Education and Lifelong Development Research

Vol. 1, No. 1, p. 3-14, 2024

Original article

The effects of immersion in virtual reality environment on oral English learning for Chinese university students

Jian Li[1,2]*, Tian Qiu[1], Changjing Li[1], Chenhan Xu[1], Peixin Cheng[1], Yun Tang[3,4], George Kyriacou Georgiou[5]

[1]*Faculty of Psychology, Beijing Normal University, Beijing 100875, P. R. China*

[2]*Beijing Key Laboratory of Applied Experimental Psychology, Beijing 100875, P. R. China*

[3]*School of Psychology, Central China Normal University, Wuhan 430079, P. R. China*

[4]*Key Laboratory of Adolescent Cyberpsychology and Behavior, Wuhan 430079, P. R. China*

[5]*Department of Educational Psychology, University of Alberta, Alberta 510632, Canada*

Keywords:
Virtual reality
oral English learning
immersion
second language

Cited as:
Li, J., Qiu, T., Li, C., Xu, C., Cheng, P.,Tang, Y., Georgiou, G. K. (2024). The effects ofimmersion in virtual reality environmenton oral English learning for Chineseuniversity students. Education andLifelong Development Research, 1(1), 3-14.
https://doi.org/10.46690/elder.2024.01.02

Abstract:
Most Chinese adults encounter difficulty in second language acquisition due to the lack of a second language environment to practice. The aim of this study was to examine the effects of immersive virtual reality learning environments on oral English learning for Chinese university students. A total of 43 freshmen participated in this study. Participants' oral English performance, engagement, emotional and motivational status were compared between the fully immersive VR learning condition (using a head-mounted device) and the non-immersive condition (using a personal computer). Results indicated that participants in the fully immersive VR condition showed higher speaking accuracy, lower anxiety, and higher speaking-efficacy in the posttest than participants in the non-immersive VR condition. Participants in the fully immersive VR condition also showed more engagement than those in the non-immersive VR condition in all four experiment days. Theoretical and practical implications for applying VR technology in second language learning are discussed.

1. Introduction

Virtual reality (VR) is a computer-generated environment that simulates the real or imaginary world (Lee & Wong, 2008; Cakiroğlu & Gökoğlu, 2019). VR technology allows people to be immersed in a 3D virtual environment and even interact with people or objects in it. The intuitive experience that highly resembles the reality helps people learn facts and procedural skills. Therefore, VR technology has been increasingly used in education and training. More specifically, it has been applied in learning of many subjects such as history (Morgan, 2013), geometry (Song & Lee, 2002), language (Yang et al., 2010), and training of professionals such as pilots (Hawkins, 1995), doctors (Gutiérrez et al., 2007), and teachers (Barmaki & Hughes, 2015).

In this study, we are interested in the application of VR technology in second language acquisition because it might help Chinese to overcome their difficulties in oral English language learning. Therefore, the current study used a Chinese sample of university students to explore the influence of VR technology on their oral English language learning. In the following sections, we will first specify the challenges and difficulties of oral English language learning in China. Then, we will elaborate on the role of immersive VR on learning. Last, we will discuss the possible effects of immersive VR technology on oral English language learning.

Yandy Scientific Press

*Corresponding author.
E-mail address: jianli@bnu.edu.cn (J. Li); 202111061061@mail.bnu.edu.cn (T. Qiu); chauncey_li@qq.com (C. Li); xanxus707@qq.com (C. Xu); tangyun@ccnu.edu.cn (Y. Tang); 1176667066@qq.com (P. Cheng).

Received January 10, 2024; revised February 1, 2024; accepted February 16, 2024; available online February 20, 2024.

1.1 Oral English language learning in China

Second language proficiency refers to a person's ability to use a language effectively, accurately, and fluently in various contexts. It encompasses a range of skills, including speaking, listening, reading, and writing. Previous studies have construed proficiency as an index of general abilities across language processing domains (Marian et al., 2007), including literacy-oriented proficiency, grammatical proficiency, vocabulary knowledge, and discourse abilities (Bachman, 1990; Harley et al., 1990).

According to the English proficiency report released by the well-known English training company EF Education First in 2018, China ranked 47th among 88 non-native countries and areas of English and was rated as one of "low proficiency" countries (EF Education First, 2018). This result suggested that many Chinese still cannot use English effectively in social life after they have spent more than 10 years learning English from elementary school to college.

One main manifestation of low proficiency is "dumb English" (Zeng, 2012), which refers to the low level of oral English language competence. A number of Chinese students can pass English examinations in school, however, they are not able to communicate with others in English. According to previous studies, there are two explanations for this paradox. The first possible explanation relates to the method of second language teaching and learning in China. From the perspective of usage-based approach, actual language use shapes linguistic forms and is the foundation of second language learning (Tyler, 2010). Oral English competence is more likely to develop in meaningful communications (Long, 1983). Specifically, an effective English learning environment might at least have pragmatic language examples and an authentic communication situation, so that learners can learn to use contextually-appropriate language to communicate with others and solve real-life problems (Belz, 2007; Jauregi et al., 2011). However, most Chinese learners of English focus on memorization of vocabulary and grammar in the classroom and have few opportunities to speak English both inside and outside the school. Learning a second language without the authentic communication contexts could be a reason for Chinese students' low proficiency in oral English.

Second, the anxiety that accompanies learning of a different language might also affect the development of learners' oral English language competence. Some researchers using samples of Chinese college students revealed that English learners have high levels of anxiety (e.g., Liu et al., 2012) and low self-efficacy (e.g., Peng, 2006) in oral English learning. Foreign language anxiety refers to "the feelings of tension and apprehension specifically associated with second language contexts, including speaking, listening and learning" (MacIntyre & Gardner, 1994). Anxiety has been generally recognized as a debilitating factor that leads to negative learning processes and performance (Jackson, 2002; Huang & Hung, 2013). Self-efficacy refers to people's beliefs about their capabilities to produce desired outcomes (Bandura, 2012). Self-efficacy could influence people's choices, efforts, emotions and the perseverance they use to deal with difficulties and challenges (Bandura, 1997). Studies showed that individuals with higher self-efficacy in English learning are more likely to make efforts to improve their English skills and have less anxiety (Sardegna et al., 2018). These findings demonstrated that anxiety and self-efficacy in English language learning are not only indicators of emotional and motivational states, but also can be predictors of oral English performance.

