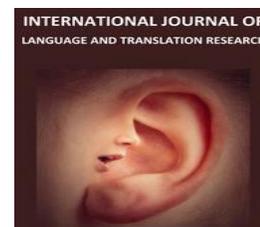# Impact of Online-learning Activities on Improving IELTS Speaking Performance of Iranian EFL Learners: An Experimental Comparative Study

Omid Rezaei Dastgerdi[1] *

[1]Independent Researcher, Toronto, Canada

Abstract

This paper investigated the possible impacts of using online activities and the speaking ability of Iranian EFL learners. For this reason, 30 participants were selected randomly and divided into two experimental groups and one control group. The participants were given a speaking IELTS test at first. After that, the experimental groups received online-learning activities regarding speaking abilities in two various forms with teacher's intervention and without teacher's intervention; whereas the control group did not receive. Then, the three groups were given a post-test on speaking ability to compare the possible impacts of the online- learning activities. The results showed that the participants in the experimental groups yielded more positive results. In addition, between two ethe experimental groups, the experimental groups with the teacher's active role had better results. The findings of this research can have very useful implications for IELTS courses in Iran.

Introduction

Modern technology and its advances are increasingly influencing educational settings in a way that language teachers and language learners are more and more dependent on these technologies (Dalgaro & Lee, 2010). As a matter of fact, due to the time restrictions and language learners have and in line with the fast growing of internet, e-learning language teaching and learning is going

---

from strength to strength and many students prefer to enroll in online language classes in line with the interests of teachers in holding online courses (Al-Rawahi & Al-Mekhlafi, 2015). In addition, with the enhancement in the need for learning and teaching with communication tools and via virtual space and thanks to advances in internet technology educational institutions are increasingly adding online classrooms into their syllabi and more courses are being offered in this way (Khayef & Amini, 2015).

Among four skills of language, speaking is regarded as the most noticeable of all in second language teaching and learning (Sadiku, 2015). As Richards & Renandya (2002) assert: "A large percentage of the world's language learners study English in order to develop proficiency in speaking" (p. 201). Regarding language testing, speaking is different from other skills in that it is interactive and has to be tested in an interactional mode in which there is a mutual dialogue and/ or discussion between the tester and the teste.

Speaking assessment has been a part and parcel of worldwide large-scale language proficiency tests like IELTS, TOEFL, and Cambridge exams like FCE and CAE. In IELTS, the speaking is composed of various sections which are together between 11–14 minutes. These sections include general questions about the applicant and some other related topics, talking about a topic that has been written in a card, and talking about more abstract concepts in relation to a topic in the card. All of these sections are carried out in an interactive mode in which there is a constant interaction between the applicant and the interviewer (Heidari Tajan, 2016).

Teaching and learning a second language with technology and its differences from conventional language teaching has been the topic of many hot debates in the last decades (Allen & Seaman, 2007). In other words, since its advent in distance learning and teaching, online learning has been the subject of many questions about the best way to use this technology in a foreign language learning context and in such tests as IELTS (Rahnavard & Heidar, 2017). It is commonly argued that since online classes suffer from the lack of effective communication channels, therefore, they are less capable of representing the 'social presence' of participants in online classes (Georgieva & Smrikarov, 2004).

Since the application of technology in second language teaching and learning, there has been much research reported on the application of online activities and language learning and teaching and its differences from conventional/traditional language teaching activities (see for example

Baron, 2008; Chua & Don, 2013; Kruk, 2014; Wang & Liao, 2018). However, since applying technology in second language teaching and learning and comparing it to conventional language teaching with an emphasis on the IELTS speaking test and sub-components like fluency and accuracy has not been reported, to the best knowledge of the authors, and due to the paucity of experimental research in this area, this research was put into practice as an erroneous effort to fill this gap.

Research Questions

Q1. Do online-learning activities, as compared to conventional activities have any significant effect on Iranian IELTS candidates' speaking fluency?

Q2. Do online-learning activities, as compared to conventional activities have any significant effect on Iranian IELTS candidates' speaking accuracy?

Q3. Do online-learning activities, as compared to conventional activities, have any significant effect on Iranian IELTS candidates' speaking ability?

Q4. Does teaching with teachers' interference or without teachers' interference have statistically more significant effects on Iranian IELTS candidates' fluency, accuracy and speaking ability?

Methodology

Design

The design of the research was quasi-experimental as it was impossible to assign random sampling. More, this study enjoyed a comparative group design as there were two experimental groups with various treatments.

