Education and Information Technologies (2024) 29:1379–1405
https://doi.org/10.1007/s10639-023-11738-0

The effects of virtual reality on EFL learning: A meta-analysis

Xiang bin Qiu[1] · Cheng Shan[1] · Jin Yao[2] · Qing ke Fu[1]

Received: 28 May 2022 / Accepted: 16 March 2023 / Published online: 28 March 2023

Abstract

In recent years, an increasing number of teachers and researchers have used virtual reality (VR) to enhance English as a foreign language (EFL) learning, but the learning effects they found varied. Because of these differences, we conducted a meta-analysis. This study aims to compare the effectiveness of VR-based EFL learning methods with traditional EFL learning methods, and to determine what factors led to these results. We searched for articles published from 2015 to 2021 using Web of Science, ERIC, and Google Scholar. A total of 23 randomized controlled trials (RCTs) or quasi-experimental studies were included, and the effect size (ES) was calculated. The results show that the VR-based EFL learning method is significantly better than the traditional EFL learning method and has a small positive effect ($g = .445$). We also coded the characteristics of the samples to examine their moderating effect on the results. We found that educational levels, country/area, and EFL learning outcomes have significant differences that can explain the variance in ES among the samples. According to the analysis results, we also provide suggestions on which learning outcomes are most helpful, how to choose the best educational stages, learning materials, and type of VR to promote EFL learning performance, and directions for future research and practice.

Keywords Virtual reality · English as a foreign language (EFL) · Language learning · Meta-analysis

✉ Xiang bin Qiu
qxbin@zjhu.edu.cn

[1] School of Teacher Education, Huzhou University, Huzhou, China

[2] School of Education Science, Huaiyin Normal University, Huaiyin, China

1 Introduction

1.1 EFL learning

English has become the most widely used language in the world due to its widespread use in science, business, commerce, and technology (Su, 2006). In today's society, many countries attach great importance to teaching and learning English from primary school. For example, the majority of students in Asian countries are required to learn English as a foreign language (EFL) to cultivate their global perspective as citizens and essential competencies for the twenty-first century (Lan, 2015). Because in these countries, English proficiency is considered essential to compete in the international community. However, today's English teaching still faces many challenges.

The process of learning language as a communication bridge between people and society is time-consuming and complicated. Atkinson (2002) emphasized that language is a tool for society. Therefore, a suitable environment is a necessary condition for language learning. Donato and Mccormick (1994) agreed that the best way for foreign language learners to learn a language is to use it in an authentic language environment. However, the lack of an authentic learning environment is an obstacle that cannot be ignored in EFL learning. Some studies have shown that second or foreign language learning is more efficient when it takes place in a natural environment that includes as much physical interaction with the environment as possible (Macedonia, 2014). However, in general, EFL learning takes place in the classroom environment of the school, which does not have sufficient authenticity.

Several problems have long been identified with EFL learning in the classroom. Underdeveloped communicative skills in EFL usually lead to fear of communication, self-consciousness, or silence, which then leads to foreign language anxiety (FLA) (Aslan & Thompson, 2021). In addition, low proficiency EFL learners will have greater English anxiety (Thompson & Lee, 2013). Demotivation is also a daunting problem that weakens EFL students' interest in language learning (Pathan et al., 2020). Factors such as teacher behavior, course content or materials, and lack of self-confidence can potentially have a negative impact on students' language learning outcomes (Krishnan & Pathan, 2013).

In EFL learning, students also have problems in writing, listening, reading, vocabulary, and so on (Jiang & Eslami, 2021; Pérez-Segura et al., 2020; Tai et al., 2020a, b). In particular, the lack of an authentic language expression environment has led to a generally poor listening and speaking proficiency of EFL learners. To reduce or solve the above problems, many scholars are trying to use certain technologies such as robots, augmented reality, and educational games to create an interactive learning environment for students (Almusharraf, 2021; Cheng et al., 2020; Lin et al., 2020).

The social and autonomous requirements of the language learning process are often not met, resulting in unsatisfactory language learning outcomes. Today, the use of technology in education is becoming increasingly sophisticated. The use of

technology for language learning has become a major focus of research and offers new solutions for improving language learning methods. In this study, we will explore the practical implications of applying VR technology to EFL learning.

1.2 VR technology

Virtual Reality (VR) offers an intriguing new way for people to perceive the world and has received considerable attention since its inception. The origins of this technology can be traced back to the twentieth century (Sutherland, 1968). Bridges et al. (2007) proposed the following core definition of VR: a system that aims to simulate real-life experiences by providing topography, motion, and physics that create the illusion of being there. This means that what users see is all virtual through VR equipment. VR can use either the same tools as traditional display technologies (such as televisions, computers, and smartphones) or dedicated head-mounted displays (HMDs), but its fundamental purpose is to simulate or build the environment, and it is designed to revolutionize the interaction between the viewer, the display, and the environment (Yin et al., 2021). Over time and with the development of technology, VR holds great promise to revolutionize healthcare, communication, entertainment, education, manufacturing, and other fields (Chang et al., 2020).

VR systems can be divided into 3 major categories, non-immersive, semi-immersive, and immersive, based on VR hardware equipment and development tools. The first to emerge was non-immersive VR. Non-immersive VR systems, also called desktop VR systems, have the least immersion and the lowest cost, and are currently the most widely used. Users can interact with the virtual world using devices such as a computer mouse and keyboard. Semi-immersive VR systems, such as CAVE (Cave Automatic Virtual Environment), provide a high level of immersion while maintaining the simplicity of desktop VR (Barrett et al., 2011). The immersive VR system is the most expensive and provides the highest level of immersion. Its components include an HMD, tracking devices, and data gloves that surround the user with computer-generated 3D animations that create a sense of being part of the virtual environment (Bamodu & Ye, 2013). Various HMDs such as Oculus Rift, Samsung Gear VR, Google Cardboard, PlayStation VR, and HTC Vive are available on the market (Acar & Cavas, 2020). Unlike the virtual world seen from the outside, HMDs allow for personal viewing, creating a sense of immersion in learning (Wu et al., 2020).

Currently, VR is most widely used in medicine, education, business, and gaming. The application of VR in education is not a new topic. Accordingly, there has been a lot of research on the application of VR to EFL learning. Compared with traditional learning methods, EFL learning based on VR has many advantages. For example, VR can provide a more realistic learning environment, or VR can be repeated for training and reduce costs. A lot of research has been done in the field of EFL in both non-immersive and immersive virtual reality. However, semi-immersive virtual reality systems are not suitable for use in EFL learning because they are uncomfortable to use.

1.3 VR for EFL learning

VR's greatest contribution to improving EFL learning is to create an immersive environment for EFL learners. This enhancement also shows its unique advantages over traditional multimedia. We can use VR software to create the necessary EFL learning environments. Meanwhile, VR technology can be used in all age groups without any restrictions, which also allows VR-based EFL teaching to be carried out in all age groups. Liu (2021) used SL to build a virtual environment for students in their EFL learning, which can not only provide students with learning opportunities without time and space constraints, but also provide a game-like scenario for EFL learning. He conducted research on 132 primary school students, and his results showed that the sentence and conversation levels of the primary school students were significantly improved.

VR technology with fun and reusable features is ideal for use in EFL learning. For example, speaking is very important in EFL learning. However, developing oral presentation skills is often considered to be a time-consuming activity (Chan, 2011). Improving this skill requires considerable simulation and practice to achieve good results. Meanwhile, heavy course loads leave instructors little time to teach oral presentation skills, and creating more opportunities to practice and receive the necessary feedback through more human assistance is neither feasible nor affordable. (Bankowski, 2010; Schneider et al., 2017). In this circumstance, researchers hope to help students solve their EFL learning problems by using devices that can attract them to practice after class or enable more efficient EFL teaching in the classroom. Obviously, VR technology can be one of them. Moreover, if VR can eliminate the need for other training materials or even trainers, it can even be used as a cost-saving measure (Howard & Gutworth, 2020).

Of course, these studies are based on the most basic features of VR. There are also many researchers who have further explored the characteristics of VR combined with learning theory in their experiments, and since learning activities supported by VR emphasize sociality, which is also necessary in EFL learning, VR can promote different aspects of learners (Qiu et al., 2021). Some researchers claim that the realistic nature of VR allows participants to go through psychological processes and act out behaviors similar to what they would experience in a real-life scenario (Coffey et al., 2017). The environment created by VR software could allow students to move around and interact directly with the environment itself or to interact, communicate, and collaborate through avatars, which are still used in learning through virtual environments such as Second Life (SL, a type of VR game) or multiplayer online games (Christoforou et al., 2019). Yu et al. (2020) used SL as a platform to manage collaborative projects for second language teaching and learning among university students. After the experiment, the students' speaking performance was significantly improved due to the rich conversational opportunities provided by the SL tasks.

