

Review

# The Effects of Virtual Reality-Assisted Language Learning: A Meta-Analysis

Bing Chen , Yunqing Wang and Lianghui Wang \* 

Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University, Jinhua 321004, China; chen\_bing@zjnu.edu.cn (B.C.); 202125200558@zjnu.edu.cn (Y.W.)

\* Correspondence: wlh@zjnu.cn; Tel.: +86-13566997797

**Abstract:** Existing literature reflects that VR technology is widely used in language learning settings. Although many studies have identified the multiple benefits and affordance of Virtual Reality (VR) technologies in language learning, most studies are qualitative studies that do not provide substantial evidence to investigate the impact of this technology on language learning. To this end, this study conducted a meta-analysis of 21 quantitative studies with 1144 participants published between 2010 and 2021. The study's main purpose was to examine the effects of VR on students' language learning academic performance, including linguistic gains and affective gains. The results indicated that VR-assisted language learning had a medium effect on the linguistic gains (Hedges'  $g = 0.662$ , 95% CI [0.398–0.925],  $p < 0.001$ ) and affective gains (Hedges'  $g = 0.570$ , 95% CI [0.309–0.831],  $p < 0.001$ ) of students compared to non-VR conditions, respectively. Furthermore, the study further analyzed the impact of several moderator variables such as education levels, hardware types, language skills, target language, and L1/L2 on language learning gains. The research indicates that VR technology has a great potential to improve language learning as an educational resource and provides suggestions for further research and practice on the use of VR-assisted language learning.

**Keywords:** meta-analysis; virtual reality; language learning; VR-assisted language learning



**Citation:** Chen, B.; Wang, Y.; Wang, L. The Effects of Virtual Reality-Assisted Language Learning: A Meta-Analysis. *Sustainability* **2022**, *14*, 3147. <https://doi.org/10.3390/su14063147>

Academic Editors: Rosabel Roig-Vila, Jordi M. Antolí-Martínez, Antonio Cortijo, Vicent Martines, Santiago Mengual Andrés, Elena Sánchez-López, Fabrizio Manuel Sirignano and Alexander López Padrón

Received: 27 January 2022

Accepted: 6 March 2022

Published: 8 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Language serves as a bridge between humans and society to communicate. With the growing trend of globalization, language education has also developed rapidly. For example, the recognized international language, English, has become a compulsory subject in many countries and regions [1]. However, language learning is a time-consuming and challenging process for many students as far as the current situation is concerned. The main reason for language learning difficulties is the lack of an authentic language environment, and learners cannot personally contact relevant contexts to use the target language to achieve learning goals [2]. That is, it is essential that language learners are provided an authentic learning environment and meaningful tasks [3]. The advancement of computer technology as a learning tool has provided new methods and created real-world environments to improve language learning [4]. As a novel technology, virtual reality (VR) has provided numerous alternative learning opportunities to language learners in the past decade [5].

VR technology and device appeared in the early 1960s, and many studies on VR and its applications have already been carried out in recent years [6]. Virtual Reality (VR) refers to a three-dimensional (3D) environment generated by computer technology, which can provide a context similar to visual simulation and other senses [7]. It allows users to communicate with people, machines, and other entities in the virtual environment by using computers and various devices [8,9]. There are two main types of virtual reality devices, namely immersive VR and non-immersive VR. Immersive VR involves high immersion and costs, such as cave automatic virtual environment (CAVE) and headset VR devices.

Non-immersive VR commonly refers to the desktop VR that interacts with computers through the mouse, control handle, and other devices [7,10]. The dramatic reduction in the cost of devices and technology has driven a rapid growth of VR applications in educational fields such as medicine, science, and mathematics in recent years that has been proven to be positive [11–13]. Learners feel the actual situation through sensory organs, which can help them improve their motivation, participation, and learning ability [5]. Moreover, VR has also been applied to language learning and has shown the importance and potential of applications to support language learning.

