Language Teaching and Educational Research

e-ISSN 2636-8102

Volume 3, Issue 2 | 2020

A Comparative Study of Two Ways of Presentation of Listening Assessment: Moving towards Internet-based Assessment

Ghasem Modarresi
Kaveh Jalilzadeh

To cite this article:

Modarresi, G., & Jalilzadeh, K. (2020). A comparative study of two ways of presentation of listening assessment: Moving towards internet-based assessment. *Language Teaching and Educational Research (LATER), 3*(2), 176-194. https://doi.org/10.35207/later.742121

LATER
http://dergipark.org.tr/later
Language Teaching and Educational Research
e-ISSN: 2636-8102 *LATER*, 2020: 3(2), 176-194
Turkish JournalPark ACADEMIC

Research Article

A comparative study of two ways of presentation of listening assessment: Moving towards internet-based assessment

Ghasem Modarresi* ORCID
Assistant Professor, Islamic Azad University, Quchan Branch, Department of English, Quchan, IRAN
Kaveh Jalilzadeh ORCID
Lecturer, Istanbul University-Cerrahpasa, Department of Foreign Languages, İstanbul, TURKEY

Abstract

Despite the theoretical importance of internet-based assessment, there is a paucity of experimental research into it. The present study, as an experimental study, is an attempt to compare internet-based and paper-based assessment of listening comprehension for secondary students. In so doing, 36 male students who were studying English language at a secondary school in Tehran participated in this study. The students were divided into two groups: one control group which was exposed to paper-based assessment and one experimental group which was exposed to internet-based assessment developed by the researchers providing the students with the listening quizzes and tests and guidelines including assessment techniques such as leading questions and hints. Following the treatment which lasted for fifteen sessions, the results of the one-way ANCOVA confirmed that there was a significant difference between the two groups on post-test scores on the listening test. Indeed, internet-based group outperformed paper-based group in their listening scores. Having examined the significant difference between the students' scores in internet-based assessment obtained from Time 1, Time 2 and Time 3, as measured by the ANOVA, the results indicated that there were statistically significant differences at the $p<.05$ level in students' scores for the three sets of scores.

Received
24 May 2020

Accepted
08 September 2020

Keywords
internet-based assessment
paper-based assessment
listening comprehension

Suggested APA citation: Modarresi, G., & Jalilzadeh, K. (2020). A comparative study of two ways of presentation of listening assessment: Moving towards internet-based assessment. *Language Teaching and Educational Research (LATER), 3*(2), 176-194. https://doi.org/10.35207/later.742121

Note(s) from the author(s)
» *Corresponding author
» The ethics committee approval was obtained from Molla Sadra High School in Tehran Provincial Directorate of Education with a document number of 100/11/936 on 27/04/2020.

Author(s)' statements on ethics and conflict of interest
Ethics statement: We hereby declare that research/publication ethics and citing principles have been considered in all the stages of the study. We take full responsibility for the content of the paper in case of dispute.
Statement of interest: We have no conflict of interest to declare.
Author contribution disclosure: All authors contributed equally to this study.
Funding: None
Acknowledgements: None

Dinleme becerisini değerlendirme yöntemlerine yönelik karşılaştırmalı bir çalışma: İnternet tabanlı değerlendirmeye geçiş

Öz

İnternet tabanlı değerlendirmenin kuramsal önemine rağmen, bununla ilgili deneysel araştırma yeterince bulunmamaktadır. Bu deneysel çalışma, ortaokul öğrencileri için dinlediğini anlamanın internet tabanlı ve kâğıt üzerinde değerlendirmesini karşılaştırmaktadır. Bu çalışmaya Tahran'da bir ortaokulda İngilizce öğrenen 36 erkek öğrenci katılmıştır. Öğrenciler iki gruba ayrılmıştır. Kontrol grubu kâğıt-kalem sınavları üzerinden değerlendirmeye tabi tutulan öğrencileri içermektedir. Deney grubu öğrencilerine ise araştırmacılar tarafından geliştirilen yönlendirici sorular ve ipuçları gibi değerlendirme tekniklerini de içeren kılavuzlar, kısa dinleme sınavları ve testlerinin uygulandığı internet tabanlı değerlendirme uygulanmıştır. On beş oturumluk uygulamanın ardından yapılan tek yönlü ANCOVA sonuçları, iki grup arasında dinleme testinde son-test puanlarında anlamlı bir farklılık olduğunu doğrulamıştır. İnternet tabanlı grubun dinleme puanlarında kağıt temelli gruptan daha iyi performans sergilediği görülmüştür. Öğrencilerin ANOVA ile ölçülen Zaman 1, Zaman 2 ve Zaman 3'ten aldıkları internet tabanlı değerlendirme puanları arasındaki farklılık incelendiğinde, sonuçlar üç puan seti için öğrencilerde $p<.05$ düzeyinde istatistiksel olarak anlamlı farklılıklar olduğunu göstermiştir.

Gönderim
24 Mayıs 2020

Kabul
08 Eylül 2020

Anahtar kelimeler
internet tabanlı değerlendirme
kâğıt temelli değerlendirme
dinlediğini anlama

Önerilen APA atıf biçimi: Modarresi, G., & Jalilzadeh, K. (2020). Dinleme becerisini değerlendirme yöntemlerine yönelik karşılaştırmalı bir çalışma: İnternet tabanlı değerlendirmeye geçiş. *Language Teaching and Educational Research (LATER), 3*(2), 176-194. https://doi.org/10.35207/later.742121

Introduction

In recent years, researchers have shown an increased interest in the integration of technology into second language teaching and learning (Ulla & Perales, 2020; Lin, Warschauer, & Blake, 2019). Actually, technology can offer effective materials for teachers to enhance their pedagogical practices in classroom settings. Among these materials, the use of Computer-Assisted Language Learning (CALL) is becoming more salient in the field of English language teaching, acting as scaffolding for the students to promote their learning development.