Different learning strategies and learning activities are also widely used in empirical research. Task-based learning, situational learning, and gamification learning have become the most widely used learning strategies. The fact that learning materials and content can be presented visually in such an environment reflects

the advantages of VR technology in supporting language learning and shows that the application of VR technology can overcome the technical limitations of traditional learning media (Qiu et al., 2021). What's more, Ebadi and Ebadijalal (2020a, b) used collaborative learning activities in their experiment. They divided students into two different groups as museum guides. The quantitative results showed that the oral performance of the students in the experimental group was significantly higher than that of the control group, and the students' willingness to communicate was improved at the end of the experiment. M. R. A. Chen and Hwang (2020a, b) used a teacher-centered approach in their experiment. The experimental results showed that VR-based learning had significant positive effects on students' oral presentations, speaking anxiety, and learning motivation.

Undoubtedly, VR has many obvious advantages in EFL learning. However, some researchers have found that the effect of VR on EFL learning is not obvious. Lin and Wang (2021) claimed that after their experiment, students' creative self-efficacy was partially improved, but the effect was not statistically significant. Dolgunsöz, Yildirim, and Yildirim (2018) found that after their treatment, learners scored slightly higher in the traditional video condition than in the VR condition. This finding is different from what most researchers have observed.

In general, teachers and researchers hope to use VR technology to create a realistic simulation environment for students to enhance their EFL learning performance.

1.4 Study purpose

The above research results show that individual empirical research from a single perspective does not help to determine the overall effectiveness of VR technology in improving students' EFL learning. Therefore, a comprehensive and reliable overall research is needed to answer how to promote EFL learning through VR technology is more effective. Similar studies have been conducted by previous researchers. However, the difference is that they all analyzed the effect of VR based on all kinds of language learning (Acar & Cavas, 2020; Wang et al., 2020).

Therefore, the EFL learning promotion effect of VR still needs to be further clarified. Based on some divergent questions, the answers need deeper systematic research. In order to reduce the uncertainty caused by a single study and obtain more reliable conclusions, in the current article, we conduct a meta-analysis to determine the overall effectiveness of VR training programs for EFL learning development.

VR-based EFL learning involves both EFL learning and VR technology. We have referred to Yassin's (2019) model of teaching language skills and Bai's (2018) model of computer-assisted language learning and teaching. To improve students' language skills, the former focuses on integrating three theories in the process of language teaching and learning, which are behaviorism, cognitivism, and constructivism. The latter focuses on how to take full advantage of technology and explore computer-assisted language learning. Each study has its own perspective. Although the design processes of these studies on VR-based EFL learning are different, they generally include some common characteristic elements. We have extracted the

Table 1 Framework of our analysis

Basic element	Explanation
Learners' background	It is important to consider the age, needs, goals, and environment of the learners. Only then can appropriate learning materials, instructional tasks and activities be developed
Learning theory	Some studies will use the appropriate research theory, such as sociocultural theory, situated learning, and immersion. There are also some studies that take atheoretical approaches without applying theories
Learning activities	This element includes the different learning activities such as personal learning, teacher centered teaching or cooperative learning
Teaching/ Learning resources	This section focuses on the use of different materials for students. Personalized resources, virtualized environment, and open system can give students a better learning effect
Learning gains	This element includes two main learning gains: linguistic gains and affective gains. Learning gains, often referred to as academic gains, include reading, writing, speaking, and listening. Affective gains include motivation, attitude, etc

common elements from their studies to form the research framework for this study, as shown in Table 1.

We aim to synthesize previous research on the learning effects of EFL in VR environments in order to advance research and practice in this area. Based on a review of the relevant literature and research frameworks, we propose five research questions (RQs):

RQ1: What is the overall learning effect of VR-based EFL learning compared to that of traditional EFL learning?
RQ2: In terms of learners' backgrounds, what are the effects of different levels of education on students' EFL learning gains in VR-based EFL learning, and are there significant differences?
RQ3: In terms of learning theories, what kind of learning theories show the best results in VR-based EFL learning?
RQ4: In terms of learning activities, what are the effects of different learning activities on students' EFL learning gains in VR-based EFL learning and are there significant differences?
RQ5: In terms of learning teaching/ learning resources, what kind of learning materials can facilitate students better?
RQ6: In terms of learning gains, what kinds of VR-based EFL learning outcomes are most helpful to students?

2 Method

The evidence from the literature revealed a widespread implementation of VR applications in EFL learning. Therefore, we conducted a meta-analysis to examine the results of our research questions presented above. Meta-analyses can

statistically aggregate previous findings and provide summary statistics. Thus, we can describe the literature and conduct a moderator analysis by coding relevant studies. For example, we could code what type of learning materials the studies used, and we could divide the analyses into two groups and compare the data. Because our hypotheses were about a new cross-cutting area of research, we felt that a meta-analysis was the most appropriate way to test our hypotheses.

In this study, we used the meta-analysis process proposed by Cooper (2015) and Hansen et al. (2022) as a design framework and referred to the five steps of meta-analysis proposed by Paul & Barari (2022), which include defining the research question, collecting data, preparing data, analyzing data, and reporting conclusions.

2.1 Study search

We searched the Educational Resources Information Clearinghouse (ERIC), Web of Science, and Google Scholar for studies published between 2015 and 2021. As mentioned above, VR has been widely used in EFL learning in recent years. Before 2015, most EFL learning studies were based on technology-enhanced instruction, such as weblogs, e-learning or digital games (Doris et al., 2015; Genzola, 2016; Tan, 2015). Only a few studies have used VR software, which is immature and rare (Berns & Reyes-Sanchez, 2021). Thus, the number of studies on VR-based EFL learning before 2015 is very small (Shadiev & Yang, 2020). In 2016, HTC, Oculus, and SONY released three types of immersive HMDs. Therefore, we believe that including only studies since 2015 would show a good response to this trend and provide an accurate understanding of the study.

Based on the above three databases, we selected appropriate keywords to search. The keywords are grouped into two categories: ("virtual reality") AND ("EFL learning"):

1. Keywords related to virtual reality: ("virtual reality" or "virtual world" or "virtual environment" or "immersive" or "simulation" or "virtual classroom").
2. EFL-related keywords: ("EFL" or "L2" or "second language" or "language education" or "language learning" or "language acquisition" or "foreign language" or "language courses" or "second language learners" or "language classroom").

The combination of these keywords was used to search the three databases. We included only the first 170 results of our Google Scholar searches in our initial coding database because the results returned after 150 were only related to the search terms (e.g., the search system returned many results for AR-based or web-based language learning). Web of Science returned 1135 results, ERIC returned 507 results, and the total from all databases was 1812 results. After removing other types of articles, including proceedings papers, early access papers, reviews, dissertations, and book chapters, the final sample was 1057 papers.

2.2 Inclusion and exclusion criteria

The studies had to meet the following criteria to be included in this meta-analysis: (a) the articles were written in English, (b) the target foreign language of the study was English, (c) the VR-based EFL learning method was compared to the traditional EFL learning method, and (d) the study was a randomized controlled trial (RCT) or quasi-experimental study. Exclusion criteria were as follows: (a) rehabilitation studies; (b) presented data that did not provide quantitative results to calculate effect size (ES); and (c) used a single group pretest and posttest.

Unrelated studies can be effectively eliminated by scanning the titles and abstracts, such as papers on AR-based EFL learning, papers in the field of rehabilitation, or papers using other languages as foreign languages. This stage reduced the list of 1057 sources to 222. Subsequently, the author read each full text of the remaining articles to assess their eligibility for meta-analysis. Finally, a total of 14 articles met all inclusion criteria and none of the exclusion criteria. All inclusion and exclusion procedures and results are shown in Fig. 1.

2.3 Moderator coding procedures

In addition to the research framework based on this study, we also refer to Zou et al. (2018) technology-based learning (TBL) framework and Godwin-Jones's (2021) complex dynamic models of language learning, with views on technology-assisted language learning based on the technical and linguistic perspectives, respectively. Finally, this study establishes its own coding scheme.

Meta-analysis requires extracting the key information from the literature and coding its characteristics in order to transform the descriptive data into quantitative data. Two trained coders independently coded all 14 studies to reduce researcher subjectivity. While coding, the coders also referred to some frameworks of previous articles. They then discussed their disagreements and reached consensus on the coding content. The following nine different codes were finally agreed upon. And the classification of the codes is shown in Table 2.

2.3.1 Country/region

This perspective was coded based on the countries or areas where the selected studies were conducted. Six countries and regions were included.

2.3.2 Educational levels

The selected studies were coded based on the educational levels of the participants, including elementary school, middle school, high school, and college.