### *1.1. Applications of VR in Language Learning*

In recent years, a large number of studies have documented the application of technology-enhanced language learning. VR has also received a great deal of attention in this field. VR technology is characteristics of immersion, interaction, and involvement [14]. It breaks through the limitations of traditional media, provides language learners with a realistic simulated language learning environment, and effectively supports their language learning. For example, Wehner, Gump, and Downey [15] investigated the effect of Second Life on the motivation of English-speaking undergraduates to learn Spanish. They found that Second Life increased participants' motivation. Huang, Hwang, and Chang [7] developed a Chinese writing SVVR learning system based on spherical video-based virtual reality (SVVR) and compared the effects of this method with traditional technology-supported learning methods in high school writing classes. It was found that, compared to a control group that participated in traditional technology-supported learning classes, the SVVR writing improved students' writing performance and their self-efficacy while also reducing their cognitive load. For example, Tai and Chen [16] randomly assigned 72 middle school students to the experimental group (mobile headset virtual device) and control group (traditional multimedia) to investigate the effectiveness of mobile headset virtual devices for learners' English listening comprehension compared with video learning. The results showed that mobile headset virtual devices were significantly higher than video. In summary, these studies indicate that VR can promote language skills and foster motivation and self-efficacy.

### *1.2. Previous Review Studies on VR-Assisted Language Learning*

As far as we know, most of the existing literature [10,17–20] has described the systematic reviews of VR application in language learning through qualitative variables. These studies have shown that VR application in language learning is steadily growing and effectively used in language learning settings. Lin and Lan [17] conducted a content analysis of VR applications in language learning from 2004 to 2013, and they found that the research topics on VR-enhanced language learning were mainly interactive communication, behavior, affections, and task-based instruction. Solak and Erdem [18] also conducted a content analysis on 40 papers published from 1995 to 2015 on foreign language learning and teaching through virtual reality and identified that the data collection tool preferred was document analysis, and that the effectiveness of VR and game-based learning were the two prominent topics in this field. Huang, Zou, Cheng, and Xie [20] analyzed the VR on language learning from five perspectives: ways of integrating VR tools in language learning; primary users; major research findings; why it can effectively promote language learning's impact. Qiu et al. [10] conducted a systematic literature review that included 150 articles on VR/AR in language learning from 2008 to 2019. The study found that VR was most applied in higher education and that the most reported advantages of VR in language learning were learners' learning behaviors, learning attitudes, and learning performance. It also showed that task-based learning and game-based learning were the most common learning strategies.

Furthermore, one meta-analysis measuring the impact of VR on language learning has been published in journals. Wang et al. [21] conducted a meta-analysis to examine the effect size of 3D virtual worlds applied to language learning in the past two decades. The

research included 13 primary studies and found that overall effect sizes of linguistic and affective gains were 0.832 and 0.531, respectively.

However, the aforementioned review studies present certain limitations. In existing literature reviews, we find that most of the studies, through a qualitative approach, describe and summarize the findings of VR technology in language learning and do not consider quantitative methods with which to measure the impact of this technology on language learning. It is not easy to use qualitative approaches to evaluate the actual overall effectiveness of VR application on language learning and how moderator variables affect the effects. Therefore, the present research has important research implications. We conducted a meta-analysis to examine and quantify studies on the effectiveness of VR-assisted language learning and to determine which moderator variables influence the effectiveness of virtual reality to better guide the future application of virtual reality technology in language learning.

### 1.3. Purpose of This Meta-Analysis

This meta-analysis used the PICO framework [22] as the research framework. The PICO framework includes four major components: (1) Population. It refers to students of different education levels who may be involved in language learning through VR applications. (2) Intervention. It is considered as the treatments (e.g., the hardware types, target language, and L1/L2) of studies on VR-assisted language learning. (3) Comparison. It refers to the teaching strategy implemented in the control groups, including traditional teaching methods and other educational technology resources. (4) Outcomes. It focuses on the student's learning performance of VR applications on language learning, such as acquiring language skills and affective gains.

Therefore, the study proposed three purposes to provide a summary on the status quo of the application of virtual reality technology in language learning, to determine the overall effect size of virtual reality applications on students' language learning outcomes, and to identify whether moderator variables influence the impact of virtual reality on students' language learning gains. As a result, the research questions were formulated as follows:

1. What is the research status of VR technology-assisted language learning? Specifically, what education levels, hardware types, language skills, target language, and L1/L2 were involved in the existing studies about this field?
2. What is the overall effectiveness of virtual reality applications on students' language learning achievement?
3. What kinds of moderator variables influence the effectiveness of virtual reality on students' language learning outcomes?