Indeed, it is documented that in the Western society students are encouraged to set their own goals, reflect on and take responsibility for their own learning, and thus become independent and autonomous learners (Langer, 2001). By contrast, L2 students are required to respect and follow their teachers, who are the authorities in the classroom, and the whole educational system tends to function for the sake of various standardized tests, even though some changes have emerged as the result of reforms from "examination-oriented education" to "quality-based education" (Langer, 2001, p. 23). It is wondered, consequently, whether and to what extent assessment, as a successful instructional and assessment tool in the Western culture, succeeds or fails in the Iranian EFL context.

Point taken, attention to the role of computerized materials and internet-based materials in foreign language settings has been increased in recent years (Modarresi & Alavi, 2014; Al-Kadi, 2018). There has been a movement in learning and teaching from the traditional text-based materials to the more hands-on computer-based materials, and the students are appealing to the use of technology in their learning and such tools as the internet, email or so. Computer-aided materials are some of the manifestation of CALL which is defined as "the search for and study of applications on the computer in language teaching and learning" (Levy, 1997, p. 1). An advantage of internet-based assessment stems from their potential for enhancing the task authenticity "through using the technology that is appropriate to a target communicative situation and using technological multimedia capabilities to highlight appropriate situational aspects of test tasks"(Douglas, 2010, p. 139). For example, the implementation of sound and video in test tasks can help to make them resemble real-life situations, such as telephone conversations, university lectures, job interviews, and the like.

Meanwhile, listening comprehension, as a means of communication, plays an important role in people's everyday lives. Listening is probably the least explicit of the four language skills, making it the most difficult skill to learn (Vandergrift, 2004). It is evident that children listen and respond to language before they learn to talk. Technically speaking, it is emphasized that listening comprehension is multi-dimensional in nature with a number of different information sources informing the comprehension process in no fixed order (Buck, 2001). It is agreed that listening strategies promote students' listening proficiency (Graham, 2017; Yeldham & Gruba, 2016). Research shows that learners do have their own listening strategies, and there are some differences in what they do in order to comprehend the listening text (Oxford, 1990). However, very few studies have focused on what teachers do in practice in the classroom for listening (Graham, 2017).

In Iran, the traditional teaching, learning, and assessment culture is still prevalent in listening classes, interpreter training classes and learning environments without focus on new assessment criteria (Khoramy Nia & Modarresi, 2019). As listening skills are still being taught

by outdated approaches in many universities, and its instructors use the "listen and repeat" method, so that each student is required to listen one or two sentences and read them aloud for the instructors and classmates. This method and the likes minimize students' participation to at most four or five students, and a boring and exhausting atmosphere dominates the whole class, while the researchers' investigations have indicated that more innovative approaches are utilized in Western universities to maximize participation and enhance motivation in teaching and learning environment that lead to better listening competence.

Specifically speaking, the role of teachers in the new era of technology is not only to transmit new knowledge, but also to give students tools to acquire knowledge and recognize the value of what they see in books and software as well as on the internet (Bancheri, 2006). The point is that the internet has become an effective means of communication, a place for experiencing different cultures and a mediator in diverse political, social, educational and economical situations (Park & Son, 2009).

Since research on internet-based assessment is in its infancy in the Iranian context, the present study is mainly an attempt to compare the conventional classes with internet-based classes at secondary school settings to see the extent to which internet-based assessment could facilitate the process of learning for students. In doing so, the researchers worked with the new materials designed and developed by means of computerized tools at secondary school settings and tried to make listening practice more interesting and challenging. To conduct the present study, the researchers mainly decided to pose two research questions as follows:

1) Is there any statistically significant difference between internet-based assessment and paper-based assessment with respect to listening skills?

2) Is there any statistically significant change in the mean scores of the students in internet-based group?

Literature Review

Assessment: Nature and definitions

There is a large volume of published studies describing the role of assessment in educational settings (Fulcher, 2012). However, what is necessary for teachers is the purpose of assessment since the purpose for which the teachers assess students determines its rationale, design, use, and interpretation of results. Popham (2014) believes that classroom-based assessment has instructional purpose (i.e., to adjust instruction to student level) and accountability purpose (i.e., to provide information to administrators). On the other hand, assessment specialists such as William (2008) classify classroom assessment purposes into two broad types: formative and summative. Assessment used for a formative purpose is typically associated with enhancing instruction and improving learning, whereas a summative purpose is relevant to summing up learning achievements to be communicated to administrators and/or other relevant stakeholders. Furthermore, classroom assessment purposes have been classified into four types recently labelled as: "assessment is for teaching", "assessment as learning" (Earl, 2013), "assessment for learning" and "assessment of learning" (Popham, 2014).

To clarify the distinction between assessment of learning and assessment for learning, the researchers believe that in assessment of learning any score is a testing score but in assessment for learning, any score is a learning score, that is, the ultimate aim of assessment is promoting learners' development. In this line, Black and Wiliam (1998) comment that

assessment for learning is an integral part of classroom practice. Today, such attitude towards assessment has resulted in an increased focus on assessment in language programs and the role it performs in enhancing learning (Rea-Dickins, 2008).

Actually, classroom assessment puts emphasis on classroom context and exclude the term "testing" which has connotations with standardized paper and pencil tests and/or large-scale tests (Rea-Dickins, 2008). Poehner (2007) also refers to classroom assessment as the procedures by which students' performance are interpreted in terms of learning goals and instruction processes, as opposed to a finished product measured by large-scale tests.

Furthermore, in discussing classroom assessment process, Popham (2014) argues that an assessment process within an educational setting typically comprises the following key components: defining the purposes of the assessment, constructing or selecting assessment methods to collect evidence of learning, interpreting assessment outcomes collected, grading decision making, recording assessment information, and reporting assessment results to relevant stakeholders comprising students, parents, administrators, potential employers and/or teachers themselves. Finally, some scholars in assessment process believe that validity and reliability characteristics play crucial roles in providing accuracy, fairness, and appropriateness of the interpretations and uses of assessment results and these quality measures should be integrated into the assessment process (Popham, 2014).