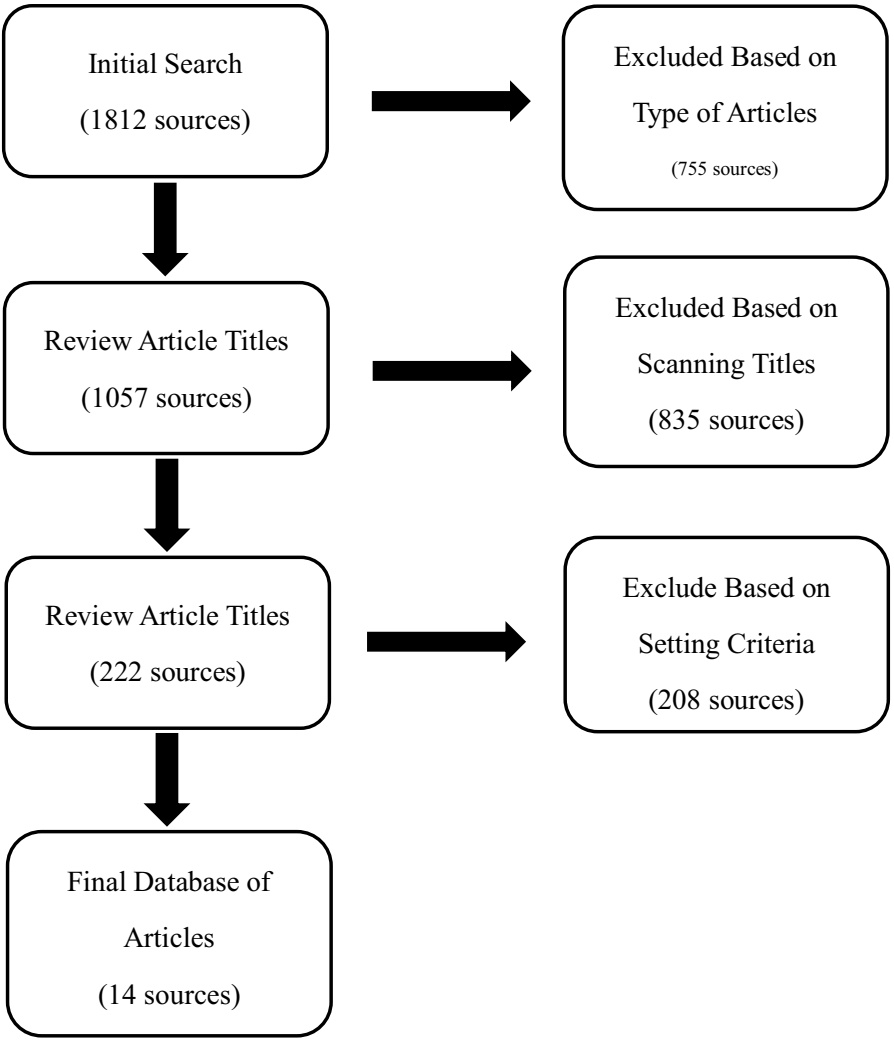

Fig. 1 Inclusion Flow Chart

2.3.3 Learning theories

This dimension was coded according to the learning theories used in the included studies. It includes atheoretical approaches, constructivism, embodied cognition, situated learning theory and social learning theory.

2.3.4 Duration of treatment

The duration and timing of each study varied. For convenient statistical analysis, we divided this variable into four categories: less than 1 month, between 1 and 3 months, more than 3 months, and not reported.

Table 2 Code table

Basis of coding	Coded objects	Code content
Learners' background	Region	Country/area
	Age	Educational levels
Experimental design	Learning theory	Learning theories
	Training duration	Treatment duration
	VR immersion	Experimental group VR type
	Learning activities	Learning activities
Teaching/ Learning resources	Source of materials	Learning materials
	Gamification	Game element
Learning gains	Linguistic gains	Learning outcomes
	Affective gains	

2.3.5 VR type of experimental Group

We recorded the VR type of the experimental group. The types of VR used in these studies were immersive and non-immersive.

2.3.6 Learning activities

According to the different learning activities, this code was categorized as cooperative learning, teacher instruction, and personal learning.

2.3.7 Learning materials

The learning materials that researchers provided to students varied. We divided them into two categories: previously existing or self-designed. If the researchers used software that already existed in the market (such as SL or some VR apps in Steam), we coded it as previously existing. If the learning materials were designed by the researchers (such as a 360° panoramic video), we coded them as self-designed.

2.3.8 Game elements

If the authors directly stated that the learning materials were a game or game-like training, or if we found that they appeared to contain multiple game elements, we coded them as "gamified". If the authors did not identify the learning materials as a game or game-like training, or if they contained no or very few game elements, we coded them as "not gamified".

2.3.9 Learning outcomes

Since each experiment had its own research design and purpose, the research findings differed based on different aspects of students' EFL learning. Therefore, we summarize seven aspects of learning outcomes.

2.4 Analysis methods

In order to determine the effectiveness of VR in improving EFL learning, the meta-analysis was conducted as follows: (a) effect size calculation, (b) homogeneity analysis, and (c) moderator analysis.

For the selected articles with multiple measures of more than one outcome or more than one experimental or control condition, the following rules were applied in calculating the effect size: (a) if the measurement of different learning outcomes was reported, one ES was extracted for each dependent variable; (b) if there was more than one experimental or control group (e.g., comparison between video, PPT, and VR conditions in one study), one ES was extracted for each comparison between VR-based EFL learning and one type of traditional EFL learning treatment. That is, an article may contain multiple samples. Therefore, the total number of samples, denoted as k (k = 23), is greater than the number of articles we selected (Sumak et al., 2011). The ES for each individual study was estimated based on the sample size, mean, and standard deviation. We used comprehensive meta-analysis software (version 3.0) to calculate the ES. Because the sample size of this study is small, Cohen's d may overestimate the ES. Finally, we chose Hedges' g (hereafter referred to as g) (Hedges, 1981) as the ES for continuous variables because it can correct for small sample bias. To account for the variability in research design and to be more likely to fit the actual sampling distribution, we chose the random effects model (REM) to calculate the overall mean ES across studies (Borenstein et al., 2010). According to Cohen's rule of thumb, ES values of 0.2, 0.5, and 0.8 were considered small, medium, and large, respectively (Cohen, 1992; Cohen, 1988).

We performed sensitivity analyses to examine the variance in ESs across studies. We calculated the Q statistic to test whether all studies in the analysis had a common effect size. Homogeneity was also assessed using I2 (I-squared, an indicator of heterogeneity), which describes the percentage of variation across studies that is due to true heterogeneity rather than sampling error (Higgins & Thompson, 2002). Based on the rule of thumb proposed by Higgins and Thompson, the cutoff values of I2 for low, moderate, and high heterogeneity were 0.25, 0.50, and 0.75, respectively.

Because publication bias may be present, a funnel plot was generated to examine the distribution of effect sizes. We also used the fail-safe N to estimate the number of undiscovered null studies that would need to be discovered before the observed effect would no longer be statistically significant (Van Wijk et al., 2008). Thus, if the null power N is sufficiently larger than our sample size, it can be assumed that the results would not be significantly affected by the discovery of additional studies. Egger's test and Begg's rank test were also used to test for publication bias.

3 Results

3.1 Outliers and publication bias

We first examined the funnel plot. The X-axis of the funnel plot represents the ES of each sample, and the Y-axis represents the standard error of each sample. At

this stage, we found one ES that was significantly different from most studies on the funnel plot, which is shown in the lower right corner of Fig. 2. In a meta-analysis, the inclusion of an outlier value may compromise the overall quality of the data (McGrath et al., 2020). We then further verified this using a sensitivity analysis method. We performed statistical analysis with and without this data point and found a large difference between the results of the two analyses, so the data point was identified as an outlier and we chose to exclude it (Aguinis et al., 2013).

After removing the outlier, Fig. 3 shows that the funnel plot is basically symmetrical, with most samples falling inside the funnel. Although there were two samples with a large ES of almost g = 2, the publication bias is not significant from the

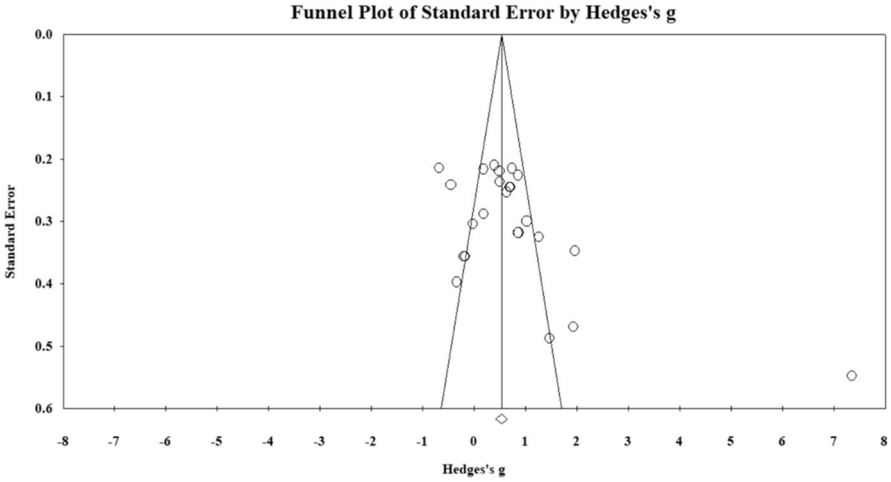

Fig. 2 Funnel plot with the outlier

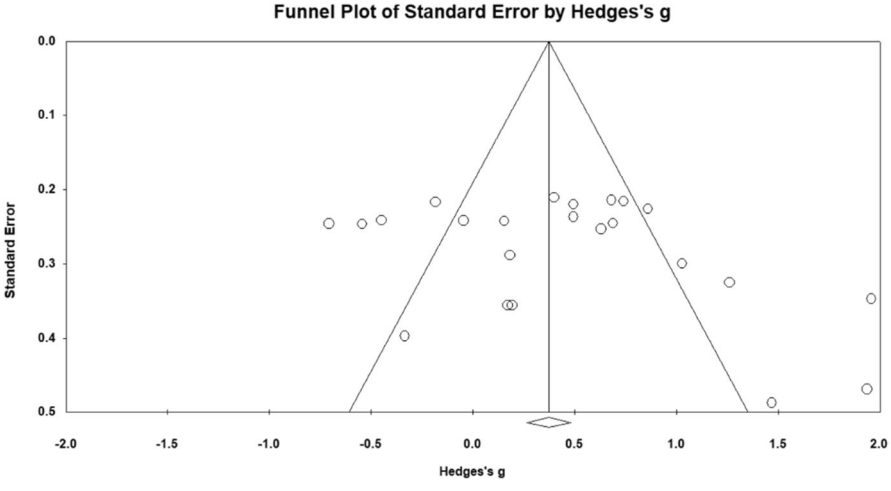

Fig. 3 Funnel plot after excluding the outlier

graph. The overall fail-safe N was 307, which means that 307 unreported null effects would have to be detected for the overall result to become insignificant. According to a previous study (Rosenthal, 1995), the number of null or additional studies needed to nullify the overall effect sizes found in this meta-analysis was greater than 5 k + 10, and the observed significant effects were considered stable.

