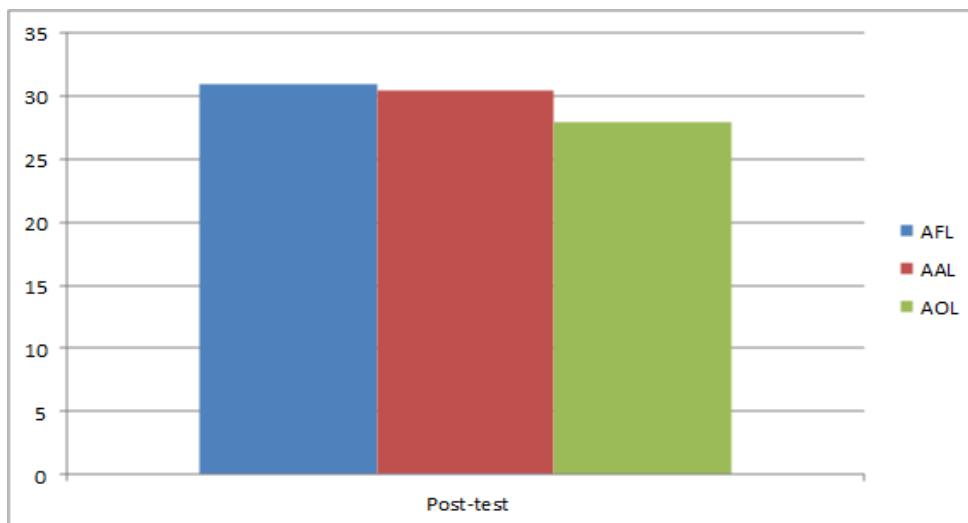
between AFL and AOL groups (p = .001) and



AAL and AOL groups ( $p = .002$ ), while no significant difference between AFL and AAL groups ( $p = .129$ ).



*Figure 1. Mean development for the pre-test of AFL, AAL, and AOL groups*



*Figure 2. shows the development of the post-test of AFL, AAL, and AOL groups*

### Addressing the Second question

The respective descriptive statistics reveal that both Integrated and AFL groups possess the highest mean score ( $M=28.00$ ), while AOL possesses the lowest mean score ( $M=27.00$ ),

showing very little difference among the three groups' mean scores of the pre-test. In order to compare the three groups' mean scores, the pre-test, one-way analysis of variance (one-way ANOVA) was run in Table 5.

**Table 5****One-way ANOVA Statistics for the Pre-Test of Three Groups**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	154.116	2	89.301	1.923	.112
Within Groups	1339.265	72	29.409		
Total	1439.381	74			

Table 5 reveals that there is no significant difference among the mean scores of Integrated, AFL, and AOL groups ( $F_{2, 72} = 1.92, p = .112$ ), which shows the similarity in

the three groups. However, in order to compare the mean scores, one-way ANOVA was run in Table 6.

**Table 6****One-way ANOVA Statistics for the Post-Test of Three Groups**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	501.238	2	314.215	8.745	.000
Within Groups	1419.571	72	36.809		
Total	1920.809	74			

There exists a significant difference among three groups ( $F_{2, 72} = 8.74, p = .000$ ) so

Games-Howell multiple comparisons was run (Table 7).

**Table 7****Games-Howell Multiple Comparison Statistics for the Post-Test of Three Groups**

(I) Groups	(J) Groups	Mean Difference (I- J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Integrated	AFL	2.00000*	1.09933	.003	.9821	4.0116
	AOL	5.00000*	1.1463	.000	1.8329	6.9901
AFL	Integrate d	-2.00000*	1.09933	.003	-.9821	-.4.0116
	AOL	3.00000*	1.4317	.001	3.7918	9.3421
AOL	Integrate d	-5.00000*	1.1463	.000	-1.8329	-6.9901
	AFL	-3.00000*	1.4317	.001	-3.7918	-9.3421



---

**\*. The mean difference is significant at the 0.05 level.**

Table 7 shows that significant difference can be observed among Integrated and AFL groups.