Listening skills

Much of the current literature on listening skills in second language acquisition pays particular attention to oral communication skills. Indeed, listening is recognized as the natural precursor to speaking; "the early stages of language development in a person's first language (and in naturalistic acquisition of other languages) are dependent on listening" (Nation & Newton, 2009, p. 37). Development of listening proficiency has been recognized as a paramount component of foreign language teaching and learning; however, many scholars are on this belief that listening comprehension is often treated as the *Cinderella* skill of second language instruction (Nunan, 1997; Vandergrift, 2004). In addition, the research in the arena of listening comprehension is "still in its infancy" (Omaggio-Hadley, 2000, p. 184). Also, a lack of second language or foreign language researches has been recurrently highlighted (Vandergrift, 2005). Vandergrift (2005) notes that second language teaching practices have predominantly highlighted merely reading, writing, and speaking as the skills one needs to develop in a successful language acquisition. This is probably due to the fact that before 1970's listening was mostly regarded as a receptive skill in language learning (Johnson, 2008). However, the significance of listening skill was not fully recognized until the early 1970's, when scholars paid an unprecedented heed to listening as a key factor in facilitating and developing language learning (Vandergrift, 2007). Such studies, however, being based on Audio-lingual method, considered listening as a passive, and receptive skill.

Despite its importance, L2 learners mostly consider listening as the most challenging language skill to learn (Hasan, 2000). As Vandergrift (2007) points out, one of the reasons might be that learners are not taught how to learn listening effectively. However, listening has changed its role from a passive activity which deserved less class time to an active process through which language acquisition takes place (Vandergrift, 2004). Listening is now widely

accepted as an essential skill that enables language acquisition to take place, both in mother tongue and in second or foreign language (Rost, 2002). However, the learning environment in foreign language learning or second language acquisition (SLA) is not as supportive as first language acquisition. While listening is now recognized as an active mental process, it is still difficult to describe (Vandergrift, 2005). In addition, Rost (2002) acknowledges the complexity of listening comprehension process, stating that "if we want to measure [assess and teach] it, we must understand how that process works" (p. 15).

Another problem related to listening comprehension studies conducted in L2 is that they mainly highlight the product-oriented models which typically measure listening ability via quantitative research methods (Rost, 2002). Vandergrift (2007) remarks that quantitative approaches are able to "tell us something about the product, i.e., the level of listening success, [but they] tell us nothing about the process; i.e., how listeners arrive at the right answer or why comprehension breaks down" (p. 192). Ur (1984) categorizes listening into two types: Listening for perception and listening for comprehension. Lund (1990) also described a listening taxonomy as listener function and listener response. He defined function as "the aspects of the massage the listener attempts to process" (p.105). He explained that the six significant functions in teaching second language are identification, orientation, main idea comprehension, detail comprehension, full comprehension, and replication.

Buck's (2001) model of the listening construct constitutes the most widely acceptable one in testing listening comprehension. Buck (2001) put forward a model for listening based on Bachman and Palmer's (1996) model of communicative language ability. In this model, many different types of knowledge are involved in listening comprehension being both linguistic knowledge (phonology, lexis, syntax, semantics, discourse, etc.) and non-linguistic knowledge (knowledge about the topic, context, and world). Buck's perspective for listening comprehension includes both types of knowledge i.e., bottom-up and top-down are needed.

A brief history of CALL

Applications of technology in education is not a recent story, but applying technology in language learning is in its infancy for language learners, teachers and scholars. Computer-assisted instruction was first used in 1950s for other purposes than language teaching (Tafazoli & Golshan, 2014). CALL's history is brief enough to be well-documented, but it points to an area of study which suffers from fragmentation and a lack of scientific rigor. CALL is a young branch of applied linguistics and is still establishing its directions (Beatty, 2013). In 1984, CALL shaped an integral part of the beginners' course, run by the German Department of Aberdeen University, using programs designed by Gordon Burgess (Beatty, 2013).

CALL's origins and development trace back to the 1960's (Delcloque, 2000). The appearance of personal computer (PC) brought computing within the range of a wider audience, causing in a growth in the development of CALL programs and a flurry of publications in the late 1970s (Davies, 2000). Due to the fact that in the previous studies CALL has developed into a symbiotic relationship between the development of technology and pedagogy, Warschauer and Healey (1998) divided the development of CALL into three phases: Behaviorist CALL, Communicative CALL and Integrative CALL (Multimedia and the Internet). Bax (2003) perceived the three phases as restricted, open and integrated, and there have been several other attempts to categorize the history of CALL. Chapelle (1998) outlines a list of 13

CALL handbooks came out during 1980s, and in both structural and communicative CALL, the teacher usually performs the role of the mediator between the computer and the students throughout the learning process.

Esling (1991) developed a menu of task-based CALL activities to enhance productive email exchanges between teachers at two Canadian secondary schools. In these activities, teachers were guided to describe photographs, give directions, or express an opinion. The role of computer software was to help deliver visual materials for description, process word documents, or provide interactive simulations. Debski (2000), using integrative CALL, aimed to utilize networked computers as a means to involve learners in meaningful, large-scale collaborative activities. He concluded a close link between learning processes, objectives, and a student ownership of the outcomes. Alongside mainstream computer-supported collaborative learning (Land & Hannafin, 2000), meaningful interaction and authentic project work are recently emphasized. Authentic discourse offers the basis for learning material.

PC-based activities

One type of CALL activity is communicating. This includes e-mail exchanges, student discussions with each other or with their teacher on LANs, MOOs (sites on the internet where student do role-playing games and talk with each other), and real-time chat (Fotos & Browne, 2004). These activities are particularly useful for foreign language teaching where students share the same L1 because they create the need to use the foreign language for authentic communication (Fotos & Browne, 2004). Another CALL activity is use of multimedia. This includes course ware presented on CD-ROM or online for study of specific skills such as pronunciation or grammar, and integrated skills-based or communicative practice where hyperlinks allow teachers to access a range of supplementary material for learning support. More recently, online courses provide opportunities for students to interact in small groups through online tools (Kiddle & Prince, 2019).