## 2. Method

We followed the procedure of meta-analysis proposed by Glass, MacGaw, and Smith [23]. The procedures included (1) collecting studies, (2) coding the features of studies, (3) calculating the effect size of each study, and (4) investigating the moderating effects of the study's characteristics.

### 2.1. Data Sources and Search Strategy

The literature search procedure of this study rigorously followed the PRISMA guidelines [24] to collect literature related to VR-supported language learning. In order to obtain relevant sufficient and scientific papers, this study searched journal articles published from 2010 to 2021. The major databases used for the searches were the Web of Science Core Collection, Springer Link, Wiley Online Library, and Scopus. Two sets of search keywords were used to search: (1) VR technology-related keywords, including *VR*, *Virtual Reality*, *Virtual Environment*, and *Virtual Worlds*; and (2) language-related keywords, including *Language*, *Second Language*, *Foreign Language*, *Second Language Acquisition*, *EFL*, *Listening*, *Speaking*, *Reading*, *Writing*, and *Vocabulary*. These two sets of keywords were used in incorporation

with Boolean operators [25]. Moreover, the major journals of educational technology (e.g., Computers & Education, Computer Assisted Language Learning, Journal of Computer Assisted Learning, British Journal of Educational Technology, Language Learning & Technology, Interactive Learning Environments, and ReCALL) and the list of collected papers' references were perused as a supplement.

## 2.2. Inclusion and Exclusion Criteria

We followed the inclusion and exclusion criteria in Table 1 to select the literature to meet the meta-analysis requirements. All the papers had to be written in English and published between 2010 and 2021. The studies had to be empirical and had to have adopted either an experimental or quasi-experimental research design in which the experimental group used VR-assisted students' language learning compared to control groups that used other learning methods. Furthermore, the studies had to provide sufficient information to calculate effect sizes, such as standard deviations, means, and sample sizes. Finally, experimental results included learning achievement (e.g., different language skills) or affective gains, and the experimental and control groups' learning contents had to be the same.

**Table 1.** Inclusion and exclusion criteria.

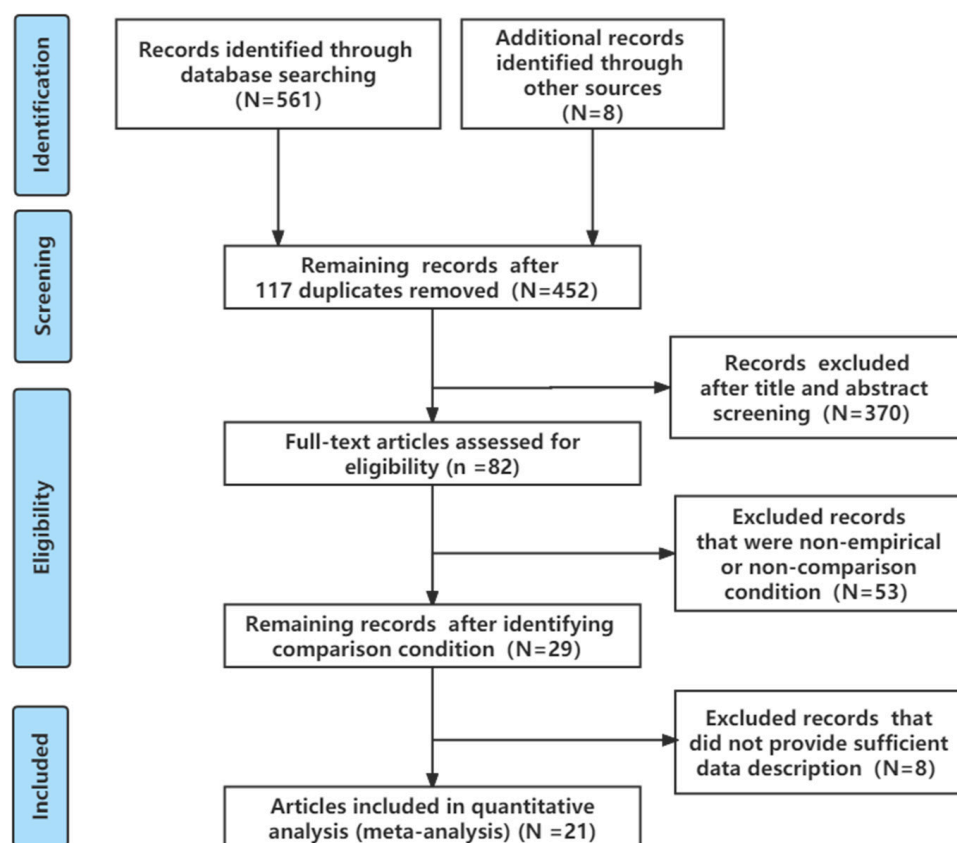
Criteria	Inclusion	Exclusion
Time	Published between 2010 and 2021	Not in this time period
Language	English	Non-English
Study type	Empirical study	Theoretical study
Intervention	VR technology	Non-VR
Research design	An experimental or quasi-experimental design that VR-assisted language learning compared to other learning methods	Non-experimental designs
Participants	Kindergarten, Elementary school, Middle school, and College	Adult
Content	The same learning content	Learning content are not the same
Results	Sufficient information to calculate the effect size	Lacked any statistical results or inadequate