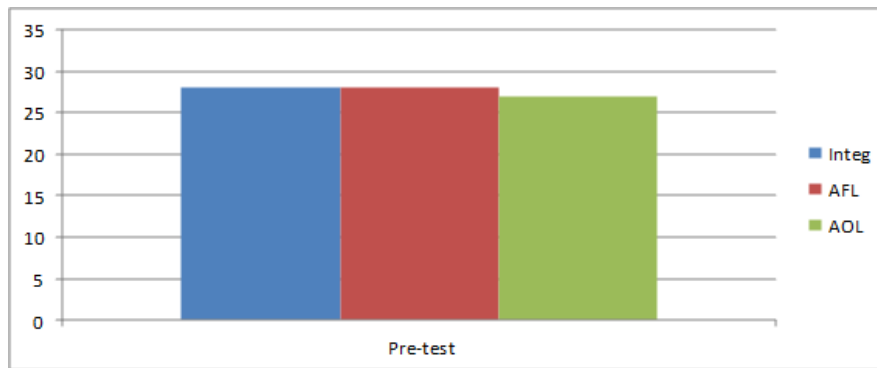


Figure 3. Mean development for the pre-test of integrated, AFL, and AOL groups

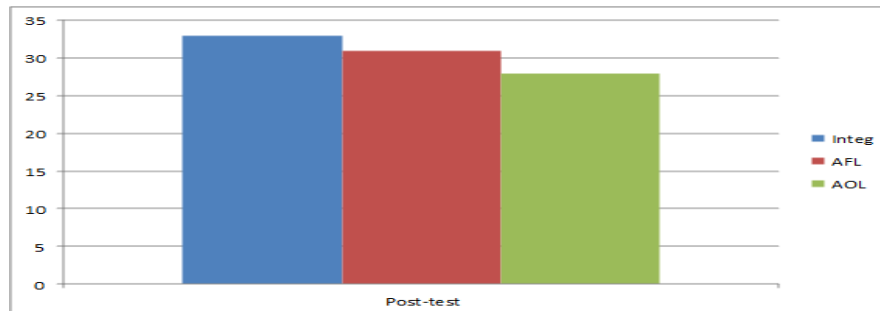


Figure 4. shows the development of the post-test of Integrated, AFL, and AOL groups

### Addressing the Third question

Descriptively, the Integrated group possesses the highest mean score ( $M=28.00$ ), while the AAL group has the lowest one ( $M=26.50$ ),

indicating minor difference among the three groups' mean scores. But, in order to compare the three groups listening comprehension pre-test, one-way analysis of variance (one-way ANOVA) was run in Table 8.

Table 8

One-way ANOVA Statistics for the Pre-Test of Three Groups

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	184.106	2	79.211	2.873	.092
Within Groups	1149.435	72	21.249		
Total	1333.541	74			

Table 8 reveals non-significant difference among the mean scores of Integrated, AAL, and AOL groups ( $F_{2, 72} = 2.87, p = .092$ ), which shows the similarity in the three groups.

The Integrated group possesses the highest mean score ( $M = 33.00$ ), while AOL possesses

the lowest one ( $M = 28.00$ ). So there exists a large difference among three groups, as supported by the respective one-way ANOVA in Table 9.

**Table 9**  
*One-way ANOVA Statistics for the Post-Test of Three Groups*

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	499.205	2	224.101	7.995	.000
Within Groups	1329.201	72	30.309		
Total	1828.406	74			

As to Table 9, there exists a significant difference among post-test of three groups ( $F_{2, 72} = 7.99, p = .000$ ).

So Table 10 shows Games-Howell multiple comparison statistics.

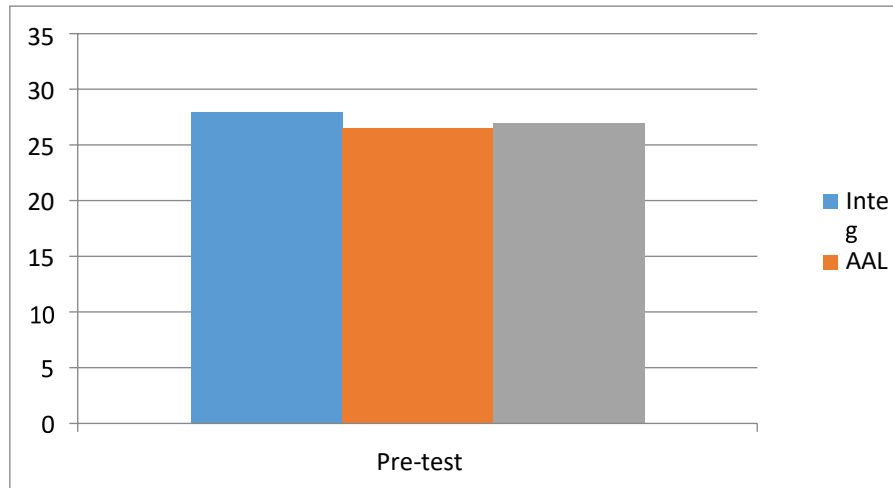
**Table 10**  
*Games-Howell Multiple Comparison Statistics for the Post-Test of Three Groups*