In the same vein, other CALL activities involve the internet, such as web searches for information and teacher construction of home pages. Related to this is the field of information literacy, a concept similar to computer literacy, referring to the ability to obtain information from the internet and processing it selectively and critically. The tremendous amount of online resources means that teacher evaluation of web sites and L2 learning materials has now become an important aspect of Internet-based activities. An additional use of CALL is concordancing and referencing, or using a corpus to examine the range of usages for grammar and vocabulary items, and using online dictionaries for definitions and usage information. Yet, another significant use of CALL is distance learning. In the United States, United Kingdom and Europe, many college professors now teach some or all of their courses online (Fotos & Browne, 2004).

Van Han and Van Rensburg (2014) declared in their study that TOEIC listening with CALL required the learners to participate in communicative and interactive activities in each lesson so that the learners can face the authentic knowledge and content of language. As being gathered from the statistics, 16% of the learners showed their interest in learning with the traditional method (listen to the teacher or CD player to do the listening tasks). In contrast, 36% of them selected communicative tasks with CALL. Startlingly, 48% of the learners felt

excited to participate in the tasks prepared with two methods at the same time. This inferred the roles of teacher as a speaker, helper and director of class activities.

The literature does not have sufficient research on the application of assessment techniques in the Iranian context and the previous works on listening skills is also underrated so that the use of new assessment techniques and the application of cyber-communication and CALL tools and activities are in dire need of more research, especially in the school settings. For, the literature lacks comparative and experimental studies on the use of technology in the classroom for improving language skills, specifically with respect to listening skills, and doing experimental works by professionals in ELT can yield fruitful results.

Method

Participants

A pool of 36 male students, who were studying English language at secondary school, were selected to participate in this study. They were selected based on convenient sampling from Molla Sadra secondary school located in Tehran, Iran. They were EEL learners at ninth grade, and all of them had studied English as a compulsory subject in the previous years. They were introduced to the book "Prospect to English 3". The age of the students was between 15 and 16. In order to make homogeneous the subjects in terms of their language proficiency, KET, developed by Cambridge English Language Assessment (2001) for schools, was used. There were 45 students, but seven of the students' scores did not fall one standard deviation below or above the means so that their scores in the pre-test and post-test were not included in this research work. However, they were not informed about it, and they were behaved as other students in the class. The students were selected from one control class and one experimental class. Each class included 18 students which is an acceptable sample size since, as agreed by several scholars in the field of applied linguistics, "the sample size for comparative and experimental studies should be 30 participants, at least 15 participants in each group" (Dörnyei, 2007, p. 99). They came to the classes two times a week and 90 minutes in each session.

Instruments

The following instruments were utilized by the researchers to gather the relevant data: The first instrument used by the researchers to measure students' listening abilities was *KET for Schools* which was designed by Cambridge English Language Assessment for school (2001). It includes the four basic language skills including listening, speaking, reading and writing. The listening and speaking section of the test were used to homogenize the students. The listening test included five sections and 25 questions. Each question carried one mark. The speaking section had 15 questions and one score was devoted for each question. The scoring procedure was calculated out of 40.

The listening sections of two KET versions, designed by Cambridge English Language Assessment for school (2001), were, again, used to measure students' performances in the pre-test and post-test. The listening test for each of pre-test and post-test consisted of 25 question items and their scores were calculated out of 25 for both pre-test and post-test.

Procedure

The study followed a straightforward procedure to conduct the study. Before the treatment phase, the students were made homogenous in terms of language proficiency by

means of KET, and following this, they were randomly assigned into two groups, including the control group and the experimental group. Students' scores on the listening section of KET was used as the pre-test scores. The control group was exposed to paper-based listening activities and the experimental group was exposed to internet-based listening activities. To do so, the teacher worked with the first group on the listening sections of their textbook "Prospect to English 3", and with the second group, the materials were presented through the internet-based tool. The teacher worked with the students on the activities and tasks during the class time, and they were asked to reproduce what was said in the audio clips in group A and in group B.

All of the participants were native speakers of Persian language. The treatment phase lasted for fifteen sessions, and the students came to the classes twice a week. Each session lasted 90 minutes, however, the students were exposed to these activities the last thirty class time to work on their listening skills. Along with the pre-test and post-test, during the treatment, the students were asked to participate in the quizzes, given by the teacher every five sessions too.

Students, who were exposed to internet-based assessment, were required to enter the site, developed by the teacher which provided the students with the listening quizzes, tests and guidelines, including assessment techniques such as leading questions and hints. Students could make use of computer-aided materials simultaneously, including the text, the audio and the test answers. The teacher worked with the students on the listening audio clips, selected from their class textbook, and the students were learning and assessing their listening skills through the internet-based activities.

The data were mainly collected from the students prior to and following the treatment phase. The students were required to participate in the pre-test of listening and, finally, in the post-test of listening. The scores obtained from the pre-test and post-test were interval and the scoring procedure was marked by 25. The steps followed are presented in Table 3.

Table 1. Steps of the procedure

Sessions	Procedure
Session 1	*Test of Homogeneity*: Dividing the students into two groups and administering the test of proficiency.
Session 2	*Pre-test of listening skills*: Administering the pre-test to the students to measure their listening abilities before the treatment phase.
Session 3 to 17	*Treatment phase*: Exposing the experimental group to internet-based assessment and the control group to the paper-based assessment including 15 sessions.
Session 18	*Post-test of listening skills*: Administering the post-test to the students to measure their listening abilities following the treatment phase.

The students, who were exposed to internet-based activities during the treatment, gained mastery over the use of internet-based listening software, and unlike the earlier sessions, in the final sessions, they could do the task with fewer hints, prompts and leading questions, provided by the teacher, and they could focus more on the listening tasks instead of

spending their times on the web site. The data obtained from the students' performance on the tests were, then, entered into the SPSS software version 22, and the researchers opted for statistical formulas including One-way ANCOVA and One-way ANOVA for the analyses of the data.