Next, Egger's test (p = 0.132) was not statistically significant, indicating that there was no evidence of publication bias. Begg's rank correlation test (p = 0.316) was also not significant. Therefore, we believe that the results indicate the absence of publication bias.

3.2 Overall effect sizes

We calculated the ES of each sample and the overall mean ES of all 23 samples. Figure 4 shows the point estimates of the ESs and confidence intervals in forest plot format. In terms of independent ESs, the values ranged from -0.706 to 1.959, with 17 (74%) positive effects in favor of the VR-based EFL learning method over other traditional EFL interventions. In terms of overall ES, VR-based EFL learning has a small effect (g = 0.445, SE = 0.131, 95% CI = [0.188,0.702], p = 0.001). The Q-value was 121.556 with p = 0.000, indicating that there were significant differences among the ESs which were due to some other factors or not a sampling error. The I2 was 81.901, indicating high heterogeneity, indicating that there was strong heterogeneity between samples and that our choice of random effects model was correct. Among the ESs of studies with substantial heterogeneity, the random effect model can avoid assigning too much weight to large samples so that the ES can be closer to the actual values (Sharifi et al., 2017; Xu et al., 2018). The heterogeneity of studies indicated that further grouping of individual ESs was needed to search for potential moderators that could explain the variance across studies.

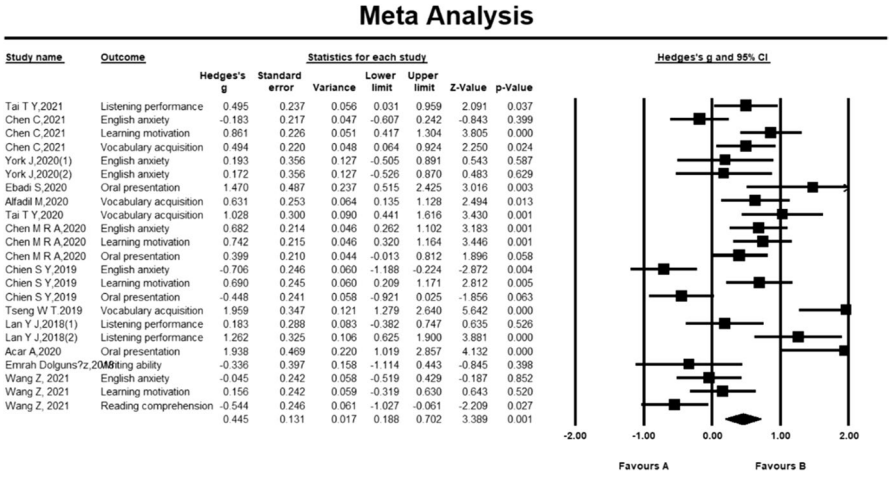

Fig. 4 Forest plot of the selected studies

Table 3 Moderator analysis

Category	*k*	*g*	*SE*	*95% CI*	*Q*	*P*
Country/area					**13.707**	**0.018**
Taiwan (China)	14	0.510	0.164	[0.188,0.833]		
Mainland of China	3	-0.142	0.207	[-0.548,0.263]		
Iran	1	1.470	0.487	[0.515,2.425]		
Saudi Arabia	1	0.631	0.253	[0.135,1.128]		
Turkey	2	0.788	1.137	[-1.440,3.015]		
Japan	2	0.183	0.252	[-0.311,0.676]		
Educational stages					**8.208**	**0.042**
Primary school	3	1.121	0.527	[0.089,2.154]		
Middle school	4	0.911	0.252	[0.417,1.406]		
High school	3	-0.155	0.427	[-0.992,0.482]		
College	13	0.296	0.135	[0.032,0.561]		
Treatment duration					**2.173**	**0.337**
≤ 1 month	12	0.366	0.224	[-0.073,0.805]		
1–3 months	6	0.700	0.159	[0.388,1.013]		
> 3 month	3	0.388	0.306	[-0.212,0.988]		
Not reported	2	0.183	0.252	[-0.311,0.676]		
Experimental group VR type					**2.039**	**0.153**
Immersive	20	0.347	0.129	[0.094,0.599]		
Non-immersive	3	1.121	0.527	[0.089,2.154]		
Learning theories					**6.981**	**0.137**
Atheoretical approaches	14	0.463	0.175	[0.119,0.806]		
Constructivism	1	1.535	0.509	[0.538,2.532]		
Embodied cognition	2	0.725	0.550	[-0.352,1.802]		
Situated learning theory	3	0.610	0.124	[0.366,0.853]		
Social learning theory	3	-0.157	0.432	[-1.003,0.69]		
Learning activities					**2.173**	**0.337**
Personal learning	12	0.268	0.201	[-0.126,0.663]		
Teacher instruction	8	0.659	0.174	[0.318,1.000]		
Cooperative learning	3	0.545	0.379	[-0.198,1.287]		
Learning materials					**0.827**	**1.363**
Previously existing	13	0.552	0.176	[0.207,0.896]		
Self-designed	10	0.305	0.206	[-0.099,0.710]		
Game elements					**0.024**	**0.878**
Gamified	6	0.481	0.132	[0.233,0.739]		
Not gamified	17	0.448	0.169	[0.117,0.779]		
Learning outcomes					**28.033**	**0.000**
Listening performance	3	0.624	0.292	[0.053,1.196]		
English anxiety	6	0.012	0.212	[-0.403,0.427]		
Vocabulary acquisition	4	0.986	0.296	[0.406,1.566]		
Oral presentation	4	0.760	0.481	[-0.182,1.703]		
Writing ability	1	-0.336	0.397	[-1.114,0.443]		
Learning motivation	4	0.621	0.153	[0.322,0.920]		
Reading comprehension	1	-0.544	0.246	[-1.027, -0.061]		

k = number of independent studies; g = mean effect size; CI = confidence interval; Q = between-group homogeneity; SE = standard error

3.3 Moderator analysis

The moderator analysis results are shown in Table 3.

The Q-statistics revealed a significant variance in ES according to the educational level of the participants as a moderator variable (Q = 8.208, p < 0.05). On average, the effect has significant differences at different levels of education, as judged by the 95% CIs. Elementary school (g = 1.121, 95% CI [0.089,2.154], k = 3) and middle school (g = 0.911, 95% CI [0.417,1.406], k = 4) have significant effects, college (g = 0.296, 95% CI [0.032,0.561], k = 13) has a small effect, and high school (g = -0.155, 95% CI [-0.992,0.482], k = 3) has a negative effect.

For the country/area moderator, the Q-statistics revealed significant variance in ES (Q = 13.707, p < 0.05). This means that judging by the 95% CI, the effects of VR-based EFL learning have significant variance across countries and regions: Iran (g = 1.470, 95% CI [0.515, 2.425], k = 1) had a large effect; Turkey (g = 0.788, 95% CI [-1.440, 3.015], k = 2), Taiwan (China) (g = 0.510, 95% CI [0.188, 0.833], k = 14), and Saudi Arabia (g = 0. 631, 95% CI [0.135,1.128], k = 1) had moderate effects, and Japan (g = 0.183, 95% CI [-0.311,0.676], k = 2) and mainland China (g = -0.142, 95% CI [-0.548,0.263], k = 3) had no apparent effects.

Regarding the experimental results, the Q-statistics showed a significant variance in ES (Q = 28.033, p < 0.05). Judging from the 95% CIs, on average, VR-based EFL learning had significant effects on students' vocabulary acquisition (g = 0.986, 95% CI [0.406,1.566], k = 4). VR-based EFL learning had medium effects on oral presentation (g = 0.760, 95% CI [-0.182,1.703], k = 4), listening comprehension (g = 0.624, 95% CI [0.053,1.196], k = 3), and learning motivation (g = 0.621, 95% CI [0.322,0.920], k = 4). However, for writing ability (g = -0.336, 95% CI [-1.114, 0.443], k = 1), reading comprehension (g = -0.544, 95% CI [-1.027, -0.061], k = 1), and English anxiety (g = 0.012, 95% CI [-0.403, 0.427], k = 6), VR-based EFL learning did not seem to have any obvious effects.