Participants

The participants of this research were 30 Iranian IELTS candidates at Danesh- Pazhoohan language institute in Isfahan, Iran. They were male and females aged between 20 and 25 years. They were assigned into three groups. One group was the control group and the other two groups were experimental groups. Each group consisted of 10 people.

Instruments

Oxford Placement Test

Oxford Placement Test is a standardized placement test for assessing the level of proficiency of learners. The test is made up of 100 items of reading and grammar of the participants. This test has a standard scoring scale including 5 different band scores for determining the current level of proficiency of learners: beginner, elementary, pre-intermediate, intermediate, and advanced.

Scale for Assessing IELTS Speaking Ability

For assessing and scoring IELTS speaking ability of the participants, there needed to have a valid and reliable scoring rubric. For this purpose, the rubric produced by the British council was used in this research as it was already valid and reliable.

Inter-rater Reliability

The speaking recordings were assessed by two raters. One rater assessed and scored the candidate simultaneously in the exam place. The second rater listened to the recordings and scored them. Both were English teachers and had experience in teaching IELTS preparation courses. They were IELTS certificate holders with a 9-band score in speaking.

Materials for Teaching IELTS

A number of various materials for both groups were used. The first material was IELTS past papers (speaking section only). The past papers were selected as they were commercially available and were authentic in nature. In addition, the instructors used some personally-developed pamphlets for teaching speaking tips for participants. The pamphlet was designed in such a way that it contained a wide range of speaking topics with specialized vocabulary for each topic.

For the experimental groups, two various classes were created. One experimental group was designed in a group in Skype software. In this group, the teacher provided students with a wide range of speaking topics like group conversations, pair conversations, form-based speaking tasks, and group dissuasions. However, the teacher did not attend the forum and did not interfere with the teachings. He only made sure that the participants were all in the forum and participated in the works provided by the teacher. The second experimental group was also created in a forum on

Skype in which the same tasks were given to the participants but the teacher played an active role in such a way that scaffolding and observation were provided by the teacher. The treatment group participated in the classes in the institution and was given traditional teachings.

Procedures

After the participants were assigned to conventional and online groups, they received teachings in class and in an online forum. They attended three times per week in the class and online forum for 90 minutes. Participants in two online forums were not required to attend the classes; they were all invited into a Skype channel which was created by the instructor and in which the same materials as the physical class were offered to them. In the online forum, the participants attended three times per week, each 90 minutes. After that, the three groups sat for an IELTS speaking test to compare the impacts of both conventional and online groups on IELTS speaking.

Data Analysis

In the first stage, a one-sample statistics was conducted on OPT to determine the level of participants was conducted.

Table 1

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
OPT	30	56.8667	5.13765	.93800

As can be seen in Table 1, the mean of the group was 56.86. As a result, the general level of the participants was intermediate. For comparing the frequency of participants in the pre-test, a descriptive statistic was conducted.

Pre-test: Fluency

Table 2

Descriptive Statistics for Frequency

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
Control Group	10	3.8500	.66875	.21148	3.3716	4.3284	3.00	5.00
Experimental A	10	4.4000	.77460	.24495	3.8459	4.9541	3.00	5.00
Experimental B]	10	3.9500	.83166	.26300	3.3551	4.5449	3.00	5.00
Total	30	4.0667	.77385	.14129	3.7777	4.3556	3.00	5.00

As can be seen in Table 2, the mean of the control groups, experimental group 1 and experimental group 2 was 3.85, 4.40, and 3.90; respectively. The total mean of the groups was 4.066. However, in order to see if the groups were homogenous in the pre-test, a test of homogeneity of variance was conducted.

Table 3

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
1.171	2	27	.325

As can be seen in Table 3, the significance was bigger than .0.05. As a result, there was an acceptable index of homogeneity in groups in the pre-test. To test the differences between the three groups, a test ANOVA was conducted.

Table 4

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.717	2	.858	1.481	.245
Groups	15.650	27	.580		
Total	17.367	29			

As can be seen in Table 4, the significance of data between and within groups was .245. In this regard, there was statistically no significant difference between and within groups in protest in terms of fluency.