Findings

Having analyzed the data obtained from the study, the researchers reports the findings as follows.

Internet-based assessment vs. paper-based assessment and listening skills

As for the first research question of the study regarding the significant difference between internet-based assessment and paper-based assessment with respect to listening skills, the researchers performed ANCOVA formula. To measure students' listening performance in the pre-test and post-test, KET was used consisting of 25 question items, and their scores were calculated out of 25 for both pre-test and post-test. Before running One-way ANCOVA, a number of assumptions were needed to be met, including the linearity for each group, the homogeneity of regression slopes between the covariate and the dependent variable for each of the groups and the assumption of equality of variance.

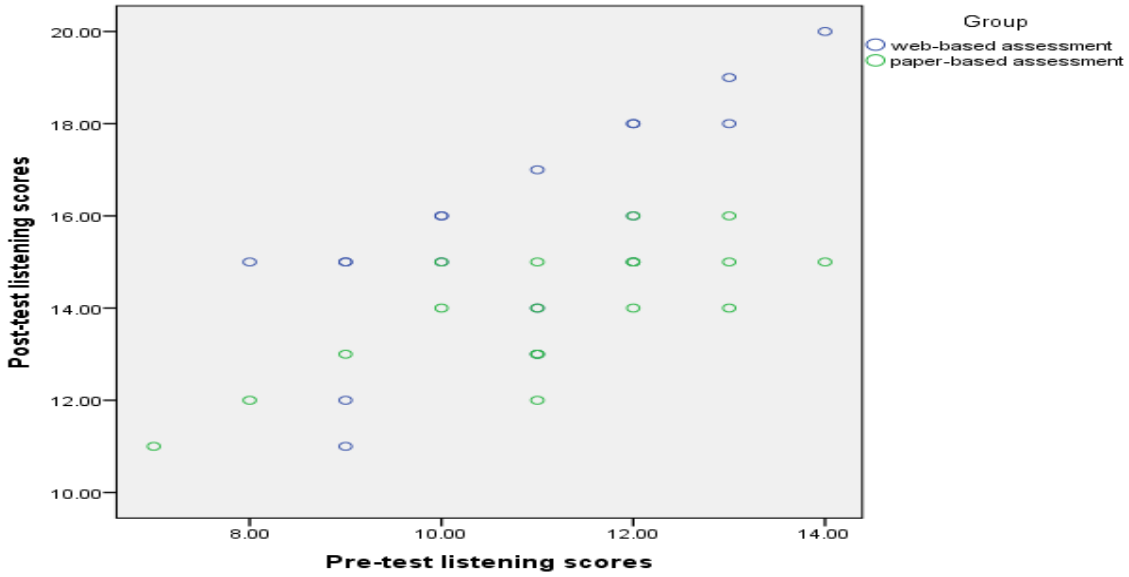

Figure 1. The linearity for internet-based assessment and paper-based assessment

First, the general distribution of scores for control group and experimental group was checked. As displayed in Figure 1, there appeared to be a linear (straight-line) relationship for each group. Indeed, there had been no indication of a curvilinear relationship. The relationship was clearly linear, so there was no violation in the assumption of the linear relationship. Moreover, there was not violation of the assumption of homogeneity of regression slopes since the Sig or probability value was .09, safely above the cut-off. This supported the earlier conclusion gained from an inspection of the scatter plots for each group. Furthermore, there

was no violation of the assumption of equality of variance because the Sig. value was .09, as indicated by Levene's test of equality of error variances for listening scores which was larger than .05.

Table 2. ANCOVA test for listening skills

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	91.43[a]	2	45.71	23.25	.00	.58
Intercept	32.67	1	32.67	16.62	.00	.33
Pre-test	64.74	1	64.74	32.93	.00	.50
Group	33.68	1	33.68	17.13	.00	.34
Error	64.86	33	1.96			
Total	8107.00	36				
Corrected Total	156.30	35				

a. R Squared = .585 (Adjusted R Squared = .560)

The main ANCOVA results were presented in Table 2. The line corresponding to the independent variable (in this case Group) was followed and read across to the column labeled Sig. Since the significant value was .00, which was *less* than .05, the groups differed significantly. Indeed, after adjusting for pre-intervention scores, there was a significant difference between the two groups on post-test scores on the listening test [F (1, 33) =17.13, p=.00, partial eta squared=.34]. That is, there was a significant difference in the students' listening scores in internet-based assessment and paper-based assessment.

Table 3. Post hoc test

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.[b]	mean	SD
internet-based assessment	paper-based assessment	1.94*	.46	.00	15.72	2.37
paper-based assessment	internet-based assessment	-1.94*	.46	.00	14.00	1.41

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

As displayed in Table 3, running Post hoc test for the exact difference between the two groups, the asterisk for the mean difference showed that the mean difference was significant; it was 15.72 for internet-based group and 14.00 for paper-based group so that the internet-based group outperformed the paper-based group. Therefore, the results revealed that students who were exposed to internet-based assessment could improve their listening skills significantly so

that the use of internet-based tools in the secondary school settings was found to be effective. Actually, the teachers could diagnose and assess the students more easily as the assessment tools were at the service of learning not testing the students. Indeed, the use of technology and internet-based materials facilitated the task of teaching for the teachers, and more importantly, they could help students become autonomous since they could do the tasks themselves with less guidance provided by the teacher in the final sessions.

Significant change in the mean scores of internet-based group

As for the second research question of the study regarding the significant change in the mean scores of the students in internet-based group, the researchers ran One-way ANOVA since there were three sets of scores for the students, who were exposed to internet-based assessment. Before running the ANOVA test, the researchers examined the normality of data.

Table 4. Tests of normality for sets of scores

Time		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
internet-based assessment	Time 1	.15	18	.20*	.95	18	.56
	Time 2	.16	18	.20*	.93	18	.25
	Time 3	.15	18	.20*	.96	18	.78

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

As displayed in Table 4, the results obtained from the Kolmogorov-Smirnov statistic showed that distribution of scores are normal. Since a non-significant result (Sig value of more than .05) indicates normality; in this case the significant value was .20 for each Time, suggesting no violation of the assumption of normality. Moreover, there was not violation of the assumption of homogeneity of regression slopes since the Sig or probability value was .34, safely above the cut-off. This supported the earlier conclusion gained from an inspection of the scatter plots for each group.