Taken together, Chinese students' oral English language proficiency could be enhanced from two perspectives. One is to build a learning environment in which individuals can learn and use contextually-appropriate language to make authentic communication. The other one is to increase individuals' willingness to learn English by reducing anxiety and heightening confidence.

1.2 The role of VR in learning

A VR environment can simulate a real-life process or situation in which individuals could be immersed. Immersion characterizing VR can be defined as "a form of spatio-temporal belonging in the world that is characterized by deep involvement in the present moment" (Hansen & Mossberg, 2013, p. 212). The level of immersion in VR ranges from non-immersive VR to fully immersive VR. Non-immersive VR can be manipulated by using a personal computer and some attachments such as a mouse and keyboard. In fully immersive VR environments, individuals need to wear special devices such as head-mounted device (HMD) and sensor gloves to completely block the real world. The virtual world will occupy all of his or her visual, aural and other perceptions (Lee & Wong, 2008).

Immersive VR could bring a sense of embodiment which makes individuals feel that they are part of the virtual world, and the avatar in VR becomes part of their own body schema (Haans & IJsselsteijn, 2012; Bailey et al., 2016). Simply put, people who are immersed in VR environments have a subjective feeling of "being there" and participating in a real event or situation.

VR has been used as a teaching tool to enhance knowledge understanding and improve performance for many years, and education takes great advantages from it. For example, teachers used VR technology to simulate historical locations and scenes in the history classroom (Morgan, 2013). Primary school students who practiced fire safety behavioral skills in VR training environments could transfer their skills to real environments (Cakiroğlu & Gökoğlu, 2019). Additionally, VR has been used in teacher training and effectively enhanced self-efficacy of student teachers (Nissim & Weissblueth, 2017). Simply put, a large number of studies have showed that VR technology contributes a lot to teaching and learning, and the sense of immersion or "being there" in VR is associated with positive learning outcomes (Gutiérrez et al., 2007; Fassbender et al., 2012).

The level of immersiveness in VR could be influenced by the hardware and software of the VR instrument. Individuals in fully immersive environments could have stronger immersive sense than those in non-immersive or partially immersive environments (Gutiérrez et al., 2007). However, the level of immersiveness in VR could also be influenced by individuals' personality traits. Researchers found that people's immersive experience in VR is correlated with personal traits such as empathy in story-telling contexts (Shin, 2018), and openness, neuroticism, and extraversion (Weibel et al., 2010).

1.3 Oral English language learning with VR technology

VR technology can simulate an authentic communication environment in which English language learners can learn how to use contextually-appropriate language in social interaction. Some learning platform has been developed in English teaching and learning. For example, Yang et al. (2010) designed the PILE system that provides a physically interactive learning environment for elementary school students. Students can manipulate virtual objects to do some language tasks such as identifying English letters and understanding phrases. Results found that students did better in English achievement tests and had higher learning motivation after using this human-computer interaction system.

Previous studies have supported that VR technology contributes to oral English language learning (e.g., Shih & Yang, 2008; Blake, 2011), but there is less evidence about how the specific variables about VR, such as immersion, influence individuals' oral English performance and learning experience. Generally speaking, high immersion could enhance authenticity of communication environments. As mentioned earlier, oral English learning emphasizes authenticity of social interaction (e.g., Jauregi et al., 2011). Therefore, we expect that VR with higher immersive level, which could make individuals feel like in the real world, would promote better outcomes of oral English language learning than VR with lower immersive level.

González-Lloret and Ortega (2014) believed that technology-mediated tasks may promote more engagement and participation, enhance their motivation, and decrease learners' anxiety. Immersive VR technology has the potential to generate these effects. First, immersive VR learning environments could increase learners' engagement. It is reasonable to think that, compared with just reading texts or watching video, learners would be more likely to concentrate on and have an interest in learning if they are immersed in social interactions such as shopping in a mall or checking in at a hotel that are highly close to real life. Second, immersive VR learning condition could reduce learners' anxiety. English learning anxiety derives from learners' competitive sense in language performance, apprehension over communication, and fear of negative evaluation from others (Ellis, 2008). Some researchers have found that computer-mediated interactions could alleviate learners' anxiety (Kern, 1995; Warschauer, 1996). The reason might be that compared with a face-to-face condition, a VR learning condition does not need learners to give an immediate response (Beauvois, 1992). Additionally, in human-computer interactions, learners do not need to worry about being laughed at because of any language mistakes. Third, learners might become more confident in their oral English competence in immersive VR learning environments. Learners' English language competence will improve as a result of more engagement (e.g., Hamari et al., 2016) and less anxiety (e.g., Huang & Hung, 2013). Learners' self-efficacy could be increased when they can speak English accurately and fluently in an authentic communication. When learners' self-efficacy is enhanced, in turn, they could make more effort to improve their speaking skills and further reduce anxiety (Sardegna et al., 2018). Thus, they would enter a virtuous circle in which oral English language competence would develop continuously.

1.4 The present study

Existing literature lacks comprehensive understanding of how variations in immersion level within VR environments influence oral English language learning outcomes, including performance, engagement, anxiety, and self-efficacy. The aim of this study was to examine the effects of immersive VR technology on oral English language learning by comparing participants' performance in non-immersive and fully immersive VR learning conditions. Based on previous research suggesting the positive impact of VR technology on English learning(e.g., Shih & Yang, 2008; Blake, 2011), we expected that:

H1: Individuals' oral English language performance would improve more in the fully immersive condition than that in the non-immersive condition.

Furthermore, considering the immersive nature of fully immersive VR environments and its potential to enhance learners' sense of presence and interactivity, we posit:

H2: Individuals' engagement in oral English language learning would increase in the fully immersive condition than that in the non-immersive condition.