4 Discussion

4.1 Discussion of overall effect sizes

The primary meta-analysis in this study answers RQ1: VR-based EFL learning methods are more effective than traditional EFL learning approaches, with a medium ES. However, with 26% of the studies having a very small effect or a negative effect, we cannot exaggerate the advantage of VR-based EFL learning. At the current stage, we believe that the development of VR is not mature enough and the operational costs are relatively too high for VR to be an obvious replacement for traditional methods of EFL learning, as the Internet and mobile devices were once expected to be (Spector, 2013). From 2016 to 2018, the Horizon Reports mentioned that VR is one of the most important technologies that will be used in education in general in the future (Samantha Adams Becker et al., 2018; Becker et al., 2016;

Freeman et al., 2017; Lan, 2020). For now, however, we see it as a promising addition that can diversify EFL learning experiences and scenarios from traditional EFL learning.

4.2 Discussion of moderating variables

To answer RQ2 to RQ6, we conducted an analysis of the moderator variables.

4.2.1 Country/area

This study analyzed the regions in which studies were conducted. The Q-test showed a statistically significant difference ($p = 0.018$). Most of the samples were published in Taiwan (China) (61%), while others were published in mainland China (13%), Japan (9%), Iran (4%), Saudi Arabia (4%), and Turkey (9%). Iran ($g = 1.470$) has large effects, Saudi Arabia ($g = 0.631$), Turkey ($g = 0.788$), and Taiwan ($g = 0.510$) have moderate effects, and Japan ($g = 0.183$) and mainland China (-0.142) have no apparent effects. However, except for the Taiwan (China) sample, the sample size is relatively small, so these results may not be fully representative. We believe there are several reasons for this phenomenon. First, the results fully reflect that the Asian region pays much attention to EFL learning based on VR, while there are few studies in other regions; the reason may be the learning environment of Asian regions, which attach great importance to EFL learning. (Lee et al., 2020; Vonkova et al., 2021). Second, other regions have many native English-speaking countries, so they have less demand for EFL learning. Third, some other regions are still underdeveloped and VR equipment is too expensive for them. When we reviewed the articles, we found that many VR-based studies promoted other kinds of foreign language learning in regions where there is no need for EFL learning. Therefore, we assume that the experimental design of this study led to this result. Of course, we urge other non-English-speaking countries to conduct similar research. In addition, given the large gap in the effects of experiments in different regions, we also suggest that researchers should conduct more academic exchanges.

4.2.2 Educational levels

The participants in the selected studies covered four educational levels: elementary school (13%), middle school (17%), high school (13%), and college (57%). From this moderating factor, we know that among the samples, college students were the most common participants. From the perspective of experimental implementation, this may be because most researchers work in colleges or higher education institutions, so college students are the easiest participants to obtain. Compared to students at other levels of education, college students can also more easily cooperate with researchers to complete the experiment. This result is similar to the findings of a previous study (Huang et al., 2021). However, VR-based EFL learning has the least impact on high school students ($g = 0.155$), with only a negative ES, while elementary school students have the greatest impact ($g = 1.121$). Middle school students

(g = 0.911) and college students (g = 0.296) have a large and small ES, respectively. The difference between the groups reached a significant level (p = 0.042). This result answers RQ2 and indicates that VR plays a better role in promoting EFL learning for middle school and elementary school students than for high school and college students.

From the data, ES generally decreases with age, which is not consistent with the findings of existing related studies (Chen et al., 2022). This may be because younger students have less exposure to VR, and VR can provide highly immersive multisensory stimuli that can capture their attention and produce better results. However, even if there are fewer effects at the college level, it does not necessarily mean that the application of VR-based EFL learning should not be extended to this level of learners. Schools can also use VR for teaching, which can improve instructional design and situational, practical, and interesting aspects of EFL learning. VR can also be applied to students' informal learning or after-school learning to help increase their interest in EFL learning and reduce their anxiety about EFL learning.

4.2.3 Learning theories

Although not all studies are based on specific learning theories to facilitate EFL learning, we still sorted through all the articles. Quantitatively, even most of the studies used atheoretical approaches (g = 0.463) and were not based on a specific theory. Studies that used atheoretical approaches had ESs. Other studies used constructivism theory (g = 1.535), embodied cognition theory (g = 0.725), situated learning theory (g = 0.610), and social learning theory (g = -0.157). This also answered RQ3, which used constructivism theory to produce the best results for the study and was a large ES. Embodied cognition theory and situated learning theory are medium ESs, while social learning theory is a negative ES. This code did not reach a statistically significant difference. This result reflects the fact that most studies did not integrate VR-based EFL learning with relevant theories, but the learning effects did not differ as a result. VR-based learning can integrate many learning theories, each with its own strengths. It is not easy to say that EFL learning based on a particular learning theory is better. Meanwhile, the sample sizes of the studies that integrated learning theories were small compared to those based on mathematical methods, and their effects on VR-based EFL learning need to be confirmed by more studies in the future to validate the effect values. The studies using atheoretical approaches are more extensive, but the overall effect sizes are small and can be combined with other experimental designs to seek better facilitation effects.

4.2.4 Treatment duration

In this moderator, durations less than 1 month (g = 0.366) and greater than 3 months (g = 0.388) have small ESs, durations between 1 and 3 months (g = 0.700) have a medium ES, and the duration not reported (g = 0.183) has no apparent effect. We found that although the treatment duration was different, the ES between the groups did not show a difference and did not reach a statistically significant level. This

is different from the discovery of the novelty effect in previous studies (Makransky et al., 2019), which means that novelty did not have a positive effect on students' EFL learning. This may be due to the warm-up course conducted by teachers to familiarize students with the devices before class, or because students are in the information age, and it is characteristic of this age that similar technologies and products are widely used (Makransky & Lilleholt, 2018). This finding is also similar to the results of a previous study (Wu et al., 2020). In addition, different treatment durations lead to similar treatment effects, which supports another conclusion that long-term use of VR may lead to fatigue or discomfort and reduce the learning effect (Chen et al., 2019). Therefore, researchers may not need to intentionally increase the duration of VR-based EFL learning.

4.2.5 Experimental group VR type

We compared the effect of different experimental group VR treatments. Although the immersive (87%) type of VR has been used in broader and deeper language learning studies, the results show that this type has only a small effect ($g = 0.347$) on students' EFL learning. In contrast, the experimental group based on non-immersive VR showed a large effect ($g = 1.121$) on students' EFL learning. The differences between them did not reach a statistically significant level. As non-immersive VR has been developed earlier and is more mature and still dominant in language learning studies, we think this result is also reasonable (Qiu et al., 2021). On the other hand, we found that the three samples using non-immersive VR all used students at the primary school level, so this result may be somewhat random. However, the ES of the immersive group was not satisfactory. This may be because the development of immersive VR is still not mature enough, and the application in the classroom is not convenient enough. The equipment needed for EFL learning based on non-immersive VR is inexpensive, easy to obtain, and easy to manage in the classroom, so it can play a greater role in EFL learning at the moment. We predict that immersive VR will be used much more widely and deeply in EFL learning studies in the near future. At present, however, it may be better to choose non-immersive VR for EFL learning.

4.2.6 Learning activities

Among the learning activities, the Q-test did not show statistical significance. Cooperative learning (13%) and teacher instruction (35%) have medium effects, and personal learning (52%) has a small effect. The largest ES size was found for teacher instruction ($g = 0.659$), followed by cooperative learning ($g = 0.545$), and personal learning ($g = 0.268$). The effect of personal learning was slightly smaller, which means that students need more supervision and coaching from teachers in VR-based EFL learning. This result answers RQ4 and indicates that the method of learning activities does not have a great impact on students' learning effect. Teachers can choose the method of learning activities according to their own or students' wishes. However, according to a previous study, it may be difficult

for young children to participate in personal learning, so it is suggested to adopt teacher guidance for primary school students (Vitta & Al-Hoorie, 2020).

4.2.7 Learning material

We categorized the samples into two types of learning materials: self-designed (43%) and previously existing (57%). The number of samples for the two variables was nearly equal, but their effects were different. The moderator of the previously existing group had a medium ES (g = 0.552), and the moderator of the self-designed group had a small ES (g = 0.305), with no significant differences at the 95% confidence level. This result addresses RQ5. In the included studies, the previously existing learning materials (such as SL) and the self-designed learning materials (such as a 360° panoramic video) used by the researchers have their own advantages and disadvantages (Repetto et al., 2021a, b). The existing materials are easy to obtain and can be quickly incorporated into the experiment. However, most of them are universal and cannot be used specifically to train students' skills in a certain aspect. At the same time, to attract more users, learning materials need high interactivity, aesthetics and entertainment, but it is difficult to consider the educational goals. The self-designed materials have a significant purpose and can help students to train a certain skill. At the same time, there is no irrelevant content (such as advertisements) in the materials, which may have a negative impact on learning. For example, the learning materials of Acar's (2020) experiment are self-designed, which is highly professional and educational, but has poor interactivity (less content for students to interact with). Thus, ensuring the quality of the materials and improving the user experience for students is a challenge for researchers.