Table 5. Descriptive statistics for three sets of scores

	N	Mean	Std. Deviation	Std. Error
Time 1	18	11.66	1.64	.41
Time 2	18	12.61	1.68	.35
Time 3	18	15.72	2.37	.33
Total	54	13.33	1.89	.24

As shown in Table 5, the results of Descriptive Statistics showed that the means and standard deviations of the scores were as follows: Time 1 (M= 11.66; SD=1.64), Time 2 (M= 12.61; SD=1.68), and Time 3 (M= 15.72; SD=2.37).

Table 6. ANOVA test for internet-based assessment

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	162.11	2	81.05	21.77	.000
Within Groups	189.88	51	3.72		
Total	352.00	53			

Using the statistical formula of one-way ANOVA, the researcher examined the significant difference between the students' scores obtained from Time 1, Time 2 and Time 3, as measured by the ANOVA. There was a statistically significant difference at the $p<.05$ level in students' scores for the three sets of scores [$F(2, 51) = 21.77$, $p=.00$] (see Table 6).

Table 7. Multiple comparisons for three sets of scores

		Mean			95% Confidence Interval	
(I) Time	(J) Time	Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Time 1	Time 2	-.94	.64	.31	-2.49	.60
	Time 3	-4.05*	.64	.00	-5.60	-2.50
Time 2	Time 1	.94	.64	.31	-.60	2.49
	Time 3	-3.11*	.64	.00	-4.66	-1.55
Time 3	Time 1	4.05*	.64	.00	2.50	5.60
	Time 2	3.11*	.64	.00	1.55	4.66

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

As displayed by Table 7, Post-hoc comparisons, using the Tukey HSD test, indicated the exact difference between the mean score for students' scores that was significantly different between the three Time span. The asterisks (*) next to the values listed showed that scores obtained at Time 3 were significantly different from Time 1 and Time 2 at the p<.05 level. The exact significance value was given in the column labelled *Sig.* In the results presented above, the students at Time 3 showed significant improvement in their listening skills in internet-based assessment. Therefore, the results confirmed that the students did not improve significantly in internet-based group during the first five sessions and the second five sessions, but their scores improved effectively following the third five sessions. This shows that the students were spending their time more on how to learn by means of internet-based tools, and then, after ten sessions, they could handle the task and gained mastery over the web site so that

they could focus more on the listening tasks, and the use of internet-based tools enhanced their listening skills in this time span.

Discussion

The results of the present study revealed that the students who were exposed to internet-based materials could perform better than those who were exposed to paper-based materials, following the treatment phase. The results of the study confirmed that there was a statistically significant difference in the mean scores of the students who were exposed to internet-based materials after ten sessions. The point is that attention to computer-aided and internet-based materials and programs in second language teaching and learning had recently been increased in the domain of second language acquisition. According to Lai and Kritsonis (2006), while the use of task-based materials has been explored in SLA, there are many more areas of synergy that arguably deserve further attention among which computer-aided materials with focus on learner-based education. Computer-aided materials could be further developed to take account of a branch of education that has largely been overlooked thus far: Computer-aided education.

Although some previous research on the issue shows that measuring students' language knowledge is hard by means of internet-based resources in the domain of education (Johnson, 2002), the present study illustrated that working on internet-based assessment with the students highlights the complexity of interactional and instructional factors in language teaching and learning. That is why there is a challenge to the idea that CALL is the same thing across all kinds of conditions (Street, Pishghadam, & Zeynali, 2015). Moreover, the present study is in line with the previous works such as the studies undertaken by Shams and Modarresi (2015) and Fotos and Browne (2004) who found that students would like to work with the computer and computer resources as well as computer-based tools such as e-mail and additional software tools. Indeed, internet-based resources can provide a new environment for the students that would be more interesting to them in comparison to the traditional text-based instruction, and this viewpoint has already been supported by the previous research conducted by Cameron (1999) and Lambropoulos, Christopoulou and Vlachos (2006) who confirmed that computerized means enhance the opportunities for interactive learning.

Actually, secondary school students studying English language in their academic careers are not well familiar with the use of software tools in their classroom practice, and they have indeed phobia in employing and working with software tools. This may be due to the fact that they are not equipped with technological tools in the classroom. They seem to be accustomed to the paper-based materials in the classes while the rapid expansion of technology brings the need for the students to learn how to work with them. To mitigate the problems inherent in our educational system and teaching English agenda, professionals in ELT can make use of the alternative learning and assessment tools one of which is computerized and internet-based materials. As it is mentioned earlier, CALL, as a new teaching and learning tool, is to be used as a bridge. The educational system uses this bridge to terminate the traditional learning and teaching system and to arrive at the modern environment in translation pedagogy.

Finally, according to Lamy and Goodfellow (1999), the idea that CALL programs can act as scaffolding providing the learners with appropriate feedback is interesting and arguably deserves further attention. The results of the present study also yield that computer-based tools

not only can offer remedial help but also can help students assess themselves and become self-directed and approach autonomy in learning by means of self-awareness and self-observation. As commented by Preece, Rogers and Sharp (2002), providing both support from the teachers and feedback from the automatic computer contribute well to the development of learning.

Conclusion and Implications

Major conclusions can be drawn from the present study. Indeed, with the increasing expansion of technology in educational settings, the focus on computerized mediated tools and CALL literacy at the secondary school environment in our country would help students to learn English language more effectively. Actually, the potentiality for computer-enhanced language learning have recently been underscored by educators in the context of schools in Iran. Today, the technological tools, which are facilitative in the everyday life transactions, has increased the applicability of computer-aided materials in important aspects of life among which is academic life. Nearly most of the secondary school students are familiar with communication software programs such as Telegram, WhatsApp, Instagram, or so. The researchers concluded that students are fond of new software programs and so the teachers should equip them with those internet-based programs that are more helpful to them in their lessons.