Moreover, given the potential of fully immersive VR environments to create a more authentic and supportive learning atmosphere, we suggest:

H3: Individuals' learning anxiety would decrease in the fully immersive condition than that in the non-immersive condition.

Last but not least, as individuals become more immersed and actively participate in oral English language learning activities within fully immersive VR environments, the following hypothesis is posited:

H4: Individuals' self-efficacy would increase in the fully immersive condition than that in the non-immersive condition.

2. Method

2.1 Participants

We used G-power (Faul et al., 2007) to determine the number of participants for our project. According to the previous meta-analysis in the field of immersive learning technologies in English language learning (Altun & Lee, 2020), the overall effect size was proved to be positive and large (Hedge's g = .84). Given a power of 0.8 and a large effect size (Cohen's d = 0.8) for independent T-test, the experimental group and control group would need at least 21 participants each.

Forty-three Mandarin-speaking Chinese students (37 females, 6 males; M_{age} = 19.05 years, SD = 0.57) were recruited through advertising at a Chinese university campus. Twenty-two participants were randomly assigned to the fully immersive condition (experimental group) and 21 to the non-immersive condition (control group). All participants were blind to the experimental purpose and were paid 200 yuan (about $30 US) for their participation. All participants were taking the same English language course, which was a compulsory course for all freshmen enrolled at the university. Ethics permission for the project was obtained from the Ethics Review Committee of Faculty of Psychology, Beijing Normal University, and written consent was given by each participant prior to testing.

2.2 Task

VR English Learning in Hotel (VELH) was used as a learning platform. VELH is an application developed by Yuandian Tech, a Chinese company focusing on VR education. This application can run on both HMD and PC monitor. When participants are equipped with HMD, they can use the helmet and controllers for human-computer interaction. However, when using PC monitor as the output method, participants will use mouse and earbuds for interaction. The most important reason for choosing VELH is that learners will face the exact same scenarios and tasks in both fully immersive and non-immersive VR learning conditions.

Fig. 1. A dialogue between the learner and the NPC in the room.

When playing VELH, participants are put in several simulated hotel scenarios. The difficulty level of the daily dialogue in the task matched to the typical difficulty level found in college English courses. They can see the surroundings and subtitles, hear the voice of the non-player character (NPC), and speak to the NPC. Their oral language performance will be automatically scored (see Fig. 1). Then the system will provide feedback, including the correct texts and accurate pronunciations. Participants may choose the back button to repeat some sentences to achieve a higher score, or the next button to move on to the next dialogue.

2.3 Measures

2.3.1 Oral English performance

Participants' learning performance was automatically scored by machine algorithms built in the VELH. Two indicators were selected in the present study. One was accuracy which reflected whether participants' pronunciation of the words presented on the screen was right; the other one was reading fluency which assessed language rhythm, mainly referring to pause and continuity when reading a sentence.

2.3.2 Engagement

Two indicators were used to assess participants' engagement level. One was recording time which was saved in the log file. When the participants joined a certain conversation with the NPC, they had the choice to record their own voices in the conversation. If they did, the system would provide feedbacks for each sentence. Then the participants could choose to record the certain sentence repeatedly to improve their accuracy or fluency. Therefore, recording time could reflect how much effort participants were willing to make to learn English. The other indicator was heart rate measured by the heart rate monitor Polar m430 that was produced by the Polar company. Heart rate could reflect participants' arousal level (Wu & Xu, 1994; Yang et al., 2017). The higher level of arousal suggested that participants were paying more attention on the learning task (Chen, 2017). In the present study, heart rate change (HRC) which equaled to average heart rate during learning minus the base line measured before learning, was used as the indicator of learning engagement.

2.3.3 Learning anxiety

A 36-item measure of second language-skill-specific anxiety scales (Cheng, 2017) was adopted to assess participants' level of English learning anxiety in listening, speaking, reading and writing. Participants were asked to rate such items as "When speaking in English, I often worry that I will make language mistakes" on a 5-point Likert scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*). A higher score indicated higher anxiety. In the pretest of the current study, reliability coefficients of the whole scale and four dimensions were .90, .88, .83, .75 and .75, respectively. In the posttest, reliability coefficients were .89, .84, .83, .75, and .72, respectively.

2.3.4 Self-efficacy

The questionnaire of English self-efficacy (Wang et al., 2013) was used to measure participants' general level of self-efficacy in using English. The questionnaire consisted of 32 items and four dimensions (i.e., listening, speaking, reading and writing). Participants were asked to answer items like "Can you tell a story in English" on a 7-point Likert scale ranging from 1 (*I cannot do it at all*) to 7 (*I can do it* *very well*). A higher score reflected higher self-efficacy. In the pretest of the current study, reliability coefficients of the whole scale and four dimensions were .95, .88, .81, .90 and .85, respectively. In the posttest, reliability coefficients were .95, .89, .83, .92, and .86, respectively.

2.3.5 Personality

Agreeableness, conscientiousness, and openness were measured by the 50-item IPIP representation of the Goldberg (1992) markers for the Big-Five factor structure (see Administering IPIP Measures, with a 50-item Sample Questionnaire, Retrieved October 18, 2021). Participants were asked to rate on a 5-point Likert scale ranging from 1 (*not like me at all*) to 5 (*like me very much*). In the current study, reliability coefficients of the five dimensions ranged from .77 to .91.

2.3.6 Previous English language achievement

The score of English language course final examination for the last semester was used as the measurement of participants' English language achievement before taking part in the present study.

2.4 Procedure

Participants were randomly assigned to the non-immersive and fully immersive VR learning conditions. In the fully immersive condition, participants wore the HTC VIVE-VVR-H RATING head-mounted display unit and its accessories. In the non-immersive condition, participants learned the same content on a personal computer (see Fig. 2). In both conditions, participants took four lessons in four days during two weeks with a three-day interval. Each lesson focused on a specific hotel scenario and consisted of three stages. At the first stage, participants would hear a conversation from the earbuds and see the text of each sentences on the screen. After each sentence, participants were asked to repeat that sentence. Second, the participants practiced the same conversation by speaking to the NPC (see Fig. 3). Feedbacks would be provided on the screen after each sentence, including the score of the whole sentence and the detail information for each word. For example, green words showed good pronunciation and fluency, while red words represented bad performance. The participants could choose the option to re-play the correct voice, and then they might decide to repeat it for several times according to the feedback. At the third stage, participants needed to listen to the whole conversation between themselves and the NPC, which was recorded at stage two.