Table 5

Multiple Comparisons

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Control Group	Experimental A	-.55000	.34048	.257	-1.3942	.2942
	Experimental B]	-.10000	.34048	.954	-.9442	.7442
Experimental A	Control Group	.55000	.34048	.257	-.2942	1.3942
	Experimental B]	.45000	.34048	.396	-.3942	1.2942
Experimental B]	Control Group	.10000	.34048	.954	-.7442	.9442
	Experimental A	-.45000	.34048	.396	-1.2942	.3942

As can be seen in Table 5, the mean between the control group and the experimental group was .257. Therefore, there was statically no significant difference between them. There was also no statistically significant difference between the control group and the experimental as the level of significance was .954. between experimental a and control group and experimental b, was no statistical significance as the level of significance was .257 and .396; respectively.

Table 6

Statistics for Pre-test Fluency

Group	N	Subset for alpha = 0.05
		1
Control Group	10	3.8500
Experimental B]	10	3.9500
Experimental A	10	4.4000
Sig.		.257

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 10.000.

As the data in Table 6 represents, there was statically no significant difference between the three groups in protest at the level of significance is .275 which is bigger than .0.05. However, like the pre-test, statistics were conducted for the post-test.

Table 7

Descriptive Statistics

Post-test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
Control Group	10	4.1000	.73786	.23333	3.5722	4.6278	3.00	5.00
Experimental A	10	5.3500	.94428	.29861	4.6745	6.0255	4.00	7.00
Experimental B]	10	5.8500	1.20301	.38042	4.9894	6.7106	4.00	8.00
Total	30	5.1000	1.20631	.22024	4.6496	5.5504	3.00	8.00

As the data in Table 7 represent, the mean of the control group was 4.100. The mean of experimental A and experimental B were 5.35 and 5.85; respectively. All in all, the total mean of the groups was 5.100 in the post-test. For comparing between-group statistics, an ANOVA test was conducted.

Table 8

ANOVA for Post-test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	16.250	2	8.125	8.454	.001
Groups	25.950	27	.961		
Total	42.200	29			

As the data in Table 8 represent, the significance of the between and within groups was .001. As a result, there found a statistically significant difference between and within groups. For comparing the differences between control groups and experimental groups in Post-test, multiple comparisons were conducted.

Table 9

Multiple Comparisons

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Control Group	Experimental A	-1.25000*	.43843	.022	-2.3371	-.1629
	Experimental B]	-1.75000*	.43843	.001	-2.8371	-.6629
Experimental A	Control Group	1.25000*	.43843	.022	.1629	2.3371
	Experimental B]	-.50000	.43843	.498	-1.5871	.5871
Experimental B]	Control Group	1.75000*	.43843	.001	.6629	2.8371
	Experimental A	.50000	.43843	.498	-.5871	1.5871

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

As can be seen in Table 9, the level of significance between the control group and experimental groups A and B was .022 and .001; respectively. As a result, there was a statistically- significant difference between them in the post-test. The level of significance between experimental A and the control group and experimental B was .22 and .498. Therefore, there was a statistically significant difference between them.

Pre-test: Accuracy

Like fluency, the mean of the groups was considered in terms of accuracy.

Table 10

Descriptive Statistics for Accuracy

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Control Group	10	4.2000	1.08525	.34319	3.4237	4.9763	2.00	5.50
Experimental A	10	3.8500	.74722	.23629	3.3155	4.3845	3.00	5.00
Experimental B]	10	4.3500	.81820	.25874	3.7647	4.9353	3.00	5.50
Total	30	4.1333	.88992	.16248	3.8010	4.4656	2.00	5.50

As can be seen in Table 10, the means of the control group, experimental A, and experimental B were 3.850 and 4.350 respectively.

Table 11

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.644	2	27	.533

As the data in Table 11 show, the significance of variance was .533. There was an acceptable index of homogeneity between them.

Table 12

ANOVA between and Within Groups

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.317	2	.658	.821	.451
Groups	21.650	27	.802		
Total	22.967	29			

As can be seen in Table 12, the significance of the groups was .451. In this regard, there was a statistically-significant difference between and within groups in the pre-test.

Table 13

Multiple Comparisons

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Control Group	Experimental A	.35000	.40046	.661	-.6429	1.3429
	Experimental B]	-.15000	.40046	.926	-1.1429	.8429
Experimental A	Control Group	-T35000	.40046	.661	-1.3429	.6429
	Experimental B]	-.50000	.40046	.436	-1.4929	.4929
	Control Group	.15000	.40046	.926	-.8429	1.1429

| Experimental | Experimental A | .50000 | .40046 | .436 | -.4929 | 1.4929 |
| B] | | | | | | |

As the data represents, the difference between the control group and experimental groups A and B was .661 and .926; respectively. Therefore, there was statistically no significant difference between them in the pre-test. The significance between experimental A and the control group and experimental B was .400 and. 400; respectively. Therefore, there was statistically no significant difference between them in the pre-test.