Nevertheless, we believe that the materials designed by researchers themselves deserve further research and development. Researchers can consider spatial awareness, collaboration, and cognitive processes to develop appropriate VR learning materials that most effectively stimulate the development of cognitive processes (Horváth, 2019). Personalized learning materials may be more appropriate for teaching, just as different teachers have their own teaching materials for the same class. However, self-designed learning materials require researchers to iterate and optimize to improve their quality, which is a very time-consuming task. We suggest that researchers also learn from other fields, such as medicine and engineering, so that they can put forward their own requirements and adapt the corresponding learning materials from teams in other professions.

4.2.8 Game elements

We tested two aspects of the learning materials: whether they contained game elements or not. The results show that the non-gamified samples (74%) are much more common than the gamified samples (26%), but their effects are very similar, meaning that they both have a small effect. The non-gamified (g = 0.448) and gamified (g = 0.481) samples did not reach a statistically significant difference. In

general, game elements can attract participants' attention and improve their learning interest, but we doubt whether game elements can promote students' EFL learning effectiveness. On the other hand, this conclusion supports some studies that found that the pedagogical benefits of game elements have not yet been scientifically confirmed (Dichev & Dicheva, 2017). Moreover, it is possible that the potential of game elements has not been fully exploited in current learning materials. Therefore, researchers should carefully select learning materials that do or do not include game elements.

4.2.9 Learning outcomes

Table 3 shows that there is a significant difference in the scores (p = 0.000). Among these outcomes, vocabulary acquisition (17%, g = 0.986) has a large effect. Motivation to learn (17%, g = 0.621), oral presentation (17%, g = 0.760), and listening comprehension (13%, g = 0.624) have medium effects. English anxiety (26%, g = 0.012) has no apparent effect, but reading comprehension (4%, g = -0.544) and writing ability (4%, g = -0.336) have no positive effects. Among these studies, the sample sizes used to measure reading comprehension and writing ability were relatively small, highlighting the need for future research to fill these gaps.

We observed that the ES of writing ability is negative, which is consistent with previous research findings (Huang et al., 2021); we believe that this finding may be coincidental due to the small sample size. We also speculate that the inability to improve writing ability may be because writing requires more training in the real world, rather than being improved only in the VR world simulated by audiovisual sensory input. Therefore, it is difficult for the ES of writing ability to outperform other learning outcomes. This finding is also consistent with the results of a previous study (Wang et al., 2020).

We also observed that the ES of reading comprehension was negative. One possible reason is that traditional VR is based on visual stimulation, such as images or videos, or not based on words, which is not helpful for training reading comprehension. On the other hand, given the small sample size, this effect may not be representative. In this included paper that tested reading comprehension, Wang et al., (2021) also improved traditional VR into visual prompt scaffolding-based VR (VPS-VR). He found that compared with the traditional VR group and the control group, the use of VPS-VR can better promote students' reading comprehension. At the same time, the VPS-VR approach still has positive effects on students' EFL learning motivation and English learning anxiety. However, VPS strategies in the current study can only improve the lower level of reading comprehension skills. This conclusion is consistent with another included study, which found that the VR learning approach did not significantly affect all aspects of students' learning behaviors (Chien et al., 2020a, b). Therefore, other strategies to improve higher level reading comprehension skills need to be investigated. In future research, researchers can try to use VPS-VR or develop more instructional strategies to improve EFL students' reading comprehension.

Regarding the ES of the English anxiety sample, although VR-based EFL learning has a small positive effect on reducing English anxiety, and students also report that they are very happy while learning, this effect seems to be limited. York et al. (2020) showed that the avatar effect or anonymity did not seem to play an important

role in improving participants' positive perceptions. Therefore, we believe that VR can be used as a complementary method for students to reduce English anxiety.

Students' learning motivation, listening performance, oral presentation, and vocabulary acquisition were significantly improved. The VR environment immerses students and allows them to experience good practice in these aspects. These three aspects cooperate with each other, which can promote students' EFL learning. RQ6 is also answered here, that is, vocabulary acquisition is the most helpful learning outcome for students in this study. This conclusion is also similar to that of a previous study (Hao et al., 2021).

5 Conclusions and suggestions

5.1 Research conclusion

In this study, we conducted a meta-analysis of studies on the use of VR in EFL learning. The results show that the VR-based EFL learning method has a better overall effect than traditional EFL learning methods, which contributes to the research on VR in EFL learning. By analyzing the characteristics of the studies as moderator variables, we found that VR has been widely used in EFL learning. The application of non-immersive VR is more mature, but immersive VR is also worth promoting. We also found that the effects of VR in arousing situational interest may not be as large as we expected, and longer VR exposure may not produce better results. There are no noticeably different effects of using different materials on students; however, we still think that the researchers who designed the materials and used them in their experiments are worth encouraging. We also found that VR-based EFL learning has a good effect on improving students' vocabulary acquisition, oral presentation, learning motivation, and listening comprehension. Researchers can apply these aspects to the classroom or further explore the other aspects with poor improvement effects.

For the related research area, this paper further focuses on EFL learning based on previous studies. For EFL researchers, more intuitive research results can be obtained. With the development of time and technology, the development of immersive VR is progressing rapidly, and the comparison of immersive VR and non-immersive VR never stops. We continue the previous study by once again comparing two different kinds of VR, and two groups still do not show significant differences. In addition, this study examined for the first time the effect of VR learning materials on student learning outcomes. Although the effects did not differ between materials, we believe it is still worth exploring.

5.2 Limitations and suggestions

However, there are limitations to this meta-analysis. Because the language of this study was limited to English, the sample size was small, which may have affected the results. In addition, the moderator analysis may be more or less subjective due to

our personal judgment of some moderator variables, such as learning activities (Lin & Lin, 2019). For example, the classification of learning activities is not detailed enough. A more detailed classification, such as task-based instruction, can be done in later studies (Wehner et al., 2011).

At the same time, we also found some problems in the existing articles. First, researchers did not test how VR affects language learning with some key issues, such as VR scene effects (the level of immersion and presence), higher-order cognition assessment (high-level skills and cooperation), and students' satisfaction in VR use. Secondly, VR has no standardized application model in EFL learning and has obstacles in its application and promotion. Therefore, we call on the academic community to conduct relevant research and standardize the use of VR. At the same time, there is also a lack of focused research and clear teaching programs in the development of VR learning materials and resources. This indicates that the integration of VR into education requires long-term efforts. Only by conducting multiple studies will it be possible to gradually build a consensus and ensure the effective development of VR in EFL research.

Therefore, in the future, researchers need to focus on developing high-quality VR materials that focus not only on enhancing students' development of higher-order skills, but also on adapting VR technology to all aspects of EFL teaching. In addition, many authors assume that VR can produce beneficial effects and include atheoretical approaches in their study. That is, many do not use a specific theory to justify its use. Therefore, we call for future research in which researchers can integrate more relevant theory into the application and in-depth study of VR-based EFL learning. Some of the included studies were developed in combination with some learning strategies or learning methods that we did not explore in depth and could be explored in more detail by subsequent researchers. Accordingly, the use of VR in EFL learning has not yet to become a mature field. This paper also encourages researchers to explore more effective ways of implementing VR-based EFL learning in the future. Future research should further explore these elements, and the current meta-analysis is only a small step in this area of research.

Author contributions All authors contributed to this study. More information can be given when needed.

Funding This study was supported in part by Zhejiang province Soft Science Research Project (No.2023C35027), the Collaborative Education Project of the Ministry of Education of P. R. China (No. 202002234058), and the Key Laboratory of Brain Science and Children's Learning of Huzhou, Huzhou University, Huzhou.

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval The study comply with Ethical Standards.

Consent to participate The participants all agreed to take part in this study.

Consent for publication The publication of this study has been approved by all authors.

Conflicts of interests/Competing interests The authors declare that they have no conflicts of interests.

References

Acar, A., & Cavas, B. (2020). The Effect of virtual reality enhanced learning environment on the 7th-grade students' reading and writing skills in English. *Malaysian Online Journal of Educational Sciences, 8*(4), 22–33.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods, 16*(2), 270–301.

Almusharraf, N. (2021). Incorporation of a game-based approach into the EFL online classrooms: students' perceptions. *Interactive Learning Environments* (2), 1–14. https://doi.org/10.1080/10494820.2021.1969953

Aslan, E., & Thompson, A. S. (2021). The interplay between learner beliefs and foreign language anxiety: Insights from the Turkish EFL context. *Language Learning Journal, 49*(2), 189–202. https://doi.org/10.1080/09571736.2018.1540649

Atkinson, D. (2002). Toward a sociocognitive approach to second Llanguage acquisition. *The Modern Language Journal, 86*(4), 525.

Bai, Y. (2018). On modern computer assisted language learning facilities and its integrated teaching. *International Journal of Emerging Technologies in Learning (online), 13*(11), 225.

Bamodu, O., & Ye, X. (2013). *Virtual reality and virtual reality system components* (Vol. 765, pp. 1169–1172): Trans Tech Publications Ltd.

Bankowski, E. (2010). Developing skills for effective academic presentations in EAP. *International Journal of Teaching and Learning in Higher Education, 22*(2), 187–196.