At Lesson 1 (day 1), participants filled in the scales of learning anxiety, self-efficacy and their demographic information, and provided their English examination score. They also tried the three learning stages to learn how to use the software before the formal experiment. They completed the three-stage learning process once after the trial. At lesson 2 (day 2), lesson 3 (day 3), and lesson 4 (day 4), participants conducted the three steps at least twice so that learning time could reach 25 minutes. In addition, at day 2 and day 4, participants took a test without subtitles, in which they needed to organize the language to complete the conversation. Personality scale was completed after English learning at day 3. At the end of day 4, they completed the scales of learning anxiety and self-efficacy as the results of posttest. Participants had to take a 10-minute rest before learning every day, and their heart rates were recorded.

Fig. 2. Participants in the fully immersive condition (up) and non-immersive condition (down).

2.5 Data analysis

We employed SPSS 22.0 to achieve the data analysis. A 2 (non-immersive vs fully immersive) ×2 (Day 1 vs Day 4) repeated measures ANOVA was conducted to analyze the variation of each dependent variable in two learning conditions.

3. Results

Descriptive statistics of the main study variables (i.e., oral English language performance, engagement, learning anxiety, self-efficacy) are presented in Table 1. An inspection of the distributional properties of the measures revealed that they were within acceptable levels (Tabachnick & Fidell, 2012). Next, repeated measures were performed to examine whether participants in the fully immersive VR condition did better in engagement, learning anxiety, self-efficacy and oral English language performance than those in the non-immersive condition.

3.1 Descriptive analysis

3.2 Oral English language performance

Given the possible influences of English competence and personality traits such as openness (Laidra et al., 2007;

Table 1. Means and standard deviations of each dependent variables in two conditions at different times.

Dependent variable	Test day	Fully immersive (n=22)		Non-immersive (n=21)		Total (n=43)	
		Mean	SD	Mean	SD	Mean	SD
Accuracy	2	75.23	6.32	75.52	4.99	75.37	5.64
	4	78.32	6.24	74.24	6.43	76.33	6.59
Fluency	2	86.68	2.89	86.86	1.71	86.77	2.36
	4	87.00	2.37	86.38	2.92	86.70	2.64
Recording time	2	637.27	67.38	604.33	68.32	621.19	69.07
	4	753.36	70.97	689.04	84.94	721.95	83.74
Heart rate change (HRC)	1	9.36	6.68	-2.02	5.69	3.81	8.42
	2	10.40	5.76	-0.25	4.69	5.20	7.49
	3	14.60	6.26	1.21	5.27	8.06	8.87
	4	17.73	9.32	-2.23	10.02	7.98	13.90
Speaking anxiety	1	3.26	0.69	3.27	0.80	3.26	0.74
	4	2.84	0.64	3.09	0.82	2.96	0.74
Listening anxiety	1	2.79	0.95	2.84	0.66	2.81	0.81
	4	2.57	0.78	2.83	0.64	2.70	0.72
Reading anxiety	1	2.25	0.49	2.15	0.60	2.20	0.54
	4	2.04	0.54	2.37	0.55	2.20	0.56
Writing anxiety	1	2.42	0.70	2.45	0.48	2.44	0.59
	4	2.43	0.62	2.61	0.53	2.52	0.58
Total anxiety	1	2.68	0.59	2.68	0.38	2.68	0.50
	4	2.47	0.54	2.72	0.40	2.59	0.49
Recording time	2	637.27	67.38	604.33	68.32	621.19	69.07
	4	753.36	70.97	689.04	84.94	721.95	83.74
Listening efficacy	1	4.01	0.82	3.76	0.96	3.88	0.89
	4	4.21	0.80	4.04	1.21	4.13	1.01
Reading efficacy	1	4.44	0.58	4.35	0.99	4.39	0.80
	4	4.60	0.65	4.56	0.99	4.58	0.83
Writing efficacy	1	4.56	0.76	4.50	1.07	4.53	0.92
	4	4.96	0.75	4.64	1.07	4.80	0.92
Recording time	1	4.38	0.72	4.27	0.88	4.33	0.79
	4	4.64	0.67	4.39	0.94	4.52	0.81

Fig. 3. Screenshot of a participant in the fully immersive condition.

Weibel et al., 2010), conscientiousness (Wagerman & Funder, 2007; Vedel, 2014) and agreeableness (Papamitsiou & Economides, 2017; Shin, 2018) on immersion or academic achievement, we controlled for their effects statistically. Specifically, we used 2 × 2 repeated measures ANOVA to examine the influence of learning condition on accuracy and fluency respectively after controlling for potential confounding effects of English competence and three personality traits. Learning condition (non-immersive vs fully immersive) served as a between-subject factor, and assessment time (Day 2 and 4) was taken as a within-subject factor.

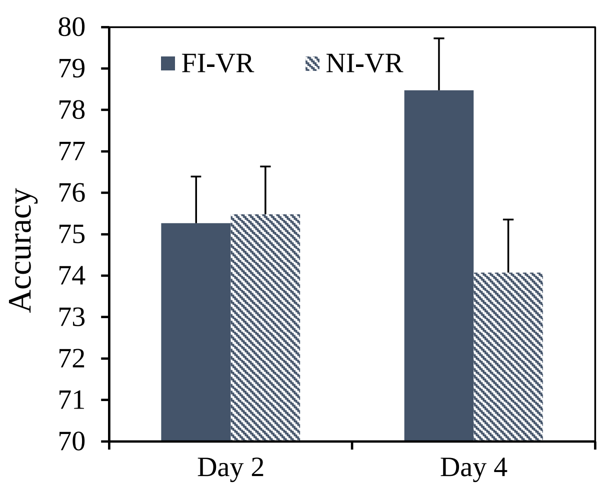

Fig. 4. Simple effect analysis demonstrating the influence of immersive VR condition on accuracy.