Table 14

Results of Accuracy Pre-test

| | | Subset for alpha = 0.05 |
Group	N	1
Experimental A	10	3.8500
Control Group	10	4.2000
Experimental B]	10	4.3500
Sig.		.436

As the data in Table 14 represent, the significance between experimental A control group and experimental B was .436. Therefore, there was statically no significant difference between them in the pre-test.

Post-test: Accuracy

Table 15

Descriptive for Post-test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
Control Group	10	5.0000	.78174	.24721	4.4408	5.5592	4.00	6.50
Experimental A	10	5.1000	.77460	.24495	4.5459	5.6541	4.00	6.00
Experimental B]	10	6.5000	.81650	.25820	5.9159	7.0841	5.00	8.00
Total	30	5.5333	1.03335	.18866	5.1475	5.9192	4.00	8.00

As the data represented in Table 15, the mean of the control group, experimental A, and Experimental B was 5.00, 5.100, and, 6.500; respectively. As a result, there was a statistically-significant difference between them. All in all, the mean of the three groups was 5. 53.

Table 16

ANOVA for Post-test for Accuracy

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	14.067	2	7.033	11.237	.000
Groups	16.900	27	.626		
Total	30.967	29			

As the data in Table 16 represents, the significance of between and within groups was .000. As a result, there was a statistically-significant difference between and within groups inaccuracy in the post-test.

For comparing the differences between the three groups, multiple comparisons were conducted.

Table 17

Multiple Comparisons in Post-test

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Control Group	Experimental A	-.10000	.35382	.957	-.9773	.7773
	Experimental B]	-1.50000*	.35382	.001	-2.3773	-.6227
Experimental A	Control Group	.10000	.35382	.957	-.7773	.9773
	Experimental B]	-1.40000*	.35382	.001	-2.2773	-.5227
Experimental B]	Control Group	1.50000*	.35382	.001	.6227	2.3773
	Experimental A	1.40000*	.35382	.001	.5227	2.2773

As can be seen in Table 17, the difference between the control group and the experimental A was .975. As a result, there was statistically no significant difference between them. The data show a significant difference between the control group and experimental B. As a result, there were

statistically significant differences between them. There was a statically-significant difference between experimental A and experimental B (.0001).

Speaking Ability

Like fluency and accuracy, the speaking ability of the subject was tested in pre-test and post-test. The results are shown in the following Tables.

Table 18

Descriptives: Pre-test: Speaking

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Control Group	10	3.9500	.45338	.14337	3.6257	4.2743	3.50	4.50
Experimental A	10	4.0500	.61010	.19293	3.6136	4.4864	3.00	4.75
Experimental B]	10	4.1500	.61464	.19437	3.7103	4.5897	3.00	4.75
Total	30	4.0500	.55086	.10057	3.8443	4.2557	3.00	4.75

As can be seen in Table18, the mean of the control group, experimental groups A and B was 3.95, 4.04, and 4.15; respectively. The total mean was 4.05. For ensuring that there was homogeneity between variances, a test of homogeneity was conducted. The results are shown in Table19.

Table 19

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.332	2	27	.721

As can be seen in Table 19, the significance was .721. As a result, there was an accepTable index of homogeneity between variances in the pre-test.

Table 20

ANOVA for Speaking in the Pre-test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.200	2	.100	.314	.733
Groups	8.600	27	.319		
Total	8.800	29			

As can be seen in Table 20, the level of significance is .733, As a result, there was statistically no significant difference between and within groups in the pre-test. For understanding the differences between groups in the pre-test, multiple comparisons were conducted.

Table 21

Multiple Comparisons Pre-test

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Control Group	Experimental A	-.10000	.25240	.917	-.7258	.5258
	Experimental B]	-.20000	.25240	.711	-.8258	.4258
Experimental A	Control Group	.10000	.25240	.917	-.5258	.7258
	Experimental B]	-.10000	.25240	.917	-.7258	.5258
Experimental B]	Control Group	.20000	.25240	.711	-.4258	.8258
	Experimental A	.10000	.25240	.917	-.5258	.7258

As can be seen, the difference between the control group and the experimental was .917. Therefore, there was no statistically significant difference between them. There was statistically

no significant difference between control A and experimental B (.711). There were statistically no significant differences between experimental A and B (.917).