(I) Groups	(J) Groups	Mean Difference (I- J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Integrated	AAL	2.50000*	1.11051	.003	1.0121	4.2116
	AOL	5.00000*	1.1463	.000	1.8329	6.9901
AAL	Integrated	-2.50000*	1.11051	.003	-1.0121	-4.2116
	AOL	2.50000*	0.9915	.002	2.8497	7.6619
AOL	Integrated	-5.00000*	1.1463	.000	-1.8329	-6.9901
	AAL	-2.50000*	0.9915	.002	-2.8497	-7.6619

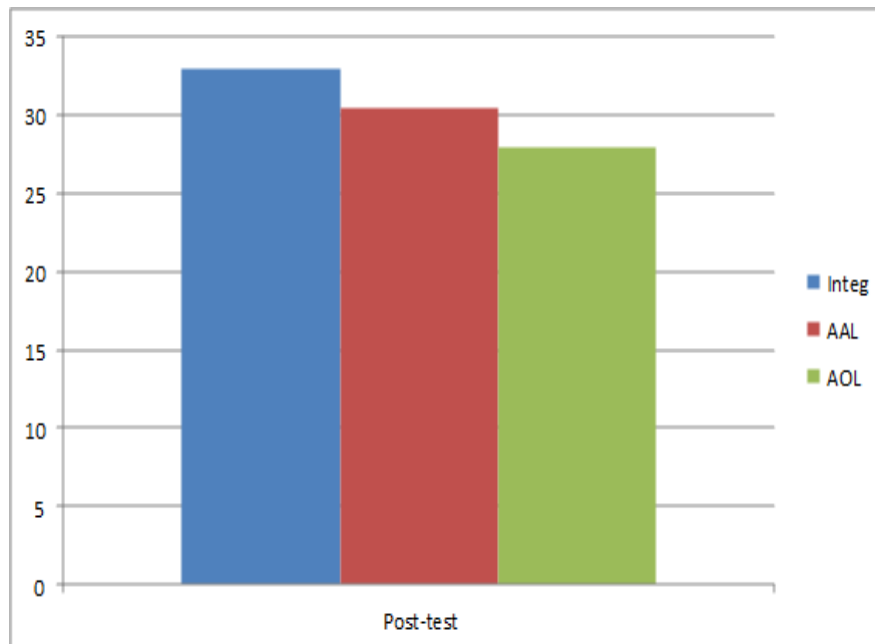
\*. The mean difference is significant at the 0.05 level.

Table 10 shows that there was a significant difference between the effects of the integrated

purposeful assessment and AAL-oriented instruction.



*Figurer 5. Mean development for the pre-test of integrated, AAL, and AOL groups*



*Figure 6. shows the development of the post-test of Integrated, AAL, and AOL groups*

## DISCUSSIONS

The study's findings indicated no significant difference between AFL- and AAL-oriented instructions concerning the learners' listening

comprehension ability. It appears that the commonalities between the two make the learning process more communicative for the learners, and listening comprehension is achieved under such interactive learning

environments when students are engaged in conscious-based, self-reflective, and autonomous learning environment. As each assessment scenario tries to bold the learning environment by using various assessment tools, it might be rather challenging to have a clear-cut comparison of the two and say which one can be better than the other one.

It might be assumed that those (Blaich & Wise, 2011; Dunn & Mulvenon, 2009) favouring AFL might direct the attention toward the teacher as the assessor of the learners' achievement. They believe that assessment is a complex task, which should be done by the teacher who is also the provider of both descriptive and evaluative feedback types. On the other hand, AAL supporters (Graziano-King, 2007; Sendzuik, 2010) conversely argue that sharing the responsibilities with the learners can help them construct their identity and develop their metacognitive awareness. Hence, there is no consensus regarding the superiority of one approach over the other, demonstrating that each should be carried out in its appropriate context and concerning teacher's and learners' pedagogic and realistic needs. This is in agreement with Xiao and Yang's (2019) study who demonstrated that under the guidance of their teachers, the students perceived the classroom formative assessment activities they experienced. The feedback they received to be helpful in the development of their deep understanding and capability for self-regulation in English language learning. Carless (2007) also emphasized timely and forward-looking feedback to support current and future student

learning.