## 2.3. Search Results

The process of literature identification, screening, eligibility, and eventual inclusion in this meta-analysis is shown in Figure 1. We found 561 journal articles related to VR-assisted language learning published from 2010 to 2021 and 8 articles from other sources during the first phase of literature selection. After removing 117 duplicates, two researchers read the title and abstract of each article and then identified whether or not the article was relevant to VR-assisted language learning. Ultimately, a total of 82 full-text articles were selected during the identification and screening process.

In the second phase, the included articles of the first stage were further screened. Only experimental and quasi-experimental studies involving VR-assisted language learning were carefully examined to identify whether the VR technology and other teaching/learning methods were included. At this stage, theoretical studies, qualitative research, and survey research were all excluded. After the second screening stage, 29 articles remained for consideration in the meta-analysis.

In the final stage, we examined the experimental data reported in the study. Studies that did not provide sufficient data to calculate effect sizes (e.g., mean, standard deviation, sample size) were excluded. After this stage, only 21 articles were included in the further meta-analysis. Of the 21 articles, only two were affective gains, and six included both linguistic and affective gains.



**Figure 1.** PRISMA flowchart of article selection.

#### 2.4. Coding Scheme

The coding scheme of the present study was composed of four main categories: basic research information, research participants' level, control treatment, research treatments, and learning outcomes. In addition to the basic research information, the other four items and their moderator variables correspond to the components of the PICO framework. The coding scheme was described in detail as follows:

##### 2.4.1. Basic Research Information

This referred to the author's last name, year of publication, region, and article's title.

##### 2.4.2. Research Participants' Level

The research participants' level corresponded to the "Population" of the PICO framework and was coded by their educational levels, including elementary school, middle school (junior or senior), and college.

##### 2.4.3. Control Treatment

The control treatment of the included studies corresponded to the "Comparison", referring to the different control group treatments. Previous meta-analyses had considered "control treatments" as moderator variables to compare experimental treatments with different control treatments [26,27]. Garzón and Acevedo [27] divided the control treatments into three types: multimedia, traditional lectures, and traditional pedagogical tools. In this meta-analysis, the control group treatments were classified into two categories: (1) traditional, which refers to traditional lectures and traditional pedagogical tools, such as curriculum, conventional teacher introduction, or textbooks, or (2) multimedia, which refers to educational resources accessed using videos, images, animation, or computer-assisted instruction.



#### 2.4.4. Treatments

In the meta-analysis, for all the included studies, the treatments corresponded to the PICO framework's "Intervention", which was considered as the hardware types, language skills, target language, and L1/L2. The specific descriptions for these treatments were as follows:

1. Hardware types. According to Qiu et al.'s [10] classification of the immersion level of the VR devices, VR can be divided into non-immersive and immersive devices. The non-immersive devices included desktop VR, smartphones, and tablet computers. Cave VR and head mount display were classified as immersive devices, such as Samsung Gear VR, phone cardboard, and cave-like VR.
2. Language skills. According to Hwang and Fu [4], Qiu et al. [10], and Sung et al. [28], the language skills were divided into vocabulary acquisition, listening, speaking, reading, writing, and mixed thereof.
3. Target languages. This study classified the coding variables of different target languages into English, Chinese, Spanish, German, and Korean.
4. L1/L2. According to Hwang and Fu's [4] review study on mobile language learning applications, the L1 referred to the native and first language that the learners learned from birth. The L2 usually referred to the learner's subsequent language after acquiring their first language.