For accuracy, there was a significant interaction effect between learning condition and assessment time, F (1, 37) = 10.44, p = .003, $\eta^2_{partial}$ = .22. Results of simple effect analysis showed that participants' accuracy level in the fully immersive VR condition was significantly higher than participants' accuracy level in the non-immersive VR condition on Day 4 [F (1, 39) = 7.10, p = .011, $\eta^2_{partial}$ = .15], while there was no significant difference in accuracy between the two conditions in Day 2 [F (1, 39) = 0.08, p = .780, $\eta^2_{partial} < .01$]. The interaction pattern was depicted in Fig. 4, which showed that participants in the fully immersive VR condition made more progress in accuracy than those in the non-immersive VR condition.

For fluency, both the main effects of learning condition [F (1, 37) = .83, p = .37, $\eta^2_{partial}$ = .02] and assessment time [F (1, 37) = .01, p = .92, $\eta^2_{partial} < .01$] were not significant as well as their interaction effect [F (1, 37) = .74, p = .40, $\eta^2_{partial}$ = .02), suggesting that there was no significant difference in fluency between two learning conditions.

3.3 Engagement

Similarly, repeated measures were used to examine the influence of immersive VR learning condition on participants' engagement (i.e., recording time and heart rate change), after controlling for the potential confounding effects of English competence and personality traits. Learning condition (non-immersive vs fully immersive) served as a between-subject factor, and assessment time was taken as a within-subject factor.

For recording time, we conducted a 2 (non-immersive vs fully immersive) ×2 (Day 2 vs 4) repeated measures ANOVA and found that the main effect of learning condition was significant, F (1, 37) = 5.62, p = .023, $\eta^2_{partial}$ = .13. That is, recording time in the fully immersive VR condition was longer than in non-immersive VR condition, suggesting that participants in the fully immersive condition did more speaking exercises than those in the non-immersive condition. The main effect of assessment time [F (1, 37) = 2.94, p =

.095, $\eta^2_{partial}$ = .07] and the interaction effect [F (1, 37) = 3.62, p = .065, $\eta^2_{partial}$ = .09] were not significant.

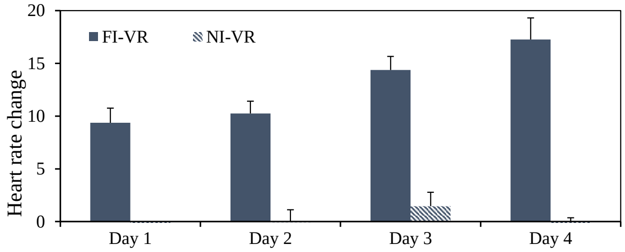

Fig. 5. Simple effect analysis demonstrating the influence of immersive VR condition on HRC.

For HRC, we conducted a 2 (non-immersive vs fully immersive) ×4 (Day 1 to Day 4) repeated measures ANOVA and found that the main effect of learning condition was significant [F (1, 37) = 87.21, p < .001, $\eta^2_{partial}$ = .70], while the main effect of assessment time was not significant [F (3, 111) = 1.93, p = .151, $\eta^2_{partial}$ = .05]. The interaction effect between learning condition and assessment time was significant, F (3, 111) = 4.00, p = .022, $\eta^2_{partial}$ = .10. Results of simple effect analysis showed that HRC in the fully immersive VR condition were higher than those in the non-immersive VR condition in all four days [Day 1: F (1, 40) = 33.31, p < .001, $\eta^2_{partial}$ = .45; Day 2: F (1, 40) = 43.50, p <.001, $\eta^2_{partial}$ = .52; Day 3: F (1, 40) = 55.69, p < .001, $\eta^2_{partial}$ = .58; Day4: F (1, 40) = 45.64, p < .001, $\eta^2_{partial}$ = .53]. The interaction pattern was depicted in Fig. 5.

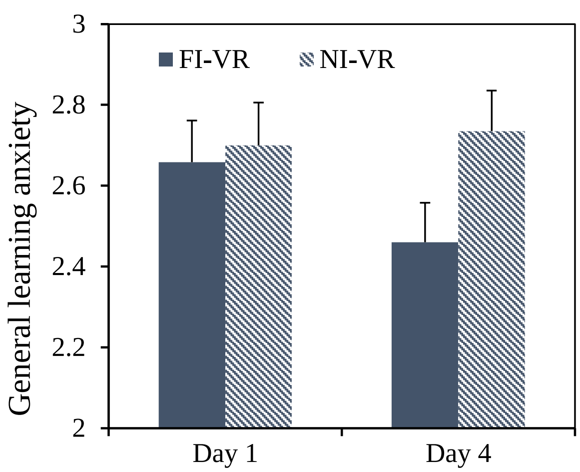

Fig. 6. Simple effect analysis demonstrating the influence of immersive VR condition on general learning anxiety.

3.4 Learning anxiety

First, a 2 (non-immersive vs fully immersive) ×2 (Day 1 vs Day 4) repeated measures ANOVA was used to analyze the variation of general learning anxiety in two learning conditions after controlling for the potential confounding effects of English competence and personality traits. Results showed that the interaction effect of learning condition and assessment time was marginally significant, F (1, 37) = 3.81, p = .059, $\eta^2_{partial}$ = .09. The simple effect analysis revealed that the score in Day 4 was significantly lower than the score in Day 1 in the fully immersive VR condition [F (1, 39) = 4.83, p = .034, $\eta^2_{partial}$ = .11], while the difference between the scores in Day 1 and 4 was not significant in the non-immersive VR condition [F (1, 39) = 2.23, p = .143, $\eta^2_{partial}$ = .005]. The interaction pattern was depicted in Fig. 6.