Other experimental investigations such as Zarei and Yasami (2016), Bayat et al. (2017) also revealed that by using formative assessment, the students could retain more information, understood concepts more quickly and were more interested in what they were learning. The evidence from Li (2018) also indicated that students' self-assessment, as a specific form of AAL, correlated positively with their instructor assessment and promoted positive learning attitudes among students. On the other hand, AAL supporters (Graziano-King, 2007) conversely argue that sharing the responsibilities with the learners can help them construct their personal identity and develop their metacognitive awareness. Quantitative measures of the listening comprehension tests revealed that integrated purposeful assessment provided an optimal condition for the EFL learners to significantly outperform the AFL group in doing their listening comprehension tasks. When assessment is integrated, it becomes more focused, which probably results in more academic achievement by the learners. The point should be considered that apparently integrated type of assessment produces positive results; however, it seems that more important than the integration is the teacher's expertise in productively implementing the assessment materials according to the specific guidelines each approach follows (Davis & Neitzel, 2011). Therefore, teachers should be provided with awareness programs to make the assessment-oriented instruction more purposeful than it suggests.

It is also noteworthy that the teacher should



be cautious in creating a balance between AAL and AFL in the sense that the two assessment scenarios complement each other. Although integration was found to be effective concerning the findings of the study, it might oppositely result in undesirable findings as the integrated approach appears to put much load of responsibilities on the learners' shoulders, which can run the risk of having less-motivated and mentally-tired language learners due to their challenges in writing diaries, keeping their logs, and doing the listening tasks simultaneously. On the whole, the integrated purposeful assessment scenarios might worth a try since it definitely brings about learners' autonomy in commitment in learning. These findings are in alignment with those of Sadeghi and Rahmati (2017) who used an integrated assessment as, for, and of learning model for improving the students' writing skill and indicated that the integrated assessment group outperformed the non-integrated assessment participants. Investigating the learners' pre- and post-test scores of listening comprehensions revealed the outperformance of the integrated group compared to AAL group in doing the listening comprehension tasks. In fact, the integration of AFL and AAL can boost the learners' cognitive and metacognitive capacities in the listening class and benefiting from an autonomous language learning environment.

Although there seems to a paucity of literature regarding the usefulness of the integrated approach in assessment, Davis and Neitzel (2011) implied that both AFL and AAL can be simultaneously adopted by a teacher who

has enough expertise in implementing the target assessment materials not to exaggerate the use of descriptive and evaluative types of feedback. The integrated approach is apparently effective in fostering learners' involvement within the assessment procedure –making them more self-regulated assessors of their learning process – although the teacher's role is to a very large extent undeniable in not flooding the learners with assessment tasks, probably leading to having less-engaged language learners due to their boredom caused by varieties of activities. In fact, when learners are using both AFL and AAL, it appears that both teacher and his/her learners are simultaneously involved in the assessment procedures, which demands the implementation of the logical mechanism to carry out the assessment tasks more purposefully, resulting in the learners' success in both learning and assessment. Findings can be well supported by Chappuis (2009) and Gibbs (2006) who encouraged language teachers to benefit from AFL and AAL assessment types, respectively, resulting in the effective employment of assessment tasks at the service of learning gains.

## CONCLUSION

In the light of the present study, more areas of inquiry were identified to help multiple stakeholders and consumers. It is necessary to revise and redesign pedagogy to balance the tensions among assessment *as, for, and of* learning and to use the advantages of each to improve learning and teaching (Mok, 2012), regardless of the statistically non-significant

differences between the two scenarios examined in this study as this very achievement can't imply assessment lacks advantages. The findings of this research may assist policymakers in emphasizing the significance of using different approaches to the skills evaluation. Moreover, it seems that students, teachers, and researchers can also benefit from the outcomes of the present study.

Learners are considered as the first beneficiary of the study findings. Many learners appear to be worried about their listening ability in the process of language learning and are usually concerned with their listening skill as well as their grades in listening exam. Being assessed through a purposeful method of assessment, learners can overcome listening difficulties since they are exposed to practice and interactive learning environment when are productively assessed and consciously involved within the assessment procedure. In fact, when learners are aware of their listening skill, they can take necessary action to solve the possible deficiencies in listening as well as strengthening their listening ability and awareness raising. Since listening comprehension is a demanding task for language learners, purposeful assessment scenarios enable learners to be in charge of their listening progress by being involved in the assessment process. Nature of AFL and AAL indicates teachers' moderation and modelling on one hand, and learners' participation in the process of being active, acting as a source of data, cooperating in expressing their metacognitive strategies, responding to feedback loops, and playing their participatory

role on the other facilitate the process of self-evaluating their performance.