Table 22

Statistics for Speaking: Pre-test

		Subset for alpha = 0.05
Group	N	1
Control Group	10	3.9500
Experimental A	10	4.0500
Experimental B]	10	4.1500
Sig.		.711

Means for groups in homogeneous subsets are displayed.

As can be seen in Table 22, the overall significance in speaking was .711. Therefore, there was statistically no significant difference between groups in the pre-test.

Post-test: Speaking

For determining the differences between pre-test and post-test, statistics were run.

Table 23

Descriptives for Post-test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
Control Group	10	4.5500	.28382	.08975	4.3470	4.7530	4.00	5.00
Experimental A	10	5.2250	.66091	.20900	4.7522	5.6978	4.00	6.00
Experimental B]	10	6.1750	.71734	.22684	5.6618	6.6882	5.00	7.50
Total	30	5.3167	.88311	.16123	4.9869	5.6464	4.00	7.50

As the data in Table 23 represents, the mean of the control group was 4.55. The mean of experimental A and experimental B was 5.22 and 6.17; respectively. The total mean was 5.31.

Table 24

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	13.329	2	6.665	19.375	.000
Groups	9.288	27	.344		
Total	22.617	29			

As the results of the ANOVA show, the significance was.000. Therefore, there were statistically significant differences between and within groups.

For understanding the differences between groups, multiple comparisons were conducted.

Table 25

Multiple Comparisons

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Control Group	Experimental A	-.67500*	.26229	.041	-1.3253	-.0247
	Experimental B]	-1.62500*	.26229	.000	-2.2753	-.9747
Experimental A	Control Group	.67500*	.26229	.041	.0247	1.3253
	Experimental B]	-.95000*	.26229	.003	-1.6003	-.2997
Experimental B]	Control Group	1.62500*	.26229	.000	.9747	2.2753
	Experimental A	.95000*	.26229	.003	.2997	1.6003

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

As can be seen in Table 25, the level of significance between the control group and the experimental A is .041. Therefore, there was a statistically significant difference between them in the post-test. Also, there was a statistically significant between the control group and experimental

B (.000). More, the significance between experimental groups A and B was .0003. As a result, there was a statistically significant difference between them.

Results and Discussion

Response to the First Research Question

The first research question of this study was to investigate the possible impact of online-learning activities on Iranian IELTS candidates' speaking fluency. As the data in Tables 3, 4 and 5 represent, before the treatment, there were statistically no significant differences between the three groups in terms of fluency (.245). In addition, the result of the total significance of the three groups in Table 6 represents that the significance is not statistically significant in the three groups (.257). However, after the treatment, the three groups differed in terms of fluency. In other words, as the data in Tables 8 and 8 demonstrate, the level of significance in the three groups is .001. In this regard, it can be concluded that the fluency of the participants in the experimental groups differed statistically-significant from that of the control group. In other words, the null hypothesis of this research is rejected.

Response to the Second Research Question

The second research question of this research was to explore the impacts of online-learning activities on Iranian IELTS candidates' speaking accuracy. As the data in Tables 10, 11 and 12 reveals, before the treatment, there were statistically no significant differences between the three groups in terms of accuracy (.451). In addition, as the data of multiple comparisons in Table 14 reveals, there were statistically no significant differences between the groups as the total significance was .436. However, after the treatment, there found differences between the control group and two other experimental groups as shown in Tables 15, 16 and 17 (.000). In addition, as the data in Table 17 multiple comparisons represent, the three groups differ in terms of accuracy. In this regard, there was a positive relationship between teaching online courses and enhancing the accuracy of IELTS candidates. As a consequence, the null hypothesis of the second research question is rejected.

Response to the Third Research Question

The third research question of this research was to investigate the possible impacts of online-learning activities on Iranian IELTS candidates. As the data in Tables 18,19 and 20 represent, before the treatment, there was statistically no significant difference between the three groups in terms of accuracy (.733). However, as the data in Tables 22, 23 and 24 represent, the participants differed in terms of speaking ability (.000). In other words, the students speaking ability differed after the treatment they received as shown in Table 25, multiple comparisons. In this regard, it is said that there was a statistically significant difference between online teaching activity and the speaking ability of IELTS candidates. In this regard, the third hypothesis of this research is rejected.

