Enhancing Speaking Fluency and Self-Perceived Communication Confidence (SPCC); a Comprehensive Approach to Integrating AI Solutions in EFL Online Flipped Classrooms

Authors:
Maryam Mansourpour[1,*]
[1] *Ph.D. student of Applied Linguistics, Faculty of Foreign Languages and Literature, The University of Tehran, Researcher at Institute for Trade and Studies Research*
Corresponding Author:
Maryam Mansourpour
Ph.D. student of Applied Linguistics, Faculty of Foreign Languages and Literature, The University of Tehran, Researcher at Institute for Trade and Studies Research

ABSTRACT:
Among various technological tools, artificial intelligence (AI) has emerged as a strong drive in the area of the education, especially in the realm of oral communication. This study investigated the effect of AI-based instruction in online flipped classrooms on Iranian EFL learners' speaking fluency and SPCC. To this end, 44 out of 113 Iranian EFL learners from three language institutes in Iran during 2024 academic year were selected based on a placement and divided into two groups as one experimental group and one control group. The data were obtained via TOEFL placement test and Self-Perceived Communication Competence (SPCC) Scale developed and designed by McCroskey and McCroskey (1988). The data were analyzed based on the descriptive statistics and factorial ANCOVA. The results revealed that that AI-based instruction in online flipped classrooms has a positively significant effect on Iranian EFL learners' speaking fluency and SPCC. Finally, the theoretical and pedagogical implications of the study are provided.

Keywords: Artificial Intelligence; EFL Learners; Online Flipped Classrooms; Self-Perceived Communication Confidence (SPCC); Speaking Fluency

1.INTRODUCTION:

Speaking appears instinctively the foremost imperative skill for communication. Individuals who know a language are considered as speakers of that language, as in the event that speaking include all other sorts of aptitudes, and numerous, on the off chance that not most outside language learners are fundamentally inquisitive about learning to talk (Ur, 2006). Speaking is an intelligently handle of developing meaning that includes creating, accepting and handling data (Burns & Joyce, 1997). Its frame and meaning are subordinate on the settings in which it happens, counting the members themselves, their collective encounters, the physical environment, and the purposes for speaking. It requires that learners not as it were know how to create particular focuses of language such as linguistic use, elocution, or lexicon (phonetic competence), but too they get it when, why and in what ways to create language (sociolinguistic competence) (Cunningham, 1999). Wilson (1997) claimed that children who can interpret their contemplations and concepts into words are more likely to succeed in school. Students who do not create great tuning in and speaking ability will have life-long consequences because of their shortage.

He pointed out that speaking aptitudes do not have to be be instructed as an isolated subject. These abilities can effectively be coordinates into other subject matter. Typically since, students learn Speaking, clarify considerations by Speaking, comprehend way better with dialog of reading, write way better after speaking amid composing conferences, create certainty by speaking before peers, and provide a window to their claim considering through their conversation. Speaking is the capacity and ready to consider it as a beneficial ability to utilize skill in arrange to interact with others. Agreeing to Cole et al. (2007), "speaking is a fundamental collaborative and intelligently prepare" (p. 12). It is a trade. We may wrap up each other's comments, hinder, oppose this idea with or expand what is said. Support (2000) pointed out that "an expertise by which they [individuals] are judged whereas to begin with impressions are being formed″ (p. 261). Speaking has need since of its critical part within the prepare of educating and learning a outside skill. As Ur (2006, p. 54) declares,
of all the four abilities [tuning in, speaking, perusing and composing], speaking appears instinctively the foremost imperative: individuals who know a skill are

alluded to as speakers of that skill, as on the off chance that speaking included all other sorts of knowing.

Classroom activities are vital since they can offer assistance the method of instructing and learning to make strides. As Scrivener (2005, p. 41) clarified, "a key educating expertise is to effectively get ready, set up and run a single classroom movement or task". Moreover, Thornbury (2005, p. 90) said "classroom speaking exercises that include negligible help, where learners can take dangers and boost their certainty, give an imperative dispatch cushion for consequent real-world skill use". So, by the use of classroom speaking exercises the students are able to hone genuine world skill by taking dangers and making strides their certainty. Now numerous ESL instructors accept that students learn to talk within the moment and remote skill by "connection". Communicative skill educating and learning offer assistance this point. Communicative skill educating is based on real-life circumstances that needs communication. By utilizing this strategy in ESL classes, students will have the chance of communicating with each other within the skill. Therefore, ESL instructors ought to make an environment where students have real-life communication, true exercises, and significant errands that advance verbal skill. This happens when students take an interest in aptitudes to do the task.