This study has provided a deeper insight into the application of computer-aided resources in the form of internet-based activities, and in identifying the major effectiveness in this regard, the statistical analyses of the data revealed that these types of activities assist students in improving their listening skills and teacher can assess students better by means of these sorts of activities. The researchers believe that technological tools are integrating into the individual life of the learners and help them learn independent of time and place.

The researchers also came to the conclusion that the time interval is also important in this regard, and students need to work with the internet-based tools and resources for at least two or three months so that they feel they are progressing, and acquaintance with the computer-aided tools would make them accustomed to these types of materials by which they can work on the language skills interactively at home through web. As Kern and Warschauer (2000) noted, the development of the social constructivist language pedagogy in CALL and L2 learning have created teaching principles according to which the target language is acquired through interaction among the learners. In the near future, they will have no phobia to take part in national and international computerized tests. The fact of the matter is that students who are more accustomed to working with computer-aided tools are more successful in TOEFL IBT since they should reply to the questions in front of the computer and their familiarity with and their mastery over computer software help them much to have a better performance on their examination.

The results of the present study can add credence to the significance of CALL in Iran and offer pedagogical implications for students, teachers and syllabus designers. As for the students, they are recommended to become more familiar with computer-aided tools and try to interact with other students and their teachers via such tools as internet and computerized versions of second language acquisition. They should use computerized activities like internet-based programs and e-portfolio techniques to monitor their own learning and assess their own

development. Students can benefit from not only human resources and feedback from the teachers but also from electronic feedback and computerized assessment.

As for second language teachers, familiarizing the students with internet-based programs with its rather simple design can be a preliminary step towards the introduction of computer-aided tools that may have a more complex system. They should make students gain command over the CALL facilities step by step so that the students make use of technology in enhancing their language knowledge. Teachers need to expand their knowledge of CALL and do not just stick to the textbooks introduced by language policy in their classes, and with reference to the sociocultural context in which they are teaching, they ought to use such facilities as supplementary resources. For example, they can incorporate the neglected areas in the classrooms such as knowledge of collocations (see Modarresi, 2009) into internet-based programs.

As for syllabus designers, they can develop materials in computerized format and provide secondary school students with CALL programs and activities as complementary resources and can design materials for the students to work on English language at home and independent of time and space to improve their language proficiency and feel comfort with CALL programs. Indeed, most of the secondary school learners of EFL run into problem while developing their communicative competence beyond the classroom mainly due to the fact that they do not have a supportive learning environment to be exposed to English language for communicative purposes. For this reason, some specific endeavors are required to help Iranian students develop their language learning experiences and practice the target language outside the classroom. This need can be found in the Iranian government's special emphasis on English language proficiency and computer literacy in the spirit of globalization.

Suggestions

The present study mainly examined the significant difference between the mean scores of the group exposed to internet-based assessment and the group exposed to paper-based assessment with respect to listening skills in the context of secondary schools. Another study is needed to work on the internet-based assessment with respect to other language skills including speaking, reading and writing in secondary school settings. Also, researchers are recommended to replicate this study in other parts of the country to demonstrate its validity and also to confirm the effectiveness of internet-based assessment on the motivation and self-efficacy of the students. Finally, another study can investigate the extent to which the effectiveness of computer-aided materials vary with respect to the gender, age and socioeconomic factors of the students.

References

Al-Kadi, A. (2018). A review of technology integration in ELT: From CALL to MALL. *Language Teaching and Educational Research, 1*(1), 1-12.

Bachman, F. L., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bacon, S.M., (1992). The relationship between gender, comprehension, processing strategies, and cognitive and affective response in foreign language listening. *The Modern Language Journal*, *76* (2), 160-178.

Bancheri, S. (2006). A language teacher's perspective on effective courseware. In P. D. Randall & A. H. Margaret (Eds.), *Changing language education through CALL* (pp. 31-47). New York: Routledge.

Bax, S. (2003). CALL- past, present and future. *System, 31*, 13-28.

Beatty, K. (2013). *Teaching & researching: Computer-assisted language learning*. London: Routledge.

Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.

Cambridge English language assessment for schools (2001). *KET for schools*. Cambridge: Cambridge University Press.

Cameron, K. (1999). *CALL: Media, design and applications*. Lisse: Swets & Zeithlinger.

Chapelle, C. A. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. *Language Learning and Technology,* 2, 22-34.

Chien, C. N., & Kao, L. H. (2004). Examining the inter-relationship of metacognitive strategy training, listening comprehension, and learning attitude in EFL training. *Chung Yuan Journal*, 32(2), 241-254.

Chien, C. N., & Wei, L. (1997). A preliminary investigation of the listening strategies of EFL learners. *Chung Yuan Journal*, 25(2), 45-66.

Davies, G. (2000). *CALL (computer assisted language learning): Routledge encyclopedia of language teaching and learning.* London: Routledge.

Debski, R. (ed.) (2000). Project-oriented CALL: Implementation and evaluation. *Computer Assisted Language Learning*, *13*, 4–5.

Delcloque, P. (2000). History of CALL. Received January 3, 2017 from: http://www.history-of-call.org

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

Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2nd ed.). Thousand Oaks, California: Corwin.

Esling, J. H. (2013). Researching the effects of networking: Evaluating the spoken and written discourse generated by working with CALL. In P. Dunkel (ed.), *Computer-assisted language learning and testing: Research issues and practice* (pp. 111–31). New York: Newbury House.

Fotos, S., & Browne, C. M. (2004). The development of CALL and current options. In S. Fotos & C. M. Browne (Eds), *New perspectives on CALL for second language classrooms* (pp. 3-14). NJ: Lawrence Erlbaum Associates, Inc.

Fulcher, G. (2012). Assessment Literacy for the Language Classroom. Language Assessment Quarterly subscription information. Retrieved June 12, 2017 from: http://www.tandfonline.com/loi/hlaq20.