Response to the Fourth Research Question

The last research question of this research was to see which model of online teaching could have a more positive impact on IELTS candidates' fluency, accuracy, and speaking ability. As the data in Tables 2, 3, and 4, there were statistically no significant differences between the two modes of teaching in the experimental group (.245). In addition, as can be seen in Table 5, the level of significance between the two groups was .396; as a result, no difference was between the two groups in terms of fluency. In terms of accuracy, the results in Tables 11, 12,13, and 14 show, that there were no statistically significant differences between the two experimental groups in terms of accuracy (.436). Moreover, as far as speaking ability is concerned, there was statistically no significant differences between the two groups in the pre-test as shown in Tables 18,19, 20, and 21 (.917).

However, once the treatment was over and the participants were given a post-test, their performance increased in terms of fluency, accuracy, and speaking. Regarding fluency, Tables 8 and 9 show that there were differences between group experimental 1 and 2 (.498); however, this difference was slight. Regarding the accuracy, the data in Table 13 shows that the level of significance was.001 As a result, there were statistically significant differences between experimental group A and experimental group B in terms of accuracy. Considering the speaking ability, as the data in Table 25 represents, the level of significance between experimental group A and experimental group B is .003 which means that the two groups differ statistically.

All in all, it can be concluded that the treatments had positive impacts on Iranian IELTS candidates in terms of fluency, accuracy, and speaking. However, teaching these skills with teachers' interference and without teachers' interference, the former has more positive impacts than the latter.

Conclusions

In light of the results of this research, some conclusions can be drawn. It was found out that teaching with technology and by the virtue of Information Technology (IT) can facilitate the learning process of second language learners. In addition, it was concluded that as far as speaking ability is concerned, designing online classes with the active role of the teacher can facilitate the students' fluency, accuracy, and speaking ability.

This research has some implications. The first implications can be for English teachers in general and IELTS instructors in particular. They can use the results of this research for designing online courses in which they will supervise the students and the flow of the materials. The second group of beneficiaries is students taking such international tests as IELTS and TOEFL. They can use the results of this research to learn how to benefit from online preparatory courses as they are at home. The last, but not least group who can use the results of this research are curriculum designers. They can use the results of this research for gaining perceptions and insights on designing specialized materials for online learning and teaching.

References

Allen, I. E., & Seaman, J. (2007). Online nation: Five years of growth in online learning. Needham, MA: Sloan Consortium.

Baron, N.S. (2008). Always On. Language in an Online and Mobile World. Oxford: Oxford University Press.

Chua, Y. P., & Don, Z. M. (2013). Effects of computer-based educational achievement test on test performance and test takers' motivation. *Computers in Human Behavior*, *29*(5), 1889-1895. doi: 10.1016/j.chb.2013.03.008.

Dalgarno, B. & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual? environments? *British Journal of Educational Technology 41*(1), 10-32.

Georgiev, T., Georgieva, E., & Smrikarov, A. (2004). M-Learning - a New Stage of S-Learning. *International Conference on Computer Systems and Technologies - CompSysTech*, 1-5.

Heidari Tajan, M. (2016). Integrate Listening Activities on EFL learners' speaking Fluency. *International Journal of English Language Teaching*, *4*(7), 6-33.

Khayef, E., & Amini, D. (2015). The effect of strategic and online planning on the accuracy of L1 learners' written production. *International Journal of Applied Linguistics and English Literature*, *4*(6), 148-153. doi: 10.7575/aiac.ijalel. v.4n.6p.148.

Kruk, M. (2014). The use of internet resources and browser-based virtual worlds in teaching grammar. *Teaching English with Technology*, *14*(2), 52-67.

Omari, S. E. (Ed.). (2015). *The Effect of Computer-assisted Language Learning on Improving Arabic as a Foreign Language (AFL) in Higher Education in the United States*. (Procedia--social and behavioral sciences.

Rahnavard, F., & Mashhadi Heidar, D. (2017). The impact of computer-assisted language learning (CALL) /Web-based instruction on improving EFL learners' pronunciation ability. *International Journal of Research in Education*, 49-57.

Richards, J. C., & Renandya, W. A. (Eds.). (2002). *Methodology in language teaching: An anthology of current practice.* Cambridge: Cambridge University Press.

Sadiku, L. M. (2015). The Importance of Four Skills Reading, Speaking, Writing, Listening in a Lesson Hour. *European Journal of Language and Literature*, *1*(1), 29. doi:10.26417/ejls. v1i1.p29-31.

Wang, Y. H., & Liao, H. C. (2017). Learning performance enhancement using computer-assisted language learning by collaborative learning groups. *Symmetry*, *9*(141), 1-16.