Barrett, M., Blackledge, J., & Coyle, E. (2011). Using virtual reality to enhance electrical safety and design in the built environment. *ISAST Transactions on Computers and Intelligent Systems, 3*(1), 1–9.

Becker, S. A., Brown, M., Dahlstrom, E., Davis, A., DePaul, K., Diaz, V., & Pomerantz, J. (2018). *NMC horizon report: 2018 higher* (education). Educase.

Becker, S. A., Freeman, A., Hall, C. G., Cummins, M., & Yuhnke, B. (2016). *NMC/CoSN horizon report: 2016 K-12 edition*. Austin, Texas: The New Media Consortium.

Berns, A., & Reyes-Sanchez, S. (2021). A review of virtual reality-based language learning apps. *Ried-Revista Iberoamericana De Educacion A Distancia, 24*(1), 159–177. https://doi.org/10.5944/ried.24.1.27486

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97–111. https://doi.org/10.1002/jrsm.12

Bridges, C., Hummel, J., Hursthouse, J., Moss, R., Antilla, T., Book, B., . . . Farmer, R. (2007). A Cross-Industry Public Foresight Project.

Chan, V. (2011). Teaching oral communication in undergraduate science: Are we doing enough and doing it right? *Journal of Learning Design, 4*(3), 71–79.

Chang, C., Bang, K., Wetzstein, G., Lee, B., & Gao, L. (2020). Toward the next-generation VR/AR optics: A review of holographic near-eye displays from a human-centric perspective. *Optica, 7*(11), 1563–1578. https://doi.org/10.1364/optica.406004

Chen, B., Wang, Y., & Wang, L. (2022). The effects of virtual reality-assisted language learning: A meta-analysis. *Sustainability, 14*(6), 3147. https://doi.org/10.3390/su14063147

Chen, M. R. A., & Hwang, G. J. (2020a). Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2020.1734626

Chen, X., Chen, Z., Li, Y., He, T., Hou, J., Liu, S., & He, Y. (2019). ImmerTai: Immersive motion learning in VR environments. *Journal of Visual Communication and Image Representation, 58*, 416–427. https://doi.org/10.1016/j.jvcir.2018.11.039

Cheng, Y. W., Wang, Y., Yang, Y. F., Yang, Z. K., & Chen, N. S. (2020). Designing an authoring system of robots and IoT-based toys for EFL teaching and learning. *Computer Assisted Language Learning*(1), 1–29. https://doi.org/10.1080/09588221.2020.1799823

Chien, S.-Y., Hwang, G.-J., & Jong, M.S.-Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education, 146*, 103751. https://doi.org/10.1016/j.comped-u.2019.103751

Christoforou, M., Xerou, E., & Papadima-Sophocleous, S. (2019). *Integrating a Virtual Reality Application to Simulate Situated Learning Experiences in a Foreign Language Course*. Research-Publishing.

Coffey, A. J., Kamhawi, R., Fishwick, P., & Henderson, J. (2017). The efficacy of an immersive 3D virtual versus 2D web environment in intercultural sensitivity acquisition. *Etr&d-Educational Technology Research and Development, 65*(2), 455–479. https://doi.org/10.1007/s11423-017-9510-9

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112*(1), 155–159.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Journal of the American Statistical Association, 67*(3), 1007–1007.

Cooper, H. (2015). *Research synthesis and meta-analysis: A step-by-step approach* (Vol. 2). Sage publications.

Dichev, C., & Dicheva, D. (2017). Gamifying education: What is known, what is believed and what remains uncertain: A critical review. *International Journal of Educational Technology in Higher Education, 14*(1), 9. https://doi.org/10.1186/s41239-017-0042-5

Dolgunsöz, E., Yildirim, G., & Yildirim, S. (2018). The effect of virtual reality on EFL writing performance. *Journal of Language and Linguistic Studies, 14*(1), 278–292.

Donato, R., & Mccormick, D. (1994). A Sociocultural Perspective on Language Learning Strategies: The Role of Mediation. *Modern Language Journal, 78*(4), 453–464. https://doi.org/10.2307/32328584

Doris, U., Bolliger, Daniel, Mills, Jeremy, . . . Kohyama. (2015). Japanese Students' Perceptions of Digital Game Use for English-Language Learning in Higher Education. *Journal of Educational Computing Research, 53*(3), 384–408. https://doi.org/10.1177/0735633115600806

Ebadi, S., & Ebadijalal, M. (2020a). The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency. *Computer Assisted Language Learning* (3), 1–25. https://doi.org/10.1080/09588221.2020.1854311

Freeman, A., Becker, S. A., & Cummins, M. (2017). *NMC/CoSN horizon report: 2017 K–12 edition*. Austin, Texas: The New Media Consortium.

Genzola, A. E. (2016). Weblogs on Language Learning: A Technology-Enhanced Instruction in a Tertiary-Level EFL Classroom in China. *Social Science Electronic Publishing, 6*(4), 389–407. https://doi.org/10.2139/ssrn.2847467

Godwin-Jones, R. (2021). Evolving technologies for language learning. *Language Learning & Technology, 25*(3), 6–26.

Hansen, C., Steinmetz, H., & Block, J. (2022). How to conduct a meta-analysis in eight steps: A practical guide. *Management Review Quarterly, 72*(1), 1–19. https://doi.org/10.1007/s11301-021-00247-4

Hao, T., Wang, Z., & Ardasheva, Y. (2021). Technology-Assisted Vocabulary Learning for EFL Learners: A Meta-Analysis. *Journal of Research on Educational Effectiveness, 14*(3), 645–667. https://doi.org/10.1080/19345747.2021.1917028

Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect size and Related Estimators. *Journal of Educational Statistics, 6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Horváth, I. (2019). *How to develop excellent educational content for 3D VR.* Paper presented at the 2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom).

Howard, M. C., & Gutworth, M. B. (2020). A meta-analysis of virtual reality training programs for social skill development. *Computers & Education, 144*. https://doi.org/10.1016/j.compedu.2019.103707

Huang, X. Y., Zou, D., Cheng, G., & Xie, H. R. (2021). A Systematic Review of AR and VR Enhanced Language Learning. *Sustainability, 13*(9). https://doi.org/10.3390/su13094639

Jiang, W., & Eslami, Z. R. (2021). Effects of computer-mediated collaborative writing on individual EFL writing performance. *Computer Assisted Language Learning* (3), 1–30. https://doi.org/10.1080/09588221.2021.1893753

Krishnan, K. S. D., & Pathan, Z. H. (2013). Investigating demotivation in learning English: An extension to Sakai and Kikuchi's (2009) framework. *Advances in Language and Literary Studies, 4*(2), 124–131. https://doi.org/10.7575/aiac.alls.v.4n.2p.124

Lan, Y. J. (2015). Contextual EFL Learning in a 3D Virtual Environment. *Language Learning & Technology, 19*(2), 16–31.

Lan, Y. J. (2020). Immersion, interaction, and experience-oriented learning: Bringing virtual reality into FL learning. *Language Learning & Technology, 24*(1), 1–15.

Lee, S., Kuo, L. J., Xu, Z., & Hu, X. (2020). The effects of technology-integrated classroom instruction on K-12 English language learners' literacy development: A meta-analysis. *Computer Assisted Language Learning, 35*(5–6), 1106–1137. https://doi.org/10.1080/09588221.2020.1774612

Lin, J. J., & Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. *Computer Assisted Language Learning*, 1–42. https://doi.org/10.1080/09588221.2018.1541359

Lin, V., Liu, G. Z., & Chen, N. S. (2020). The effects of an augmented-reality ubiquitous writing application: a comparative pilot project for enhancing EFL writing instruction. *Computer Assisted Language Learning* (4), 1–42. https://doi.org/10.1080/09588221.2020.1770291

Lin, Y. J., & Wang, H. C. (2021). Using virtual reality to facilitate learners' creative self-efficacy and intrinsic motivation in an EFL classroom. *Education and Information Technologies* (6), 1–19. https://doi.org/10.1007/s10639-021-10472-9

Macedonia, M. (2014). Bringing back the body into the mind: gestures enhance word learning in foreign language. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01467

Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and Cognitive Benefits of Training in Immersive Virtual Reality Based on Multiple Assessments. *Journal of Computer Assisted Learning, 35*(6), 691–707. https://doi.org/10.1111/jcal.12375

Makransky, G., & Lilleholt, L. (2018). A Structural Equation Modeling Investigation of the Emotional Value of Immersive Virtual Reality in Education. *Educational Technology Research and Development, 66*(5), 1141–1164. https://doi.org/10.1007/s11423-018-9581-2

McGrath, S., Zhao, X., Steele, R., Thombs, B. D., Benedetti, A., & Collaboration, D. S. D. (2020). Estimating the sample mean and standard deviation from commonly reported quantiles in meta-analysis. *Statistical Methods in Medical Research, 29*(9), 2520–2537. https://doi.org/10.1177/0962280219889080

Pathan, Z. H., Ismail, S., & Fatima, I. (2020). English language learning demotivation among Pakistani university students: Do resilience and personality matter? *Journal of Applied Research in Higher Education*. https://doi.org/10.1108/jarhe-04-2020-0087

Paul, J., & Barari, M. (2022). Meta-analysis and traditional systematic literature reviews—What, why, when, where, and how? *Psychology & Marketing, 39*(6), 1099–1115. https://doi.org/10.1002/mar.21657

Pérez-Segura, J. J., Ruiz, R. S., González-Calero, J. A., & Cózar-Gutiérrez, R. (2020). The effect of personalized feedback on listening and reading skills in the learning of EFL. *Computer Assisted Language Learning* (1), 1–23. https://doi.org/10.1080/09588221.2019.1705354

Qiu, X. Y., Chiu, C. K., Zhao, L. L., Sun, C. F., & Chen, S. J. (2021). Trends in VR/AR technology-supporting language learning from 2008 to 2019: a research perspective. *Interactive Learning Environments* (3), 1–24. https://doi.org/10.1080/10494820.2021.1874999

Repetto, C., Di Natale, A. F., Villani, D., Triberti, S., Germagnoli, S., & Riva, G. (2021a). The use of immersive 360° videos for foreign language learning: a study on usage and efficacy among high-school students. *Interactive Learning Environments*, 1–16. https://doi.org/10.1080/10494820.2020.1863234

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*(2), 183–192.