Graham, S. (2017). Research into practice: Listening strategies in an instructed classroom setting. *Language Teaching*, *50* (1), 107-119.

Hasan, A. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture and Curriculum*, *13*, 137-153.

Hosseini, H. S., & Modarresi, Gh. (2015). Construction, validation, and application of CALL knowledge scale for EFL learners and its relationship with their L2 writing skills. *Foreign Language Research Journal, 5 (2),* 309-311.

Johnson, E. M. (2002). The role of computer-supported discussion for language teacher education: What do the students say? *CALICO Journal, 20*(1), 59-79.

Johnson, K. (2008). *An introduction to foreign language learning and teaching* (2nd ed). Harlow, England: Pearson.

Kern, R., & Warschauer, M. (2000). Theory and practice of network-based language teaching. In M. Warschauer & R. Kern (eds.), *Network-Based Language Teaching: Concepts and Practice* (pp. 1–19). Cambridge; Cambridge University Press.

Khorami, F. & Modarresi, Gh. (2019). A Rasch-based Validation of the Evaluation Rubric for Consecutive Interpreting Performance. *Sendebar, 30*, 221-244.

Kiddle, T., & Prince, T. (2019). Digital and online approaches to language teacher education. In S. W. Walsh, & S. Mann (Eds.), *The Routledge handbook of English language teacher education* (pp. 111-124). London: Routledge.

Lai, C. C., & Kritsonis, W. A. (2006). The advantages and disadvantages of computer technology in second language acquisition. *National Journal for Publishing and Mentoring Doctoral Student Research, 3*(1), 1-6.

Lambropoulos, N., Christopoulou, M., and Vlachos, K. (2006). Culture-based language learning objects: A CALL approach for a ubiquitous world. In P. Zaphiris, & G. Zacharia, *Computer-aided language learning* (pp. 22-43). London: Information Science Publishing.

Lamy, M. N., & Goodfellow, R. (1999). Reflective conversation in the virtual language classroom. *Language Learning & Technology, 2*(2), 43-61.

Land, S. M., & Hannafin, M. J. (2000). Student-centered learning environments. In D. H. Jonassen & S. M. Land (eds.), *Theoretical foundations of learning environments* (pp. 1–23). Mahwah, NJ: Lawrence Erlbaum.

Langer, J. A. (2001). Beating the odds: teaching middle and high school students to read and write well. *American Educational Research Journal*, *30*(4), 837-880.

Levy, M. (1997). *Computer-assisted language learning: Context and conceptualization*. New York: Oxford University Press.

Lin, C.-H., Warschauer, M. & Blake, R. (2016). Language learning through social networks: Perceptions and reality. *Language Learning & Technology, 20*(1), 124-147.

Lund, R. J. (1990). A taxonomy for teaching second language listening. *Foreign Language Annals, 23*, 105-115.

Modarresi, Gh. (2009). Collocational errors of Iranian EFL learners in written English. *TELL, 3* (1), 135-154.

Modarresi, Gh., & Alavi, S. M. (2014). Designing and validating a test battery of computerized dynamic assessment of grammar. TELL, 14 (2), 1-29.

Nation, I. S. P., & Newton, J. (2009). *Speaking.* New York & London: Routledge.

Nunan, D. (1997). Listening in language learning. *The Language Teacher, 21* (9), 47-51.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle.

Omaggio-Hadley, A. (2000). *Teaching language in context* (2nd ed.). University of Illinois: Heinle & Heinle.

Park, C. N., & Son, J. B. (2009). Implementing computer-assisted language learning in the EFL classroom: Teachers' perceptions and perspectives. *International Journal of Pedagogies and Learning, 5*(2), 32-50.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: Beyond human-computer interaction*. New York: John Wiley & Sons.

Poehner, M. E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal*, *91*(iii), 323-340.

Popham, W. J. (2014). *Classroom assessment: What teachers need to know* (7th ed.). Boston: Pearson Education.

Rea-Dickins, P. (2008). Classroom-based language assessment. In E. Shohamy & N. H. Hornberger (Eds.) *Encyclopedia of language and education: Language testing and assessment* (pp. 257–271). New York, NY: Springer.

Rost, M. (2002). Listening tasks and language acquisition. *JALT Journal, 18,* 18-28.

Shepard, L. A. (2013). Foreword. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. xix-xxii). Los Angeles: Sage.

Street, B., Pishghadam, R., Zeinali, Sh. (2015). Changes and challenges of literacy practices: A case of a village in Iran. *IJSCL, 3*(1), 16-27.

Tafazoli, D., & Golshan, N. (2014). Review of computer-assisted language learning: History, merits & barriers. *International Journal of Language and Linguistics*, *2*(5), 32-38.

Ulla, M. B., & Perales, W. F. (2020). Integrating Internet-based applications in English language teaching: Teacher practices in a Thai university. *Issues in Educational Research 30* (1), 365-380.

Ur, P. (1984). *Teaching Listening Comprehension.* Cambridge: Cambridge University Press

Van Han, N., & Van Rensburg, H. (2014). The Effect of computer assisted language learning (CALL) on performance in the test of English for international communication (TOEIC) listening module. *English Language Teaching, 7*(2), 30-41.

Vandergrift, L. (2004). Orchestrating strategy use: Toward a model of the skilled second language listener. *Language Learning, 53*(4), 463-496.

Vandergrift, L. (2005). Relationships among motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied Linguistics, 26*(1), 70-89.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*, 191-210.

Warschauer, M., & Healey, D. (1998). Computers and language learning: an overview. *Language Teaching, 31*, 57–71.

William, D. (2008). Balancing dilemmas: Traditional theories and new applications. In A. Havnes & L. McDowell (Eds.), *Balancing dilemmas in assessment and learning in contemporary education* (pp. 267-281). New York: Routledge.

Yeldham, M., & Gruba, P. (2016). The development of individual learners in an L2 listening strategies course. *Language Teaching Research, 20*(1), 9-34.