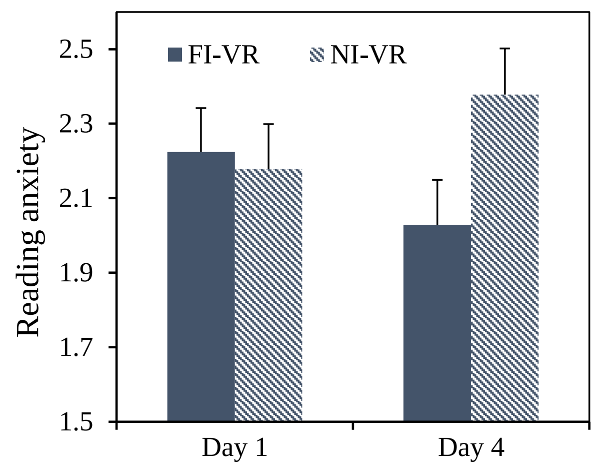

Fig. 7. Simple effect analysis demonstrating the influence of immersive VR condition on reading-anxiety.

3.5 Self-efficacy

First, a 2 (non-immersive vs fully immersive) ×2 (Day 1 vs Day 4) repeated measures ANOVA was used to analyze the variation of general self-efficacy in two learning conditions after controlling for the potential confounding effects of English competence and personality traits. Results showed that both the main effects and interaction effects were not significant. Next, we analyzed the differences of listening-, speaking-, reading- and writing-efficacy before and after learning for two groups. Results from repeated measures ANOVA showed that both the main effects and the interaction effects were not significant for listening-, reading- and writing-efficacy. However, for speaking-efficacy, the interaction effect of learning condition and assessment time was significant, F (1, 37) = 6.34, p = .016, $\eta^2_{partial}$ = .15. The simple effect analysis revealed that the score in Day 4 was significantly higher than the score in Day 1 in the fully immersive VR condition [F (1, 39) = 4.97, p = .032, $\eta^2_{partial}$ = .11], while there was no significant difference between the scores in Day 4 and 1 in the non-immersive condition [F (1, 39) = 1.86, p = .181, $\eta^2_{partial}$ = .05]. The interaction pattern was depicted in Fig. 8.

4. Discussion

In the current study, we examined the effects of immersive VR learning environments on oral English language learning. Specifically, we compared the changes of participants' oral English language performance, engagement, emotional and motivational status in the fully immersive VR learning condition with those in the non-immersive condition. Our results showed that participants in the fully immersive VR condition showed higher accuracy, lower anxiety and higher speaking-efficacy in the posttest than participants in the non-immersive VR condition, while there was no significant difference between two conditions in the pretest. Participants in the fully immersive VR condition showed more engagement (HRC) than those in the non-immersive VR condition on each assessment time. However, participants' fluency did not change in either group.

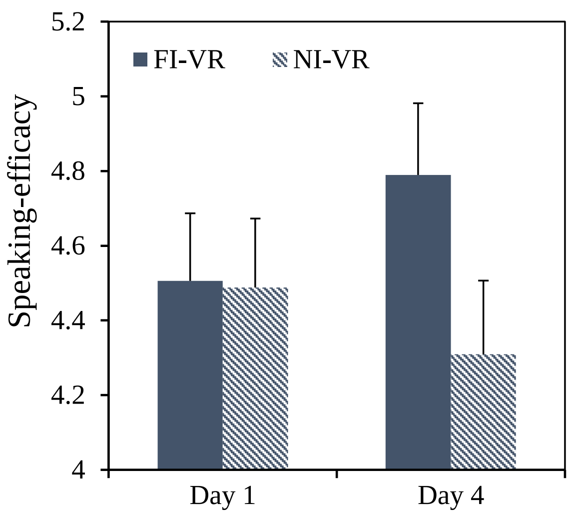

Fig. 8. Simple effect analysis demonstrating the influence of immersive VR condition on speaking-efficacy.

4.1 The role of immersive VR technology in oral English language learning

Consistent with the extant findings about positive relations between immersive VR technology and learning outcomes (e.g., Morgan, 2013; Cakiroğlu & Gökoğlu, 2019), the current study provided new evidence for it in the subject of English language learning. Our findings indicated that individuals' oral English performance, engagement, learning anxiety, and speaking-efficacy could all be improved in the fully immersive VR environments.

First, we found that highly immersive VR could enhance individuals' accuracy of oral English. A possible reason is that individuals in the fully immersive VR condition were more engaged in oral English learning. Specifically, the analysis for recording time showed that individuals who were highly immersed in learning spent more time doing more speaking exercises than those in the non-immersive environments. Additionally, for participants in the fully immersive VR condition, their heart rate differences rose from Day 1/2 to Day 3/4, suggesting their arousal level increased and they concentrated more on learning. However, as another indicator of oral English performance, fluency did not improve in VR environments, as we predicted. As mentioned earlier, fluency mainly reflected the language rhythm which refers to pause and continuity when speaking English. Individuals need to have a basic understanding of grammar and language meaning in order to make correct pause or continuity (Lennon, 1990; Kahng, 2014). Simply put, fluency taps into comprehension of a sentence, while accuracy could be acquired without understanding its meaning. Therefore, it is reasonable to think that fluency might require higher level of language knowledge and competence and need more time to practice than accuracy, which is a possible explanation for the unimproved fluency in the current study. Future studies should examine the effects of immersive VR technology on oral English fluency in a longer learning period.

Second, individuals' emotional and motivational status were improved by the immersive VR technology. Consistent with our hypotheses (see Hypothesis 3 and 4), individuals' anxiety for English learning, especially reading anxiety decreased, and speaking self-efficacy increased in the fully immersive VR environment. The current study provided the subtitles, so "reading" was understood as reading the subtitles aloud. "speaking" should be talking without the subtitles and only appeared in the test. Therefore, reasonably, it is reading anxiety, not speaking anxiety, that has a significant change in the fully immersive VR condition. However, in the items to measure reading-efficacy, specific objects such as novel, newspaper, magazine were provided so that participants could explain "reading" as reading some English materials without sound and think what they did in the study was "speaking". That might be the reason why speaking-efficacy changed significantly rather than reading-efficacy in this case.