In contrast, the teacher tries to monitor their learning behaviours (Gibbs & Simpson, 2005). The learners' involvement in the assessment procedures not only motivate the learners' development of language skills, but they also help learner to be aware of the significance of being assessed during the learning process (Archer, 2010). In addition, assessment scenarios can make the learning process more lasting since keeping track of learners' progress occurs during the term, which results in their internalization of the focused skill (Choi, Nam, & Lee, 2001).

Teachers need to provide assessment feedback to maximize its potential for students' action in line with the Learning-oriented assessment (LOA) framework. In fact, findings of the present study revealed that teachers might be encouraged to do assessments of AFL- or AAL-oriented instruction in order to track the learners' progress according to the focused language skill. No matter which language skill is concentrated, using purposeful assessment scenarios can provide a neat schedule for teachers to provide feedback for the learners based on their feedback to teachers gathered by their diary writings, interviews, or portfolios (Black & Wiliam, 2006). Findings of the study demonstrated that the teacher's expertise, as seen in the steps and procedure of conducting the tripartite scenarios of assessment (i.e., AFL, AAL, AOL, and the Integrations of the AFL with AAL, in providing the mechanism for performing purposeful assessment scenarios might be worth to attention since 'how to

implement' and 'what of' assessment can be one of the main causes of learners' success in doing the listening tasks in the assessment-based learning setting (Buck, 2001). One of the major differences between assessment and testing lies, in fact, in the teacher's expertise as the latter is much more complex than simple psychometric assessment and requires multiple skills from the teacher's side.

However, these findings should be generalized with care as the context and sample are not representative of the whole population of English learners in different settings.

Moreover, a single, commonly agreed upon definition of comprehension remains elusive (Cutting & Scarborough, 2006). Different comprehension assessments do not always generalize across items, formats, and subjects due to differing definitions of comprehension.

Therefore, further research is required to explore other variables, such as different learning environments as well as different levels of proficiency, and other language skills.

As mentioned by Dobson (2010) and Mok (2012), much more research is required to develop systems of theories and strategies for expanding LOA and to provide evidence of how AOL, AFL, and AAL improve students' learning. Moreover, AFL, AAL, or AOL is a strategic process that cannot be separated from teachers' activities. A teacher can be assessment designer, curriculum developer, and knowledge producer (Zeng, Huang, Yu, & Chen, 2018).

## References

- Alderson, C., & Bachman, L. F. (2005). Series editors preface. In G. Buck. In *Assessing Listening* Cambridge: Cambridge University Press.
- Almalki, M. (2019, No. 1 2019). Learning-oriented assessment, critical thinking and English language speaking skills: An exploratory study of Saudi EFL learners. *International Journal of English Language Education*, 7(1), 37-50. Retrieved from <http://www.macrothink.org/journal/index.php/ijele/article/view/14528>
- Archer, J. C. (2010). State of the science in health professional education: effective feedback. *Medical Education*, 44(1), 101-108. doi:10.1111/j.1365-2923.2009.03546.x
- Bayat, A., Jamshidipour, A., & Hashemi, M. (2017). The Beneficial Impact/s of Applying Formative Assessment on Iranian University Students' Anxiety Reduction and Listening Efficacy. *International Journal of Languages' Education*, 1(Volume 5 Issue 2), 1-11. doi:10.18298/ijlet.1740
- Bell, B., & Cowie, B. (2001). The Characteristics of Formative Assessment in Science Education. *Science Education*, 85(5), 536-553. Retrieved from <https://eric.ed.gov/?id=EJ633665>
- Bennett, W. J. (1984). *To reclaim a legacy: a report on the humanities in higher education*. Washington, D.C.: National Endowment for the Humanities.