#### 2.4.5. Learning Outcomes

The learning outcomes corresponded to the PICO framework's "Outcome". It referred to two categories of learning performance, including linguistic gains (e.g., language skills and knowledge or content learning) and affective gains (e.g., learning attitudes, motivation, and self-efficacy) [4,21].

The two researchers coded 21 articles according to the abovementioned coding rules. The coding process was performed in a consistent process, and whenever there was a question about coding and reaching consensus, the researchers held discussions to ensure that the coding of the article was consistent. If two researchers had different coding opinions, the article would be equally allocated to a third researcher for coding until they agree on all the code. In the end, coding was reviewed again to ensure it was accurate (Table A1).

### 2.5. Data Analysis

#### 2.5.1. Computing Effect Sizes

Comprehensive meta-analysis software was used to conduct the meta-analysis and compute the effect sizes. To compare the effect sizes, we chose to use Hedges' *g* as a standardized measure of effect sizes for computation in the present study. Hedges' *g* was the adjusted Cohen's *d* that was a measurement based on the mean difference between the two groups by the pooled standard deviation, and it was helpful for small sample size bias [29]. The calculation formula for Cohen's *d* was as follows:

$$\text{Cohen's } d = \frac{M_E - M_C}{\sqrt{\frac{(N_E - 1)S_E^2 + (N_C - 1)S_C^2}{(N_E - 1) + (N_C - 1)}}$$

where  $M_E$  and  $M_C$  were the estimated means of the experimental and control groups, respectively, with  $N_E$  and  $N_C$  being the sample sizes of both groups, and  $S_E^2$  and  $S_C^2$  the respective standard deviation. Hedges' *g* and Cohen's *d* were similar for large sample sizes, but Hedges' *g* performed best for small samples when Cohen's *d* was multiplied by a correction factor *J* (which can be adjusted for small sample bias):

$$J = 1 - \frac{3}{4(N - 2) - 1}$$

where  $N$  was the total sample size.

$$\text{Hedges' } g = J \times \text{Cohen's } d$$

Only one effect size was calculated for each article in the present study because a study contributing more than one effect size might lead to the biased overall effect size [30]. If a study had multiple effect sizes, they were combined into one value [29]. Simultaneously, the  $I^2$  test was used as heterogeneity examined, where  $I^2 < 25\%$  indicated low; 25–75% indicated moderate heterogeneity;  $\geq 75\%$  was considered substantial heterogeneity [31]. When  $I^2$  statistic  $> 50\%$ , heterogeneity was considered significant. If there was heterogeneity, the random-effects model was used; otherwise, the fixed-effect model was chosen.

### 2.5.2. Analyses of Publication Bias

Publication bias occurred when researchers published only favorable results [29]. To examine publication bias in this study, it was evaluated with the funnel plot and the fail-safe  $N$  test. We first assessed by visually observing the shapes of the funnel plot. If there was no publication bias, the plot resembles a symmetrical inverted funnel. Otherwise, the plot was asymmetric [32]. In addition to the funnel plot, the fail-safe  $N$  had been calculated. The fail-safe  $N$  refers to the number of unpublished studies required to reduce the effect size to an insignificant level [33]. There was no publication bias if the fail-safe  $N$  was larger than  $5n + 10$  (where  $n$  was the number of effect sizes).

## 3. Results

### 3.1. Descriptive Information

Twenty-one articles were included by the literature search, screening, coding, and extracting process in this meta-analysis. Table 2 describes the basic information of the 21 articles included in this study. These articles were published between 2010 and 2021, and the studies were quasi-experimental designs. Different regions had researched VR-assisted language learning, and Taiwan had the most relevant studies. The sample sizes ranged from 10 to 106 participants in both the treatment and control group, and the total sample size was 1144 (treatment group = 563 and control group = 581). The learning outcome included linguistic gain, affective gain, and both.