Luoma (2004) pointed out that from a testing point of view, speaking is extraordinary since of its intelligently nature. It is regularly tried in live interaction. In this way, speaking can be surveyed through interaction. In addition, Thornbury (2005) expressed that appraisal can happen at the starting and at the conclusion of the foremost dialect courses, or indeed amid the course itself. At situation, appraisal of speaking expertise can be done by utilizing meet which includes different verbal assignments; this arrangement test gives the premise for evaluating speaking expertise whether it points at testing advance amid the course or accomplishment at the conclusion of that course. Not at all like the test of speaking, a composed test of language structure is generally simple and time proficient. However, on the other hand, test of speaking is troublesome and it requires a long length of time because the educator has to meet each understudy exclusively; in truth, this may cause disturbance. As Brown (2004, p. 37) illustrated, "an instructor appraisal will be based on fifty or a hundred hours' introduction to the learners' dialect, in an assortment of exercises and circumstances". It implies that instructors ought to watch regularly the students' utilize of dialect in different circumstances and exercises. Griffin (2009), illustrated evaluation as: A prepare of skilling, deciphering and utilizing data around learning. This handle of skilling can take numerous shapes, from tests to exhibitions or work tests. The more often than not has a few shape of estimation or coding and their utilize leads to choices approximately instructing and learning.

One important aspect of speaking skill is fluency. The term fluency is generally used in language education and fluent is regularly appeared in language assessment, which seems the meaning of fluency is easily catchable. However, the definition of fluency is various. According to Hartmann and Stork (2021), fluent addresses a speaker who is able to utilize the correct structures of a language at normal speed, meaning speaking naturally with concentration on the content delivery, rather than focusing on the form or structure of a language.

The learners' speaking fluency is entangled with their self-perceived communication confidence (SPCC). As the name suggests, SPCC refers to learners' perceptions on the extent of their "adequate ability to pass along or give information" (McCroskey & McCroskey, 1988, p. 109). Taken in the L2 context specifically, SPCC means learners' views on how well they have the capability to communicate in L2. Highlighting the importance of SPCC in L2 communication, Clement et al. (2003) stated that even though learners' "actual competence might influence communication, it is the perception of competence that will ultimately determine the choice of whether to communicate" (p. 192). In other words, at some point, learners' SPCC plays a role in determining whether they make actual communication regardless of their real competence. One way to enhance the speaking fluency and SPCC is utilizing technology and online contexts.

As the information age has progressed, technology has permeated almost every part of our lives, including the realm of educational practice, by way of novel and smart gadgets, wireless broad-band technologies, and novel application services. An innovative, learner-centered, and individualized approach to education, called Computer-Assisted Language Learning (CALL), has emerged as a result of these fast developments in the field of education. Learning is now more genuine, context-aware, and pervasive than ever before because to the rise of new technology, an explosion of multimedia software, and a proliferation of mobile apps. In light of this, it is now possible for knowledge to be constructed in light of specific experience and practice, allowing for a wide range of personalization in the learning process. Skills, hobbies, and personal tastes. In Beatty's (2013) words, CALL is any advancement in which a learner utilizes a computer and, as a consequence, improves his or her language. Beatty (2013) saw CALL as a nebulous, unorganized field because of the rapid evolution of computers and technology, which can be seen in the shift from basic CD-ROMs to virtual reality in the field of computer science. Yaşar (2018) claims that the theory that new technologies may help children learn a language has its roots in the 1960s and 1970s, when the first personal computers were available to the public. Since then, the development of CALL (computer-assisted language learning) has gone through various transitional stages that mirror the trajectory of new media and paradigm transformations (Bax & Warschauer, 2017).

The Internet makes access to the authentic context more possible for language learners as they can easily get connected with the native speakers of the target language through the ever-advancing technologies like Facebook, Instagram, WhatsApp, Google Hangout, etc. Similarly, a research conducted by Elega et al. (2017) shows that the tendency to use mobile apps such as Google Translate, iVoice Translator Pro, and iTranslate Voice mobile among non-native students is more than other technological adaptive strategies. His finding is in consistence with Godwin-Jones (2011) statement calling "Conversation Mode" by Google Translate "an interesting experimental feature" as he maintains that the possibility of translating an utterance into the target language and reading the translated message aloud can make the communication in a target language easier. In their overview of technological materials and resources, Elega et al., (2017) argue that the role of mobile phones, among other wireless devices such as laptop, computers, etc., is more significant in opening up new language learning opportunities due to the practicality of using iPhone, Android devices, and Windows Phones almost everywhere. This has also been explored in prior studies by Aamri and Suleiman (2011) highlighting major features of mobile phones that facilitate the process of learning through helping language learners understand and develop their innate learning abilities, broadening individual and group learning experiences, as well as supporting learners to recognize lack of knowledge and skills in specific linguistic areas.

So far, pedagogical characteristics of technology such as facilitating the process of learning, developing learners' abilities to study independently, promoting visual, verbal and kinesthetic learning, as well as helping learners acquire problem-solving and critical thinking skills have been stressed by researchers providing an explanation for the prominence of CALL all over the world (Turnbull & Lawrence, 2002). One such approach that has gained considerable attention is the integration of Artificial Intelligence (AI) technologies into English language instruction. AI, characterized by its ability to simulate human intelligence, has made significant advancements in recent years, permeating various domains of society. In the field of education, AI is increasingly being explored as a promising tool to support and enhance language learning, particularly in the development of learners' communication skills. AI offers the potential for personalized, interactive, and adaptive learning experiences that cater to individual learners' needs and preferences.