- Black, P., & Wiliam, D. (2006). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. doi:10.1080/0969595980050102
- Black, P., & Wiliam, D. (2009). Developing the Theory of Formative Assessment. *Educational Assessment Evaluation and Accountability*, 21(1), 5-31. doi:10.1007/s11092-008-9068-5
- Blaich, C., & Wise, K. (Eds.). (2011). *From gathering to using assessment results: Lessons from the Wabash national study*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).
- Boud, D., & Falchikov, N. (1989). Quantitative Studies of Student Self-Assessment in Higher Education: A Critical Analysis of Findings. *Higher Education*, 18(5), 529-549. doi:10.1007/BF00138746
- Boud, D., Lawson, R., Thompson, D., & Simpson, L. (2011). *The development of student judgement: The role of practice in grade prediction*. Paper presented at the 14th Biennial EARLI Conference for Research on Learning and Instruction: Education for a Global Networked Society, Exeter, United Kingdom. Conference Paper retrieved from [http://vhost0309.web04.level27.be/conferences/EARLI\\_Biennial\\_Conferences/previous\\_Biennial\\_conferences/earli\\_2011](http://vhost0309.web04.level27.be/conferences/EARLI_Biennial_Conferences/previous_Biennial_conferences/earli_2011)
- Brookhart, S. M. (2001). Successful Students' Formative and Summative Uses of Assessment Information. *Assessment in Education Principles Policy and Practice*, 8(2), 153-169. doi:10.1080/09695940123775
- Brookhart, S. M. (2007). Feedback that Fits. *Educational leadership*, 65(4), 54-59.
- Brown, G. T. L. (2008). *Conceptions of Assessment: Understanding what Assessment Means to Teachers and Students*. New York: Nova Science Publishers.
- Buck, G. (2001). *Assessing listening*. Cambridge; New York: Cambridge University Press.
- Carless, D. (2007). Learning-Oriented Assessment: Conceptual Bases and Practical Implications. *Innovations in Education & Teaching International*, 44(1), 57-66. doi:10.1080/14703290601081332
- Carless, D. (2011). *From testing to productive student learning implementing formative assessment in Confucian-heritage settings* *Routledge research in education* v 52 (pp. xiv, 263 p.). doi:10.4324/9780203128213
- Carless, D., Joughin, G., & Mok, M. M. C. (2006). Learning-oriented assessment: Principles and practice. *Assessment & Evaluation in Higher Education*, 31(4), 395-398. doi:10.1080/02602930600679043
- Chappuis, S., & Stiggins, R. J. (2002). Classroom Assessment for Learning. *Educational leadership: journal of the*

- Department of Supervision and Curriculum Development, N.E.A.*, 60(1), 40-43.
- Choi, K., Nam, J.-H., & Lee, H. (2001). The Effects of Formative Assessment with Detailed Feedback on Students' Science Learning Achievement and Attitudes Regarding Formative Assessment. *Journal of the Korean Association For Research in Science Education*, 20(2), 28-34.
- Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension. *Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured*, 10(3), 277-299. doi:10.1207/s1532799xssr1003\_5
- Dann, R. (2002). *Promoting assessment as learning: improving the learning process* (Vol. 5). London: Routledge/Falmer.
- Dann, R. (2014). Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice*, 21(2), 149-166. doi:10.1080/0969594x.2014.898128
- Davis, D. S., & Neitzel, C. (2011). A Self-Regulated Learning Perspective on Middle Grades Classroom Assessment. *The Journal of Educational Research*, 104(3), 202-215. doi:10.1080/00220671003690148
- Dobson, S. (2010). Book review: how assessment supports learning. Learning-oriented assessment in action. *Assessment in Education: Principles, Policy and Practice*, 17(2), 105-112.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford; New York, N.Y.: Oxford University Press.
- Drummond, M. J. (2003). *Assessing children's learning* (2nd ed. ed.). London: David Fulton.
- Dunn, K. E., & Mulvenon, S. W. (2009). A Critical Review of Research on Formative Assessment: The Limited Scientific Evidence of the Impact of Formative Assessment in Education. *Practical Assessment, Research, and Evaluation*, 14(7), 11. doi:10.4324/9780203462041\_chapter\_1
- Earl, L. M. (2003). *Assessment as learning: using classroom assessment to maximize student learning* (Second edition. ed.). Thousand Oaks, CA: Corwin.
- Earl, L. M. (2013). Assessment for learning; Assessment as learning: Changing practices means changing beliefs. *assessment*, 80, 63-71.
- Earl, L. M., & Katz, M. S. (2006). *Rethinking classroom assessment with purpose in mind: assessment for learning, assessment as learning, assessment of*