Table 3 shows the different moderator variables and their corresponding percentages. Regarding control treatment, multimedia teaching was the largest proportion (57.1%). The largest proportion of studies in the educational levels included college students (47.6%), and the second largest group was middle school students (38.1%). As for hardware types, immersive (57.1%) devices had a wider usage than non-immersive (42.9%). In terms of language skills, the most frequently studied were vocabulary (26.3%), writing (26.3%), and speaking (21.1%), followed by mixed (15.8%). Finally, the main target learning was mostly English (66.6%), followed by Chinese (19.0%), and the most studied (85.7%) was L2 learning, with the rest being L1 learning (14.3%).

**Table 2.** Basic Information of the included studies.

Study (Year)	Region	Sample Size (E/C)	Learning Outcome
Acar and Cavas (2020) [34]	Turkey	26 (15/11)	linguistic gain
Alfadil (2020) [35]	American	64 (32/32)	linguistic gain
Chen and Hwang (2020) [36]	Taiwan	93 (54/39)	linguistic gain affective gain
Chen and Liao (2021) [37]	Taiwan	106 (53/53)	linguistic gain affective gain
Chen et al. (2021) [38]	Taiwan	84 (42/42)	linguistic gain affective gain
Dolgunsöz et al. (2018) [39]	Turkey	48 (24/24)	linguistic gain
Ebadi and Ebadijalal (2020) [40]	Iran	20 (10/10)	linguistic gain affective gain
Huang et al. (2020) [7]	Taiwan	65 (30/35)	linguistic gain affective gain
Lan et al. (2018) [41]	Taiwan	44 (22/22)	linguistic gain
Lan et al. (2019) [42]	Singapore	60 (26/34)	linguistic gain
Neville (2015) [43]	American	32 (13/19)	linguistic gain
Nicolaidou et al. (2021) [44]	Cyprus	40 (20/20)	linguistic gain
Tai et al. (2020) [45]	Taiwan	49 (24/25)	linguistic gain
Tai and Chen (2021) [16]	Taiwan	72 (36/36)	linguistic gain
Urun et al. (2017) [46]	Turkey	72 (36/36)	linguistic gain
Wang et al. (2012) [47]	American	55 (20/35)	linguistic gain
Wehner et al. (2011) [15]	American	40 (20/20)	affective gain
Xu et al. (2011) [48]	Korea	64 (32/32)	affective gain
Xie et al. (2019) [49]	American	10 (4/6)	linguistic gain
Yang et al. (2010) [50]	Taiwan	60 (30/30)	linguistic gain affective gain
Yang et al. (2020) [51]	China	40 (20/20)	linguistic gain

Note: E, experimental group; C, control group.

**Table 3.** Categories of the 21 included articles.

Variable	Category	Number of Studies (N)	Proportion of Studies (%)
Control treatment	Traditional	9	0.429
	Multimedia	12	0.571
Education level	Elementary school	3	0.143
	Middle school	8	0.381
	College	10	0.476
Hardware type	Immersive	12	0.571
	Non-immersive	9	0.429
Language skill	Vocabulary	5	0.263
	Listening	2	0.105
	Speaking	4	0.211
	Reading	0	0.000
	Writing	5	0.263
	Mixed	3	0.158
Target language	English	14	0.666
	Chinese	4	0.190
	Spanish	1	0.048
	German	1	0.048
	Korean	1	0.048
L1/L2	L1	3	0.143
	L2	18	0.857



### 3.2. Overall Effect Size

Table 4 shows the overall effect sizes of linguistic gains and affective gains. Because the Q statistics and  $I^2$  statistics revealed that the effect sizes in linguistic gains ( $Q = 73.476$ ,  $I^2 = 75.502$ ,  $p < 0.001$ ) and affective gains ( $Q = 14.955$ ,  $I^2 = 53.317$ ,  $p < 0.05$ ) were of moderate heterogeneity, the random-effects model was used to pool the data [30]. Furthermore, this indicated that one or more moderators were attributable to this heterogeneity other than sampling error [52], and further moderator analysis was required, as described in Section 3.3.