2. Review of Literature

In a recent study, Zhang et al. (2024) examined the influence of an AI-speaking assistant, Lora, on Chinese EFL students' foreign language enjoyment (FLE), foreign language anxiety (FLA), and willingness to communicate (WTC) in English. To this end, EFL university students participated in this study and were placed into the experimental and control groups, respectively, who filled out a pre- and post-treatment survey before and after a six-week treatment. Results unveiled significant enhancements in WTC and FLE, accompanied by a noteworthy reduction in FLA among the AI-engaged EG. These findings underscored the efficacy of AI-speaking assistants in amplifying EFL students' FLE and WTC while mitigating FLA.

In another study, Tai and Chen (2023) investigated the effects of out-of-class interactions with intelligent personal assistants (IPAs) versus human interlocutors on EFL learners' WTC in English. To this end, 100 college students participated in interactive out-of-class activities, which were held in 10-min sessions twice a week for 10 weeks. The results showed that there were significant differences in WTC. The IPA group had greater WTC than did the EL1 or EL2 group. The intergroup differences were attributable to the interaction of contextual (interlocutor and time), individual (L2 self-confidence, anxiety, and proficiency), and sociopolitical (Taiwan's K-12 education) factors. The participants in the IPA group emphasized that the mobility, convenience, interactivity, multi-functionality, and familiarity of IPAs on smartphones enabled them to practice speaking English anywhere, anytime, and at their own pace. IPAs, as supportive and patient learning partners, enhanced the learners' engagement, confidence, and thus WTC.

Qiao and Zhao (2023) inquired the effectiveness of AI-based instruction in improving L2 speaking skills and speaking self-regulation in a natural setting. The research was conducted with Chinese EFL students, randomly assigned to either an experimental group receiving AI-based instruction or a control group receiving traditional instruction. The AI-based instruction leveraged the Duolingo application, incorporating natural language processing technology, interactive exercises, personalized feedback, and speech recognition technology. Pre- and post-tests were conducted to assess L2 speaking skills and self-regulation abilities. The results of the study showed that the experimental group, which received AI-based instruction, exhibited significantly greater improvement in L2 speaking skills compared to the control group. Moreover, participants in the experimental group reported higher levels of self-regulation.

In addition, Yan (2023) adopted a qualitative approach to explore the impact of ChatGPT in L2 writing classrooms. The findings indicated that ChatGPT had the potential to enrich L2 writing pedagogy by introducing developmental features in learning activities and facilitating efficient writing. This pioneering endeavor underscored the need for further research into ChatGPT's applicability in L2 learning and the formulation of corresponding pedagogical adaptations. Further, Duolingo application, Li and Bonk (2023) conducted a study on online language learners using Duolingo outside of formal classrooms. They found that learners employed various resources

and self-monitored their learning process, relying on Duolingo's features. Intrinsic motivations, such as cultural interest and travel, drove learners more than certificates or grades. Kessler (2023) addressed limitations in mobile-assisted language learning (MALL) applications by integrating reflective e-journal activities with Duolingo. The study, grounded in metacognition theory, revealed that the journals enhanced students' metacognitive awareness in various domains, with learners finding the activity beneficial and enjoyable. In addition, Shortt et al. (2023) reviewed Duolingo's gamified MALL application, highlighting its popularity and gamification elements. They found that research focused on app design, quantitative methods, and non-probability sampling, emphasizing tool creation over learning process and outcomes. The study identified preferences for performance-based research questions, English language, and the United States as the main research context, revealing research gaps with implications for MALL and gamification practitioners and researchers.

Jia et al. (2022) presented an AI system designed to enable authentic and ubiquitous language acquisition. Their study, involving 20 participants, employed a combination of research methods to assess the system's usability and validity. The findings affirmed the efficacy of the AI system and yielded insights for potential enhancements. This research contributes to the integration of AI in language instruction and adheres to mobile learning principles, emphasizing the significance of authentic learning environments. Additionally, Subekti (2020) investigated college students' SPCC and their communication apprehension (CA), in English. To this end, 150 non-English major college students participated in the study in a survey using probability random sampling. Using descriptive statistics, the study found that learners had medium level of SPCC and medium level of CA. Through correlation formula, the study also found that there was moderate significant association between learners' SPCC and their CA. Through regression analysis, it was found that their SPCC could predict 23% of variance in their CA. In another research, Ferla et al. (2019) reported that the participants scoring high in SPCC were found to be more persistent, less anxious, to "process the learning material at a deeper level, and achieve better study results" (p. 519). In a similar vein, Denies et al. (2015), and Shahbaz et al. (2016) found that learners' SPCC is attributed to their courage to use L2 in class and contribute in class discussions. Shahbaz et al. (2016) further stated that as learners have a positive perception on their ability to make meaningful communication in one language, they will likely be more willing to communicate in that language. This confirmed Bandura's (1989) reiteration more than three decades ago that learners' perception of their competence influences their performance in which perception of low competence to perform a task can lead them to be anxious while performing that task and vice versa.