- learning*. Winnipeg, Canada: Manitoba Education, Citizenship & Youth.
- Earl, L. M., & Timperley, H. (2014). Challenging conceptions of assessment. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.). In *Designing Assessment for Quality Learning* (Vol. 1, pp. 325-336). Dordrecht: Springer.
- Evans, C. (2013). Making Sense of Assessment Feedback in Higher Education. *Review of educational research*, 83(1), 70-120.
- Falchikov, N. (2001). *Learning Together: Peer Tutoring in Higher Education*. London: RoutledgeFalmer.
- Falchikov, N. (2005). *Improving Assessment through Student Involvement: Practical Solutions for Aiding Learning in Higher and Further Education*. London: RoutledgeFalmer.
- Farhady, H., Jafarpour, A., & Birjandi, P. (1994). *Testing language skills*. Tehran: SAMT Publications.
- Gibbs, G. (2006). How assessment frames student learning. In C. Bryan, & K. Clegg (Eds.). In *Innovative assessment in higher education* (pp. 23-36). London: Routledge.
- Gibbs, G., & Simpson, C. (2005). Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education*(1), 3-31. Retrieved from <http://eprints.glos.ac.uk/3609/>
- Graziano-King, J. (2007). Assessing Student Writing: The Self-Revised Essay. *Journal of Basic Writing*, 26(2), 75-94.
- Guskey, T. R., & Bailey, J. M. (2001). *Developing Grading and Reporting Systems for Student Learning. Experts in Assessment*. Thousand Oaks, Calif.: Corwin Press.
- Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.),. In *Assessment and Learning* (2 ed., pp. 87-102). doi:10.4135/9781446250808
- Hatch, E. M., & Lazaraton, A. (1991). *The Research Manual: Design and Statistics for Applied Linguistics*. Boston, Mass.: Heinle & Heinle.
- Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2005). *Student Success in College: Creating Conditions That Matter* (s. ed. Ed.). San Francisco, Calif.: Jossey-Bass.
- Lam, R. (2008). Adopting Effective Portfolio-based Assessment: An Integrative Approach. *Modern English Teacher*, 17(1), 36-41.
- Lam, R. (2013). Formative Use of Summative Tests: Using Test Preparation to Promote Performance and Self-Regulation. *The Asia-Pacific Education Researcher*, 22(1), 69-78.
- Lam, R. (2015). Assessment as learning: examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies in Higher Education*, 1-18. doi:10.1080/03075079.2014.999317
- Lam, R., & Lee, I. (2010). Balancing the Dual Functions of Portfolio Assessment.

- ELT journal*, 64(1), 54-64.  
doi:10.1093/elt/ccp024
- Lee, I. (2016). Putting Students at the Centre of Classroom L2 Writing Assessment. *Canadian Modern Language Review*, 72(2), 258-280. doi:10.3138/cmlr.2802
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*.
- Lee, I., & Coniam, D. (2013). Introducing Assessment for Learning for EFL Writing in an Assessment of Learning Examination-Driven System in Hong Kong. *Journal of Second Language Writing*, 22(1), 34-50. doi:10.1016/j.jslw.2012.11.003
- Lee, I., & Mak, P. (2014). Assessment as Learning in the Language Classroom. *Assessment as learning. Hong Kong: Education Bureau, 2014(3)*, 66-78. Retrieved from <https://wlts.edb.hkedcity.net/a&l3.htm>
- Li, X. (2018). Self-Assessment as 'Assessment as Learning' in Translator and Interpreter Education: Validity and Washback. *The Interpreter and Translator Trainer*, 12(1), 1-20. doi:10.1080/1750399X.2017.1418581
- Liu, N.-F., & Carless, D. (2006). Peer Feedback: The Learning Element of Peer Assessment. *Teaching in Higher Education - TEACH HIGH EDUC*, 11(3), 279-290. doi:10.1080/13562510600680582
- Luongo-Orlando, K. (2003). *Authentic assessment: designing performance-based tasks*. Markham, Ontario: Pembroke Publishers Limited.
- Maki, P. L. (2012). *Assessing for learning: Building a sustainable commitment across the institution*: Stylus Publishing, LLC.
- McCarthy, M., McCarten, J., & Sandiford, H. (2004). *Touchstone Level 1*. Cambridge: Cambridge University Press.
- Mok, M. M. C. (2012). *Self-directed learning oriented assessments in the Asia-Pacific*. Dordrecht; London: Springer.
- O'Connor, K. (2002). *How to grade for learning: linking grades to standards* (2nd ed.). Arlington Heights, IL: SkyLight Professional Development.
- Penn, J. D. (2011). The Case for Assessing Complex General Education Student Learning Outcomes. *New Directions for Institutional Research*, 14, 5-14. doi:10.1002/ir.376
- Piaget, J. (1960). *The child's conception of the world*. London: Routledge & Kegan Paul.
- Pinter, A. (2006). *Teaching young language learners*. Oxford: Oxford University Press.
- Ramsden, P. (2003). *Learning to teach in higher education* (2nd ed.). London; New York: RoutledgeFalmer.
- Richards, J. C., & Trew, G. (2010). *Basic tactics for listening* (3rd ed.). New York, NY; Oxford: Oxford University Press.
- Sadeghi, K., & Rahmati, T. (2017). Integrating assessment as , for , and of learning in