Notably, previous studies showed that computer-mediated interactions could reduce learners' anxiety (Kern, 1995; Warschauer, 1996). However, the current study indicated that the immersive level of computer-mediated conditions would influence this effect. Participants' anxiety reduced only in the fully immersive condition and did not change in the non-immersive condition. Reading anxiety even increased in the non-immersive condition. The current study provided a computer-mediated environment in which learners in both conditions could avoid facing a real person directly and regulate their own learning speed, but only those who have the feeling of being in an authentic communication alleviated their anxiety. This might be because immersive feeling makes individuals more concentrate on the process of conversation and less think about others' evaluation.

4.2 Immersive VR technology and Chinese oral English language learning

One of the biggest problems that Chinese have met in English learning is that they do not have an authentic communication environment to practice English. Most of the time they learn and memorize English from textbooks, but do not have chances to do exercises of using the knowledge. Actual language use promotes second language acquisition (Tyler, 2010), so Chinese have spent a large amount of time learning English, but their English proficiency is still in a low level (EF Education First, 2018). Many Chinese have found that they are not able to understand what people are saying in English TV and films, or feel difficult to talk with English natives. The big discrepancy between the current level and learning goals is likely to make them feel anxious and unconfident, which could reduce their motivation to learn

English.

VR technology provides a solution which could help to remove the barriers that have blocked the development of Chinese oral English competence. Specifically, VR technology can make English learners be isolated from the real world and immersed in simulated social interactions highly close to real life (Lee & Wong, 2008). In the fully immersive VR environment, learners can learn authentic language examples in a particular situation and then use them to make meaningful communications in a contextually appropriate way. Our findings indicated that immersive VR learning condition has the potential to improve learners' oral English performance, alleviate their anxiety and enhance their self-efficacy. There might be more advantages that go beyond our findings. For example, language forms that are learned in an authentic situation are more likely to be retrieved and used in similar real-life communications according to context-dependent effect (Godden & Baddeley, 1975). More research and education practice are needed to further explore how immersive VR technology enhances oral English performance.

5. Limitations and future directions

The current study has some limitations worth mentioning. First, we provided relatively authentic language examples and communication environments to the learners by using the immersive VR technology, but the designed learning process lacks flexibility and fun. The content of conversations is unchangeable, and the learners cannot use different language forms or change the communicative process. Moreover, these fixed language materials make it impossible to evaluate other indicators of oral English performance, such as grammatical and lexical complexity (Fortune & Ju, 2017). Tedious and repetitive learning process could also make the learners boring and reduce their motivation. Future research may choose a learning software with more flexibility and fun to examine the improvement of learners' language competence in immersive VR environments. For example, learning can take the form of a task or game in which learners need to use language to tell information, propose a question, express an opinion and get a communicative outcome. This kind of learning environment might be more authentic and effective, so it would have more value to apply in education. Second, while we observed a notable enhancement in participants' speaking-efficacy, which is a significant challenge faced by Chinese students in achieving English proficiency, the improvement in the other three facets of self-efficacy was comparatively less pronounced. Thus, we assumed that the duration of learning may have been insufficient to fully uncover the potential impacts of immersive VR technology on oral English performance. In the current study, participants were afforded four learning opportunities, each lasting 25 minutes. Given this limited exposure, it becomes challenging to draw a comprehensive conclusion regarding the genuine impact of long-term immersive VR on factors such as accuracy, fluency, and other factors. To address this, researchers should explore the effects of immersive VR with extended learning periods to gain a more thorough understanding.

Overall, based on the results of the current study, the immersive extent of VR learning environments could exert some positive effects on oral English learning for Chinese university students. These findings provide insight into the application of VR technology in second language acquisition. We expect that future studies will reassess these effects with other samples and learning materials. More convenient devices for immersive VR are needed for its generalization in second language acquisition.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (No. 61807011).

Additional information: Author's email

georgiou@ualberta.ca (G. K. Georgiou).

Conflict of interest

The authors declare no competing interest.

References

Administering IPIP Measures, with a 50-item Sample Questionnaire. Retrieved October 18, 2021, from https://ipip.ori.org/New_IPIP-50-item-scale.htm#SampleQuestionnaire

Altun, H. K., & Lee, J. (2020). Immersive Learning Technologies in English Language Teaching: A Meta-Analysis. International Journal of Contents, 16(3), 18–32.

Bachman, L. F. (1990). Constructing measures and measuring constructs. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), The development of second language proficiency (pp. 26–49). Cambridge, England: Cambridge University Press.

Bailey, J. O., Bailenson, J. N., & Casasanto, D. (2016). When does virtual embodiment change our minds? Presence, 25(3), 222–233.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: W. H. Freeman.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. Journal of Management, 38(1), 9–44.

Barmaki, R., & Hughes, C. E. (2015, November). Providing real-time feedback for student teachers in a virtual rehearsal environment. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (pp. 531–537).

Beauvois, M. H. (1992). Computer-assisted classroom discussion in the foreign language classroom: Conversation in slow motion. Foreign Language Annals, 25, 455–464.

Belz, J. A. (2007). The role of computer mediation in the instruction and development of L2 pragmatic competence. Annual Review of Applied Linguistics, 27, 45–75.

Blake, R. J. (2011). Current trends in online language learning.