Based on related literature, it was revealed that no study has been yet conducted in Iranian EFL context to examine the effect of AI-based instruction in online flipped classrooms on EFL learners' speaking fluency and SPCC. Therefore, the present study aimed to fill in such a gap in literature by testing the following null hypothesis:

1. AI-based instruction in online flipped classrooms does not have any significant effect on Iranian EFL learners' speaking fluency.

2. AI-based instruction in online flipped classrooms does not have any significant effect on Iranian EFL learners' SPCC.

3. Method
3.1 Participants
The present study was conducted with 44 out of 113 Iranian EFL learners from three language institutes during 2024 academic year. The participants were randomly drawn from learners, who took the language proficiency test, assigned to the intermediate learners, and then randomly divided into two groups as one experimental group and one control group, ranging from 18 to 24 years old.

3.2 Research Instruments
The following instruments were utilized in the present study.

3.3 Placement Test
In order to check the level of general language proficiency of the participants at the beginning of the study and find out a homogenous sample, a sample TOEFL OPT test without its writing section will be utilized. The items of the OPT test are taken from 'Longman Complete Course for the TOEFL Test' by Philips (2022). The OPT consists of three parts: listening comprehension, collocation comprehension and grammatical structures. The test has 100 items for which the highest score is 100. Based on the standard of the test itself, and since the items of the test are time-consuming, the allotted time is 100 minutes.

3.4 Speaking Fluency Test (Pretest & Posttest)
In order to examine the participants' speaking fluency before and after the treatment, speaking sections of the OPT was used. The test consisted of 15 questions asked by two professional raters. In addition, the inter-rater reliability of the test was calculated to determine the agreement between the two raters in scoring participants' speaking fluency. Two raters, the researcher and a language expert with eight years teaching experience used OPT speaking rubrics to score the participants. The agreement of the two raters in scoring the examinees' performances was also calculated through Pearson product moment coefficient of correlation. In the current study, speaking fluency was measured in line with Ellis (1990, as cited in Ellis & Barkhuizen, 2005) in terms of the number of syllables produced per min on a task. To do so, the number of syllables produced were counted and divided by the number of min took to

produce the spoken output. In addition, the number of pauses the participants made while speaking were used to measure their fluency.

3.5 Self-Perceived Communication Competence Scale

The *SPCC* scale developed and designed by McCroskey and McCroskey (1988) was used to measure the participants SPCC. The scale is a 12-item questionnaire that measures self-perceived competence across a variety of speaking situations. Speaking situations include a variety of contexts (public speaking, meetings, groups, and dyads) and communication partners (strangers, acquaintances, and friends). Respondents read each speaking situation (e.g., "Present a talk to a group of strangers" or "Talk with an acquaintance") and write an estimate of their competence from 0 (completely incompetent) to 100 (competent). Individual items with either a shared context or communication partner are averaged to compute subscores (e.g., all items with a friend are averaged to calculate a "Friend" subscore). A total score is derived by taking the average of the Stranger, Acquaintance, and Friend subscores. Total scores above 87 are indicative of high self-perceived communication competence and scores below 59 are indicative of low self-perceived communication competence. The original study reported a high reliability (.92) for the total score and adequate to good reliability for all of the subscores except for the Dyad subscore. The *SPCC* scale was the only measure utilized to investigate self-perceived communication competence. While self-perceived communication competence has been investigated along with other communication attributes in previous literature (e.g., Rubin et al. 1997), there is precedence for single-measurement investigations, especially with the *SPCC* scale. Single-measurement investigations utilizing the *SPCC* scale have included adolescents who stutter (Blood & Blood, 2004), adults learning foreign languages (e.g. Rasekh et al., 2012), and with adults in multiple countries for cross-cultural comparisons of the scale (e.g., Dilbeck et al., 2009).

3.6 Data Collection Procedure

To conduct the study, the participants were selected from 113 Iranian EFL learners at the intermediate level of language proficiency. The initial 113 learners were given a proficiency test, and based on the OPT results 44 intermediate level learners were selected, and divided into two groups each consisting of 22 learners. Prior to starting the experiment, the learners in the two groups were given the pretest and SPCC. The experimental group was taught in online flipped class via Duolingo application. The Duolingo AI chatbot provided learners with prompts and questions in English. The learners were supposed to respond to the prompts, and the chatbot gave real-time feedback on different aspects of their spoken language, including pronunciation, fluency, vocabulary, and accuracy. This feedback was generated through machine learning algorithms that analyzed EFL learners' performance data, enabling the chatbot to offer personalized feedback tailored to each EFL learner's specific needs. The treatment also included group activities and discussions, allowing EFL learners to practice their speaking skills in a more naturalistic setting. Furthermore, learners had the opportunity to track their progress within the Duolingo application, which provided them with motivation and feedback on their overall language learning journey. In contrast, the control group was taught via traditional method for teaching speaking fluency in face-to face classes without any digital and online facility. The treatment lasted eight sessions. Finally, the speaking fluency posttest and SPCC were administered.