- a large-scale exam preparation course. *Assessing Writing*, 34, 50-61. doi:10.1016/j.asw.2017.09.003
- Sendziuk, P. (2010). Sink or Swim? Improving Student Learning through Feedback and Self-Assessment. *International Journal of Teaching and Learning in Higher Education*, 22(3), 320-330.
- Shaaban, K. (2001). Assessment of Young Learners. *Forum*, 39(4), 16-23.
- Steele, J. M., & Lutz, D. A. (1995). *Report of ACT's research on postsecondary assessment needs: American College Testing Program*.
- Taras, M. (2005). Assessment: Summative and Formative: Some Theoretical Reflections. *British Journal of Educational Studies*, 53(4), 466-478. doi:10.1111/j.1467-8527.2005.00307.x
- Wehlburg, C. M. (2007). Closing the Feedback Loop Is Not Enough: The Assessment Spiral. *Assessment Update*, 19(2), 1-2. doi:10.1002/au.192
- William, D. (2009). *Assessment for learning: why, what and how?* London: University of London, Institute of Education.
- William, D., & Thompson, M. (2008). Integrating assessment with instruction: What will it take to make it work? In (pp. 53-82): Lawrence Erlbaum Associates.
- Xiao, Y., & Yang, M. (2019). Formative assessment and self-regulated learning: How formative assessment supports language learning. *System*, 81, 39-49. doi:10.1016/j.system.2019.01.004
- Yancey, K. B. (1998). *Reflection in the writing classroom* (1 ed., pp. 1 online resource (vii, 215 p.)). doi:10.2307/j.ctt46nsh0
- Yorke, M. (2003). Formative Assessment in Higher Education: Moves towards Theory and the Enhancement of Pedagogic Practice. *Higher Education*, 45(4), 477-501. doi:10.1023/A:1023967026413
- Zarei, N., & Yasami, N. (2016). *The Impact of Formative Assessment and Remedial Teaching on EFL Learners' Listening Comprehension*. Paper presented at the Conference proceedings. ICT for language learning.
- Zeng, W., Huang, F., Yu, L., & Chen, S. (2018). Towards a Learning-Oriented Assessment to Improve Students' Learning--A Critical Review of Literature. *Educational Assessment, Evaluation and Accountability*, 30(3), 211-250. doi:10.1007/s11092-018-9281-9

students' self-regulation in English

### **Biodata**

**Ms. Elham Ghorbanpour** as a PhD candidate of TEFL at Islamic Azad University, South Tehran Branch, (Kish Int'l Campus), has published two original articles and attended national and international I conference. Her main areas of interest are language testing and assessment. She also presented at the 17th International TELLSI Conference while studying for her doctorate.

Email: [elhamghorbanpour4@gmail.com](mailto:elhamghorbanpour4@gmail.com)

**Dr. Gholam-Reza Abbasian** is an assistant professor of TEFL at Imam Ali University and a member of the Teaching English Language & Literature Society of Iran (ELLSI) Board of Directors. He has presented at (inter) national conferences, and authored and translated about 15 books. He offers language testing and assessment, research methods, and SLA courses at the Ph.D. level and supervised about 100 theses and dissertations. He acts as an external examiner of Ph.D. dissertations of Malaysian universities and as the internal manager of JOMM, and reviewer of Sage, FLA & GJER, and some other journals.

Email: [gabbasian@gmail.com](mailto:gabbasian@gmail.com)

[gh\\_abbasian@azad.ac.ir](mailto:gh_abbasian@azad.ac.ir)

**Dr Ahmad Mohseni** is an associate professor of Applied Linguistics at Islamic Azad University, South Tehran Branch. He has been teaching TEFL/TESL courses for 35 years at the undergraduate and postgraduate levels. He has carried out a number of research projects, and he is the author of 10 books, and published numerous scholarly essays in national and international academic journals. Dr. Mohseni has also been an invited professor at American Global University, College of Education in Wyoming, USA.