**Table 4.** The effect sizes of linguistic- and affective-gain in random-effects model.

Category		Effect Size		95% CI	Heterogeneity		
Dependent Variable	k	g	SE		Q-Value	df	$I^2$
Linguistic	19	0.662	0.134 ***	[0.398–0.925]	73.476 ***	18	75.502
Affective	8	0.570	0.133 ***	[0.309–0.831]	14.955 *	7	53.317

Note: k, the number of effect sizes; g, Hedges' g; SE, standard error; Q-value, Q value of the heterogeneity test between the subgroups; CI, confidence interval; \*  $p < 0.05$ ; \*\*\*  $p < 0.001$ .

Using a random-effects model to pool the effect sizes, the overall effect sizes showed a statistically significant difference in linguistic gains ( $g = 0.662$ , CI [0.398–0.925],  $p < 0.001$ ) and affective gains ( $g = 0.570$ , CI [0.309–0.831],  $p < 0.001$ ) on VR-assisted language learning compared to other learning methods. According to Cohen [53], when the effect size  $\leq 0.2$ , it indicated a small effect, while when it was between 0.2 and 0.8, it was a moderate effect, and  $\geq 0.8$  was classified at a large effect. Moreover, this suggested that VR-assisted language learning had a moderate impact on students' linguistic gains and affect in language learning classrooms. In other words, language learning conditions using VR technologies had significantly better learning outcomes than non-VR conditions in language learning. These findings provided solid evidence for the advantages of VR devices in language learning and complemented previous qualitative review research on VR-assisted language learning [20].

### 3.3. The Effect Sizes of Moderator Variables on Linguistic Gains

To learn more about the moderating effects of linguistic gains, this study analyzed the impact of moderator variables on linguistic gains. Table 5 depicts the effect sizes for moderator variables.

**Table 5.** Descriptive for the moderator analysis on the linguistic gains.

Category	k	g	SE	95% CI	Q <sub>B</sub>	df
Control treatment					1.469	1
Traditional	9	0.840	0.199 ***	[0.450–1.229]		
Multimedia	10	0.514	0.181 **	[0.158–0.869]		
Education level					1.322	2
Elementary school	3	0.468	0.350	[−0.219–1.154]		
Middle school	8	0.845	0.213 ***	[0.429–1.262]		
College	8	0.546	0.223 *	[0.109–0.982]		
Hardware type					8.178 **	1
Immersive	12	0.409	0.142 **	[0.131–0.688]		
Non-immersive	7	1.091	0.191 ***	[0.716–1.466]		
Language skill					3.620	4
Vocabulary	5	0.761	0.258 **	[0.256–1.266]		
Listening	2	0.850	0.407 *	[0.052–1.648]		
Speaking	4	1.032	0.318 **	[0.409–1.655]		

Table 5. Cont.

Category	k	g	SE	95% CI	Q <sub>B</sub>	df
Writing	5	0.319	0.260	[−0.191–0.830]		
Mixed	3	0.522	0.339	[−0.142–1.186]		
Target language					0.754	2
English	14	0.731	0.162 ***	[0.414–1.049]		
Chinese	4	0.497	0.315	[−0.121–1.115]		
German	1	0.316	0.635	[−0.928–1.561]		
L1/L2					0.605	1
L1	2	0.361	0.412	[−0.446–1.168]		
L2	17	0.701	0.145 ***	[0.416–0.986]		

Note: k, the number of effect sizes; g, Hedges' g; SE, standard error; Q-value, Q value of the heterogeneity test between the subgroups; CI, confidence interval; \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

### 3.3.1. Control Treatments

Table 5 indicates that VR had a moderate impact on language learning. However, it was necessary to verify that VR contributed to this effect. In other words, it was important to establish that this effect was not the result of other interventions but the result of intervention with VR. To make this clear, the effect of VR treatment compared to the control treatment was examined. The results shown in Table 5 suggested that using VR technologies was more effective than using other teaching resources, including traditional ( $g = 0.840$ ,  $p < 0.001$ ) and multimedia ( $g = 0.514$ ,  $p < 0.01$ ) teaching resources. These findings showed that the improvement in language learning appeared to be related to the use of VR and not just the intervention, even if there was no significant difference ( $Q_B = 1.469$ ,  $p = 0.226$ ) between the various categories of control treatment.