Schneider, J., Borner, D., van Rosmalen, P., & Specht, M. (2017). Presentation Trainer: What experts and computers can tell about your nonverbal communication. *Journal of Computer Assisted Learning, 33*(2), 164–177. https://doi.org/10.1111/jcal.12175

Shadiev, R., & Yang, M. (2020). Review of Studies on Technology-Enhanced Language Learning and Teaching. *Sustainability, 12*. https://doi.org/10.3390/su12020524

Sharifi, M., Rostami Abusaeedi, A., Jafarigohar, M., & Zandi, B. (2017). Retrospect and prospect of computer assisted English language learning: a meta-analysis of the empirical literature.*Com- puter Assisted Language Learning*, 1–24. https://doi.org/10.1080/09588221.2017.1412325

Spector, J. M. (2013). Trends and Research Issues in Educational Technology. *Malaysian Online Journal of Educational Technology, 1*(3), 1–9.

Su, Y. C. (2006). EFL teachers' perceptions of English language policy at the elementary level in Taiwan. *Educational Studies, 32*(3), 265–283. https://doi.org/10.1080/03055690600631218

Sumak, B., Hericko, M., & Pusnik, M. (2011). A meta-analysis of e-learning technology acceptance: The role of user types and e-learning technology types. *Computers in Human Behavior, 27*(6), 2067–2077. https://doi.org/10.1016/j.chb.2011.08.005

Sutherland, I. E. (1968). *A head-mounted three dimensional display.* Paper presented at the Proceedings of the December 9–11, 1968, fall joint computer conference, part I (pp. 757–764).

Tai, T. Y., Chen, H. J., & Todd, G. (2020a). The impact of a virtual reality app on adolescent EFL leaseners' vocabulary learning. *Computer Assisted Language Learning, 35*(4), 892–917. https://doi.org/10.1080/09588221.2020.1752735

Tan, P. J. B. (2015). English e-learning in the virtual classroom and the factors that influence ESL (English as a Second Language): Taiwanese citizens' acceptance and use of the Modular Object-Oriented Dynamic Learning Environment. *Social Science Information Sur Les Sciences Sociales, 54*(2), 211–228. https://doi.org/10.1177/0539018414566670

Thompson, A. S., & Lee, J. (2013). Anxiety and EFL: Does multilingualism matter? *International Journal of Bilingual Education and Bilingualism, 16*(6), 730–749. https://doi.org/10.1080/13670050.2012.713322

Van Wijk, R., Jansen, J. J. P., & Lyles, M. A. (2008). Inter- and intra-organizational knowledge transfer: A meta-analytic review and assessment of its antecedents and consequences. *Journal of Management Studies, 45*(4), 830–853. https://doi.org/10.1111/j.1467-6486.2008.00771.x

Vitta, J. P., & Al-Hoorie, A. H. (2020). The flipped classroom in second language learning: A meta-analysis. *Language Teaching Research*. https://doi.org/10.1177/1362168820981403

Vonkova, H., Jones, J., Moore, A., Altinkalp, I., & Selcuk, H. (2021). A review of recent research in EFL motivation: Research trends, emerging methodologies, and diversity of researched populations. *System, 103*, 102622. https://doi.org/10.1016/j.system.2021.102622

Wang, C. P., Lan, Y. J., Tseng, W. T., Lin, Y. T. R., & Gupta, K. C. L. (2020). On the effects of 3D virtual worlds in language learning - a meta-analysis. *Computer Assisted Language Learning, 33*(8), 891–915. https://doi.org/10.1080/09588221.2019.1598444

Wang, Z., Guo, Y., Wang, Y., Tu, Y.-F., & Liu, C. (2021a). Technological Solutions for Sustainable Development: Effects of a Visual Prompt Scaffolding-Based Virtual Reality Approach on EFL Learners' Reading Comprehension, Learning Attitude, Motivation, and Anxiety. *Sustainability, 13*(24), 13977. https://doi.org/10.3390/su132413977

Wehner, A. K., Gump, A. W., & Downey, S. (2011). The effects of Second Life on the motivation of undergraduate students learning a foreign language. *Computer Assisted Language Learning, 24*(3), 277–289. https://doi.org/10.1080/09588221.2010.551757

Wu, B., Yu, X. X., & Gu, X. Q. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology, 51*(6), 1991–2005. https://doi.org/10.1111/bjet.13023

Xu, Z., Banerjee, M., Ramirez, G., Zhu, G., & Wijekumar, K. K. (2018). The effectiveness of educational technology applications on adult English language learners' writing quality: A meta-analysis. *Computer Assisted Language Learning, 32*(1–4), 132–162. https://doi.org/10.1080/09588221.2018.1501069

Yassin, A. A., Razak, N. A., & Maasum, T. (2019). Integrated Model for Teaching Language Skills. *International Journal of English Linguistics, 9*(5), 89–97. https://doi.org/10.5539/ijel.v9n5p89

Yin, K., He, Z. Q., Xiong, J. H., Zou, J. Y., Li, K., & Wu, S. T. (2021). Virtual reality and augmented reality displays: advances and future perspectives. *Journal of Physics-Photonics, 3* (2). https://doi.org/10.1088/2515-7647/abf02e

York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2020). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL, 33*(1), 1–22. https://doi.org/10.1017/S0958344020000154

Yu, L. T., Song, J. Y., & Chiu, F. Y. (2020). Using a three-dimension virtual world to reduce language anxiety and enhance english-speaking performance of EFL university learners: a collaborative project. *Taiwan Journal of Tesol, 17*(2), 65–89. https://doi.org/10.30397/tjtesol.202010_17(2).0003

Zou, D., Xie, H. R., & Wang, F. L. (2018). Future trends and research issues of technology-enhanced language learning: A technological perspective. *Knowledge Management & E-Learning-an International Journal, 10*(4), 426–440.

Further reading (References marked with an asterisk indicate studies included in the meta-analysis)

*Alfadil, M. (2020). Effectiveness of virtual reality game in foreign language vocabulary acquis-ition. *Computers & Education, 153*, 103893. https://doi.org/10.1016/j.compedu.2020.103893

*Chen, M. R. A., & Hwang, G. J. (2020b). Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles. *Interactive Learning Environments*, 1–21. https://doi.org/10.1080/10494820.2020.1734626

[*]Chen, C. H., Hung, H. T., & Yeh, H. C. (2021). Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of English language learners. *Journal of Computer Assisted Learning*, *37*(3), 851–860. https://doi.org/10.1111/jcal.12528

* Chien, S.-Y., Hwang, G.-J., & Jong, M. S.-Y. (2020b). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education, 146*, 103751. https://doi.org/10.1016/j.compe-du.2019.103751

*Ebadi, S., & Ebadijalal, M. (2020b). The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency. *Computer Assisted Language Learning*, 1–25. https://doi.org/10.1080/09588221.2020.1854311

*Lan, Y. J., Fang, W. C., Hsiao, I. Y., & Chen, N. S. (2018). Real body versus 3D avatar: The effects of different embodied learning types on EFL listening comprehension. Educational Technology Research and Development, 66(3), 709-731. https://doi.org/10.1007/s11423-018-9569-y

*Repetto, C., Di Natale, A. F., Villani, D., Triberti, S., Germagnoli, S., & Riva, G. (2021b). The use of immersive 360° videos for foreign language learning: a study on usage and efficacy among high-school students. *Interactive Learning Environments*, 1–16. https://doi.org/10.1080/10494820.2020.1863234

*Tai, T. Y., Chen, H. J., & Todd, G. (2020b). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning35(4)*, 892-917. https://doi.org/10.1080/09588221.2020.1752735

*Tai, T. Y., & Chen, H. H. J. (2021). The impact of immersive virtual reality on EFL learners' listening comprehension. *Journal of Educational Computing Research*, *59*(7), 1272-1293. https://doi.org/10.1177/0735633121994291

*Tseng, W. T., Liou, H. J., & Chu, H. C. (2020). Vocabulary learning in virtual environments: Learner autonomy and collaboration. *System, 88*, 102190

*York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2021). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL, 33*(1), 49-70. https://doi.org/10.1017/S0958344020000154