3.7 Data Analysis

To answer the first and second questions empirically, a factorial ANCOVA (Two-Way ANCOVA) was conducted. This analysis had to be done to remove the effect of the initial difference between the two groups regarding their speaking fluency and SPCC mean scores (covariate) from their speaking fluency and SPCC posttest means. Primarily, though, the assumptions had to be checked. The assumption of linearity of relationship between the covariate and the dependent variables for each group was verified graphically as demonstrated below:

Figure 1. *Scatterplot of Relationship between Covariate and Speaking Fluency and SPCC for the Two Groups*
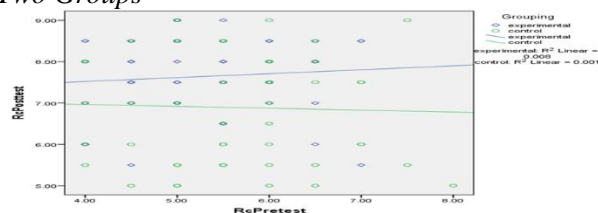
As Figure 1 demonstrates, there are two straight lines indicating linear relationships between the covariate and the dependent variables for the two groups, and as here appears to be no curvilinear relationship, the assumption is met. In addition, assumption of homogeneity of regression slopes concerns the relationship between the covariate and the dependent variables for each of the two groups, which was checked statistically. The following table shows that the assumption was not violated as the sig value corresponding to grouping*pretest turned out to be less than .58 which is less than .05 which indicates that there was no significant difference between the relationships across the groups (Table 1).

Table 1. *Tests of Between-Subjects Effects*

Dependent Variable: Posttest						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	14.504 a	2	4.835	3.592	.016	.101
Intercept	141.86	1	141.86	105	.0	.523

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
	8		8	.39 4	00	
Grouping	3.705E-005	1	3.705E-005	.00 0	.9 96	.000
Pretest	.051	1	.051	.03 8	.8 46	.000
Grouping * Pretest	.417	1	.417	.31 0	.5 79	.003
Error	129.224	42	1.346			
Total	5421.750	44				
Corrected Total	143.728	43				
a. R Squared = .101 (Adjusted R Squared = .073)						

Normality of distribution of all the sets of scores were also checked statistically (Table 2).

Table 2. *Descriptive Statistics of the Posttest Scores and the Covariate*

	N	Min.	Max.	Mean	SD	Skewness		Ratios
Ex. Speaking Post	22	5.50	8.50	7.5200	.88365	-.656	.46 4	-1.41
C. Speaking Post	22	5.00	9.00	6.9200	1.33604	-.013	.46 4	-.028
Valid N (listwise)	44							

As Table 2 shows, all sets of scores were normally distributed as the corresponding ratios were less than 1.96. Hence, the assumption is met. Then, to test the first research hypothesis the factorial ANCOVA was run.

Table 3 exhibits that the covariate did not have any significant effect on the participants' speaking fluency (F=.062, p=.805>.05). As revealed in this table also, the interaction between grouping and instruction turned out to be non-significant (F=.431, p=.513>.025), which implies that there was no significant difference between the effect of AI-based instruction in online flipped class on learners' speaking fluency. It is also shown that instruction alone was not significantly effective (F=.183, p=.67>.025), whereas there was a significant difference between the experimental and control groups collectively (F=10.26, p=.002<.025). Therefore, the corresponding null hypothesis is rejected with the effect size of .097, which means that the AI-based instruction in online flipped class could explain 9.7 percent of the variation in the participants' speaking fluency. By virtue of the total scores reported in Table 4, the experimental group outperformed the control group. Further, to examine the second null hypothesis, another factorial ANCOVA was conducted. Primarily, the normality condition was checked the result of which is shown hereunder.

Table 3. *Tests of Between-Subjects Effects*

Dependent Variables: Speaking Fluency

Source	Type III	df	Mean Squar	F	Si g.	Partial Eta

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	14.911a	2	3.728	2.749	.03 3	.104
Intercept	134.990	1	134.990	99.553	.00 0	.512
Pretest	.083	1	.083	.062	.80 5	.001
Instruction	.248	1	.248	.183	.67 0	.002
Grouping	13.910	1	13.910	10.258	.00 2	.097
Instruction * Grouping	.585	1	.585	.431	.51 3	.005
Error	128.817	42	1.356			
Total	5421.750	44				
Corrected Total	143.728	99				

a. R Squared = .104 (Adjusted R Squared = .066)

Table 4. *Descriptive Statistics of the Posttest Scores and the Covariate*

	N	Min.	Max.	Mean	SD	Skewness	Ratios
Ex. SPCC Posttest	22	46.50	81.00	65.036	10.4916	-.160	.464	-.344
C. SPCC Posttest	22	58.00	81.00	72.680	7.14417	-.832	.464	-1.79
Valid N (listwise)	22						

Table 4 reveals that all distributions of scores enjoyed normalcy as all the skeweness ratios fell below 1.96. The following table is also produced to check the homogeneity of variances condition.