Annual Review of Applied Linguistics, 31, 19–35.
Cakiroğlu, Ü., & Gökoğlu, S. (2019). Development of fire safety behavioral skills via virtual reality. Computers & Education, 133, 56–68.
Chen, J. (2017). Electronic game players' attention bias and aggressive behaviors (unpublished doctoral dissertation). Qufu Normal University, Qufu.
Cheng, Y. S. (2017). Development and preliminary validation of four brief measures of L2 language-skill-specific anxiety. System, 68, 15–25.
EF Education First. (2018). EF English Proficiency Index. Retrieved October 18, 2021, from https://www.ef.edu/epi/
Ellis, R. (2008). The study of second language acquisition (2nd ed.). Oxford, England: Oxford University Press.
Fassbender, E., Richards, D., Bilgin, A., Thompson, W. F., & Heiden, W. (2012). Vir-School: the effect of background music and immersive display systems on memory for facts learned in an educational virtual environment. Computers and Education, 58(1), 490–500.
Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175–191.
Fortune, T. W., & Ju, Z. (2017). Assessing and exploring the oral proficiency of young mandarin immersion learners. Annual Review of Applied Linguistics, 37, 264–287.
Godden, D., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. British Journal of Psychology, 66, 325–331.
Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. Psychological Assessment, 4, 26–42.
González-Lloret, M., & Ortega, L. (2014). Towards technology-mediated TBLT: An introduction. In M. González-Lloret & L. Ortega (Eds.), Technology-mediated TBLT: Researching technology and tasks (pp. 1–22). Amsterdam, the Netherlands: John Benjamins.
Gutiérrez, F., Pierce, J., Vergara, V., et al. (2007). The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. Studies in Health Technology and Information, 125, 155–160.
Haans, A., & IJsselsteijn, W. A. (2012). Embodiment and telepresence: Toward a comprehensive theoretical framework. Interacting with Computers, 24, 211–218.
Hamari, J., Shernoff, D., Row, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. Computers in Human Behavior, 54, 170–179.
Hansen, A. H., & Mossberg, L. (2013). Consumer immersion: A key to extraordinary experiences. In J. Sundbo, & F. Sørensen (Eds.). Handbook on the experience economy (pp. 209–227). Cheltenham, UK: Edward Elgar.
Harley, B., Cummins, J., Swain, M., & Allen, P. (1990).The nature of language proficiency. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), The development of second language proficiency (pp. 7–25). Cambridge, England: Cambridge University Press.
Hawkins, D. G. (1995). Virtual reality and passive simulators: the future of fun. In F. Biocca, & M. R. Levy (Eds.), Communication in the age of virtual reality (pp. 159–189). New York: Routledge.
Huang, H.-T., & Hung, S.-T. (2013). Comparing the effects of test anxiety on independent and integrated speaking test performance. TESOL Quarterly, 47, 244–269.
Jackson, J. (2002). Reticence in second language case discussions: Anxiety and aspirations. System, 30, 65–84.
Jauregi, K., Canto, S., de Graaff, R., Koenraad, T., &Moonen, M. (2011). Verbal interaction in Second Life: Towards a pedagogic framework for task design. Computer Assisted Language Learning, 24, 77–101.
Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. Language Learning, 64(4), 809–854.
Kern, R. G. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. Modern Language Journal, 79, 457–476.
Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. Personality and Individual Differences, 42(3), 441–451.
Lee, A. L., & Wong, K. W. (2008). A review of using virtual reality for learning. In Z. Pan et al. (Eds.), Transactions on Edutainment I, LNCS 5080 (pp. 231–241). Berlin Heidelberg: Springer-Verlag.
Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. Language Learning, 40, 387–417.
Liu, Z., Yao, X., & Hu, S. (2012). The structure analysis of college students' self-concept in L2, anxiety and motivaitional learning behavior. Foreign Language World, 153(6), 28–38.
Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. Applied Linguistics, 4(2), 126–141.
MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. Language Learning, 44, 283–305.
Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. Journal of Speech, Language, and Hearing Research., 50(4), 940–967.
Morgan, E. J. (2013). Virtual worlds: Integrating second life into the history classroom. The History Teacher, 46(4), 547–559.
Nissim, Y., & Weissblueth, E. (2017). Virtual Reality (VR) as a Source for Self-Efficacy in Teacher Training. International Education Studies, 10(8), 52–59.
Papamitsiou, Z., & Economides, A. A. (2017). Exhibiting achievement behavior during computer-based testing: What temporal trace data and personality traits tell us? Computers in Human Behavior, 75, 423–438.
Peng, Y. (2006). An investigation of status and problems of college students' oral English learning. College Days, 2,

62–64.

Sardegna, V. G., Lee, J., & Kusey, C. (2018). Self-efficacy, attitudes, and choice of strategies for English pronunciation learning. Language Learning, 68(1), 83–114.

Shih, Y.-C., & Yang, M.-T. (2008). A collaborative virtual environment for situated langage learning using VEC3D. Educational Technology & Society, 11(1), 56–68.

Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? Computers in Human Behavior, 78, 64–73.

Song, K. S., & Lee, W. Y. (2002). A virtual reality application for geometry classes. Journal of Computer Assisted Learning, 18, 149–156.

Tyler, A. (2010). Usage-based approaches to language and their applications to second language learning. Annual Review of Applied Linguistics, 30, 270–291.

Vedel, A. (2014). The Big Five and tertiary academic performance: A systematic review and meta-analysis. Personality & Individual Differences, 71(2), 66–76.

Wagerman, S. A., & Funder, D. C. (2007). Acquaintance reports of personality and academic achievement: A case for conscientiousness. Journal of Research in Personality, 41(1), 221–229.

Wang, C., Schwab, G., Fenn, P., & Chang, M. (2013). Self-efficacy and self-regulated learning strategies for English language learners: Comparison between Chinese and German college students. Journal of Educational and Developmental Psychology, 3(1),173–191.

Warschauer, M. (1996). Comparing face-to-face and electronic communication in the second language classroom. CALICO Journal, 13, 7–25.

Weibel, D., Wissmath, B., & Mast, F. W. (2010). Immersion in mediated environments: The role of personality traits. Cyberpsychology, Behavior and Social Networking, 13, 251–256.

Wu, Z., & Xu, S. (1994). The preliminary research on the relationship between arousal level and adult age differences of memory. Acta Psychologica Sinica, 26(2), 195–198.

Yang, J. C., Chen, C. H., & Jeng, M. C. (2010). Integrating video-capture virtual reality technology into a physically interactive learning environment for English learning. Computers & Education, 55, 1346–1356.

Yang, Z., Jia, W., Liu, G., & Sun, M. (2017). Quantifying mental arousal levels in daily living using additional heart rate. Biomedical Signal Processing & Control, 33, 368-378.

Zeng, F. (2012). The reason and strategies of “dumb English”. Education Teaching Forum, 39, 71–72.