As depicted in Table 5, there was no significant difference among variances of the groups (F=1.139, p=.337>.05). Thus, the assumption is met. Then, the descriptive statistics of each group in SPCC posttest are calculated and reported.

The main result of the factorial ANCOVA, as reported there, the interaction effect was not significant (F=.017, p=.897>.05) implying that the treatment had no significantly different effect on learners' SPCC. Additionally reported in Table 7 is the main effect of instruction which turned out to be non-significant as well (F= 1.356, p=.247>.05), and the main effect of grouping which reveals that there was a significant difference between the mean scores of the experimental group and the control group collectively (F=9.83, p=.002<.05).

By virtue of the mean scores presented in Table 6, (Experimental= 72.68 vs. Control= 66.87), and regarding the effect of instruction presented in Table 7, it is concluded that there was a significant difference between the experimental and control groups collectively (F=9.83, p=.002<.025). Therefore, the corresponding null hypothesis is rejected with the effect size of .093, confirming that AI-based

instruction in online flipped class could increase the learners' SPCC, and that it could explain 9.3 percent of the increase of SPCC of the learners.

Table 5. *Levene's Test of Equality of Error Variancesa*

Dependent Variable: SPCC Post

F	df1	df2	Sig.
1.139	2	43	.337

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

Table 6. *Descriptive Statistics of SPCC Posttest Means*

Dependent Variable: SPCC

Instruction	Grouping Post	Mean	Std. Deviation	N
	Control	66.876	8.82899	22
	Experimental	72.680	7.14417	22
	Total	69.778	8.47186	22

Table 7. *Tests of Between-Subjects Effects*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	885.182ᵃ	2	295.061	3.734	.014	.104
Intercept	472560.005	1	472560.00	5980.44	.000	.984
Instruction	107.122	1	107.122	1.356	.247	.014
Grouping	776.737	1	776.737	9.830	.002	.093
Instruction * Grouping	1.323	1	1.323	.017	.897	.000
Error	7585.683	42	79.018			
Total	481030.870	44				
Corrected Total	8470.865	43				

a. R Squared = .104 (Adjusted R Squared = .077)

4. Discussion and Conclusion

As mentioned, the current study investigated the effect of AI-based instruction in online flipped classrooms on Iranian EFL learners' speaking fluency and SPCC. The findings revealed that AI-based instruction in online flipped classrooms has a positively significant effect on Iranian EFL learners' speaking fluency and SPCC. These findings are supported by the results of previous studies (Bahrani & Sim, 2012; Junaidi, 2020; Kang, 2022; Yan, 2023). All of these studies showed the favorable impacts of AI on learners' speaking competences by increasing the learners' engagement

with AI-based instruction in online flipped contexts, which are more stimulating and interactive. AI-based instruction motivated the learners to engage in communication in a new environment, and their self-perceived communication confidence. AI-based instruction in online flipped contexts gives the EFL learners personalized and adaptive learning experiences, improving the analysis of their performance and enabling the identification of areas for increasing, as well as the provision of tailored feedback and practice tools (Yan, 2023). According to Yang (2023), this individualized approach lets EFL learners to meet their specific language needs and improve at their learning speed. Further, as stated by Jia et al. (2022), AI-based instruction in online flipped contexts provides EFL learners with extensive language input through different tools such as interactive simulations, virtual environments, and AI-powered online applications and chatbots. Engaging in real-life contexts are rich in speaking tasks within these platforms give learners to authentic language usage, which plays a pivotal role in their effective oral communication. Moreover, AI-based instruction in online flipped contexts ensures learners get continuous and immediate feedback on their speaking fluency and SPCC.

By leveraging AI-based instruction in online flipped contexts, ESL/EFL learners' pronunciation, grammar, and discourse features can be analyzed, and provide feedback (Divekar et al., 2022), allowing learners to promptly identify and rectify errors, reinforcing correct language production, and fostering their self-monitoring and self-correction skills (Loncar et al., 2023). The results of this study demonstrated that AI-based instruction in online flipped contexts fostered the development of EFL learners' interactions in the experimental group. These findings can be justified by Vygotsky's (1984) social constructivism theory in which the role of AI as a facilitator in the growth of students' self-regulation is highlighted. Consistent with Vygotsky's propositions, learners are first engaged in communicative speaking tasks with AI-based instruction in online flipped contexts, which probably helped them in regulating their own speaking skills and interactions. Then, the learners gradually move from other-regulation to self-regulation, demonstrating independent speaking fluency. Notably, the learners who exhibited self-control are able to complete their speaking tasks without relying on AI or other peers, indicating higher levels of self-regulation among the AI learners.

In addition, the results lend support for Bandura's (1989) social cognitive theory, in which learning takes place via modeling, imitation, and observation, and of others' behaviors. In the context of interactions supported by AI-based instruction in online flipped contexts, the learners had the chance to observe and engage with AI systems that demonstrate self-regulatory behaviors, such as offering adaptive feedback or guiding learners in goal-setting and planning (Zimmerman, 2002).

Regarding the the theoretical and pedagogical implications of the study, the results of this study have important implications for EFL/ESL education. Given that AI-based instruction in online flipped contexts integrate with student-centered approaches, and significantly increase EFL/ESL students' speaking fluency and SPCC, its integration is recommended in interactive EFL/ESL speaking programs. EFL/ESL educators are encouraged to utilize AI in their communicative speaking courses to improve the development of speaking skills and self-regulation among EFL students. By implementing an AI-supported classroom, EFL/ESL teachers can design engaging communicative speaking tasks involving within AI and peers. In addition, EFL/ESL learners can take advantage of an AI-infused course in flipped online classes.

Regarding the limitations of the study, the generalizability of the results may be restricted to the specific sample of Iranian EFL students, cautioning against applying these findings to the learners from various language, educational and cultural backgrounds. Secondly, the duration of the treatment was limited to eight sessions. Longer intervention periods may yield various findings, and provide a more comprehensive understanding of the effects. Additionally, using multiple measures and qualitative assessments in future studies would provide a more comprehensive understanding of the impact of AI-based instruction in online flipped contexts. It is also suggested to further investigate the long-term effects of AI-based instruction in online flipped contexts in future studies.

Conflict of interest: The authors declare that they have no conflict of interest.

REFERENCE

Aamri, E. & Suleiman, R. (2011). Effects of an intelligent tutoring system on students' learning outcomes and attitudes in the context of English as a foreign language. *Journal of Computer Assisted Learning, 34*(4), 376-393.

Bahrani, T., & Sim, T. S. (2012). Informal language learning setting: technology or social interaction? *Turkish Online Journal of Education Technology, 2*, 10–17. http://doi.org/10.7456/10203100/002

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychology. 44*, 1175–1184. http://doi.org/10.1037/0003-066X.44.9.117

Bax, E. R. & Warschauer, B. (2017). Artificial intelligence and its role in education. *Sustainability, 13*, 129-147. doi: 10.3390/su132212902

Beatty, D. (2013). Can home use of speech-enabled artificial intelligence mitigate foreign language anxiety: investigation of a concept. *World English Journal, 3*(1), 28–40. doi: 10.24093/awej/call5.3

Blood, G. W., Blood, I. M., Tellis, G., & Gabel, R. (2001). Communication apprehension and self-perceived communication competence in adolescents who stutter. *Journal of Fluency Disorder, 26*, 161–178.

Brown, H. D. (2004). *Principles of language learning and teaching* (Fourth Edi). Prentice Hall.

Burns, E. & Joyce, A. R. (1997). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*(3), 218-226.

Clement, R., Baker, S. C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effects of context, norms, and vitality. *Journal of Language and Social Psychology*, *22*, 190–209.

Cole, R. T. Bryman, A. E. Cao, Y. A. (2007). Investigating situational willingness to communicate within second language classrooms from an ecological perspective. *System*, *39*(4), 468–479. https://doi.org/10.1016/j.system.2011.10.016

Cunningham, E. (1999). Getting them speaking: Classroom social factors and foreign language anxiety. *TESOL Journal, 7*(1), 132-161.

Denies, K., Yashima, T. & Janssen, R. (2015). Classroom versus societal willingness to communicate: Investigating French as a second language in Flanders. *The Modern Language Journal, 99*(4), 718-739.

Dilbeck, K. E., McCroskey, J. C., Richmond, V. P., & McCroskey, L. L. (2010). Self-perceived communication competence in the Thai culture. *Journal of Intercultural Communication Research, 38*(1), 1–7. https://doi.org/10.1080/17475750903381598

Divekar, R. Klompas, M., & Ross, E. (2022). Life experiences of people who stutter, and the perceived impact of stuttering on quality of life: Personal accounts of South African individuals. *Journal of Fluency Disorders, 29*(4), 275–305

Elega, R. Darling, A. L., & Dannels, D. P. (2017). Preliminary investigation of daily variability of stuttering in adults. *Journal of Communication Disorders, 60*, 39–50.

Godwin-Jones, R. T. (2011). A rose by any other name is still a rose: A comparative analysis of reticence, communication apprehension, unwillingness to communicate, and shyness. *Human Communication Research, 8*(2), 99–113.

Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: focusing on the collaboration with human teachers. *Journal of Research on Technology in Education, 55*(1), 48-63.

Kang, S.-J. (2022). Dynamic emergence of situational willingness to communicate in a second language. *System, 33*(2), 277-292.

Qiao, H., & Zhao, A. (2023) Artificial intelligence-based language learning: illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology, 14*, 1255594. http://doi.org/10.3389/fpsyg.2023.1255594

Ellis, R. & Barkhuizen, G. (2005). *Analyzing learner language*. Linguistics & Education.

Ferla, S., Moore, M. R., Taylor, K. P., Surges-Tatum, D., & Hulbert-Johnson, R. (2019). The competent speaker. *Speech evaluation form* (2nd ed.).

Washington, DC: National Communication Association.

Griffin, E. (2009). Changes in communication apprehension, satisfaction, and competence in foreign dialect and stuttering clients. *Journal of Communication Disorders, 24*(5–6), 353–366

Kessler, T. E. (2023). Reconceptualizing social skills in organizations: Exploring the relationship between communication competence, job performance, and supervisory roles. *Journal of Leadership & Organizational Studies, 11*(2), 63–77

Hartmann, L. & Stork, R. (2021). Comparing the impact of online and paper-and-pencil administration of the self-determination inventory: Student report. *Journal of Special Education Technology, 35*(3), 133–144

Jia, F., Sun, D., Ma, Q., & Looi, C. K. (2022). Developing an AI-based learning system for L2 learners' authentic and ubiquitous learning in English language. *Sustainability, 14*, 15527. doi: 10.3390/su142315527

Junaidi, J. (2020). Artificial intelligence in EFL context: rising students' speaking performance with Lyra virtual assistance. *International Journal of Science and Technology. 29*, 6735–6741.

Li, Z., & Bonk, C. J. (2023). Self-directed language learning with Duolingo in an out-of-class context. *Computer Assisted Language Learning, 39*, 1–19. doi: 10.1111/jcal.12730

Loncar, M., Schams, W., & Liang, J. S. (2023). Multiple technologies, multiple sources: trends and analyses of the literature on technology-mediated feedback for L2 English writing published from 2015-2019. *Computer Assisted Language Learning. 36*, 722–784. doi: 10.1080/09588221.2021.1943452

Luoma, R. (2004). Fluctuations in willingness to communicate during a semester: A case study. *The Language Learning Journal, 49*(1), 1-12.

McCroskey & McCroskey, 1988 Examining the level of enjoyment and anxiety among Turkish EFL students. *Journal of Language and Linguistic Studies, 17*(1), 663-671.

Philips, D. W. (2022). Emotions in second language learning: Retrospect and prospect. *Modern Foreign Languages*, *47*(1), 1-13.

Rasekh, A. E., Zabihi, R., & Rezazadeh, M. (2012). An application of Weiner's attribution theory to the self-perceived communication competence of Iranian intermediate EFL learners. *Elixir International Journal of Psychology. Elixir Psychology, 47*, 8693–8697.

Rubin, E. Porter, P., & Grant, M. (1997). Towards an understanding of fluency: A microanalysis of nonnative speaker conversation. *Discourse Processes, 14*, 423-441. http://dx.doi.org/10.1080/01638539109544795

Scrivener, E. (2005). Pair taping: Increasing motivation and achievement with a fluency practice. *TESL-EJ, 5* (2), 44-59. Retrieved from http://www-writing.berkeley.edu/TESLEJ/ej18/a2.html

Shahbaz, M., Khan, M. S., Khan, R. R. M. I., & Ghulam, M. (2016). Role of self-perceived communication competence and communication apprehension for willingness to communicate in L1 and L2. *Journal of Educational and Social Research*, *6*(1), 158–166. https://doi.org/10.5901/jesr.2016.v6n1p158

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., and Akinkuolie, B. (2023). Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020. *Comput. Assist. Lang. Learn.* 36, 517–554. doi: 10.1080/09588221.2021.1933540Subekti (2020)

Support, T.Q. (2000). An intensive approach to building conversational skills. *ELT Journal, 59*(1), 14-22.

Tai, T.Y.., & Chen, H. J. (2023). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments, 31*(3), 1485-1502.

Thonrbury, S. (2005). *How to teach speaking*. Longman.

Turnbull, M. & Lawrence, G. (2002). *Computers make sense according to brain research…But what do students think*? Canadian Association of Second Language Teachers. Retrieved December 6, 2008 from http://www.caslt.org/index_en.php.

Ur, P. (2006).*A course in English language teaching.* Oxford University Press.

Wilson, E. (1997) Conceptualization of willingness to communicate in ESL. *Language, Culture and Curriculum, 16*(1), 18-38.

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: an exploratory investigation. *Educ. Inf. Technol.* 12, 1–25. doi: 10.1007/s10639-023-11742-4

Vygotsky, L. S. (1984). *Thought and language*. Cambridge: MIT Press

Yaşar, E. T. (2018). Computer-assisted EFL writing and evaluations based on artificial intelligence: a case from a college reading and writing course. *Libr. HiTech, 40*, 80–97. doi: 10.1108/LHT-05-2020-0113

Zhang, C., Meng, Y. & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. http://dx.doi.org/10.2139/ssrn.46717

Zimmerman, B. J. (2002). Becoming a self-regulated learner: an overview. *Theory of Practice. 41*, 339, 64–70. doi: 10.1207/s15430421tip4102_2