### 3.3.2. Educational Levels

With respect to educational levels, middle school had largest effect on linguistic gains ( $g = 0.845$ ,  $p < 0.001$ ), followed by the college ( $g = 0.546$ ,  $p < 0.05$ ). The educational levels of elementary school ( $g = 0.468$ ,  $p = 0.182$ ) did not reach significant effect sizes.  $Q_B$  also did not reach the 0.05 significance level ( $Q_B = 1.322$ ,  $p = 0.516$ ), suggesting that there was no significant difference between the different categories of educational levels.

### 3.3.3. Hardware Types

There was a significant difference between the categories of hardware types ( $Q_B = 8.178$ ,  $p < 0.01$ ). The non-immersive devices had the largest effect size ( $g = 1.091$ ,  $p < 0.001$ ) but the immersive devices obtained a moderate effect size ( $g = 0.409$ ,  $p < 0.01$ ).

### 3.3.4. Language Skills

There was no significant difference among the various categories of language skills ( $Q_B = 3.620$ ,  $p = 0.460$ ). Except for writing skills ( $g = 0.319$ ,  $p = 0.220$ ) and mixed ( $g = 0.522$ ,  $p = 0.124$ ), which had no significant effect size, other language skills showed a significant effect size. The effect size was largest for speaking skills ( $g = 1.032$ ,  $p < 0.01$ ), followed by listening ( $g = 0.850$ ,  $p < 0.05$ ) and vocabulary ( $g = 0.761$ ,  $p < 0.01$ ). VR can enhance the learning performance of language representations and comprehension in a short period of time, such as speaking learning. For example, Chen and Hwang [36] found that the use of VR can enhance participants' move structures and levels of learning motivation in their speaking learning compared with conventional multimedia learning.

### 3.3.5. Target Languages

English obtained a moderate-to-high effect size ( $g = 0.731$ ,  $p < 0.001$ ), while there was no significant effect size for Chinese ( $g = 0.497$ ,  $p = 0.115$ ) and German ( $g = 0.316$ ,  $p = 0.618$ ). However,  $Q_B$  also did not reach the 0.05 significance level ( $Q_B = 0.754$ ,  $p = 0.686$ ), which indicated that there was no significant difference. In summary, the results showed that

VR applications for English learning were highly effective. Chinese and German did not achieve a significant effect, probably due to the small sample size.

### 3.3.6. L1/L2

The  $Q_B$  statistics also did not reach the 0.05 significance level, indicating that there was no significant difference ( $Q_B = 0.605$ ,  $p = 0.437$ ). The effect size for L2 achieved a large level ( $g = 0.701$ ,  $p < 0.001$ ), while there was no significant effect size for L1 ( $g = 0.361$ ,  $p = 0.380$ ).

### 3.4. The Effect Sizes of Moderator Variables on Affective Gains

Regarding the moderating effects of affective gains, this study analyzed the impact of moderator variables on affective gains. Table 6 shows the results of descriptive statistics for the moderator analysis on the affective gains.

**Table 6.** Descriptive for the moderator analysis on the affective gains.

Category	k	g	SE	95% CI	$Q_B$	df
Control treatment					0.448	1
Traditional	2	0.421	0.274	[−0.116–0.959]		
Multimedia	6	0.636	0.168 ***	[0.307–0.966]		
Education level					1.603	2
Elementary school	1	0.339	0.383	[−0.411–1.090]		
Middle school	2	0.380	0.256	[−0.121–0.882]		
College	5	0.719	0.179 ***	[0.367–1.071]		
Hardware type					2.513	1
Immersive	4	0.394	0.159 *	[0.082–0.706]		
Non-immersive	4	0.780	0.184 ***	[0.419–1.140]		
Target language					0.486	3
English	5	0.521	0.198 **	[0.132–0.909]		
Chinese	1	0.834	0.438	[−0.026–1.693]		
Spanish	1	0.700	0.478	[−0.237–1.637]		
Korean	1	0.570	0.436	[−0.284–1.425]		
L1/L2					0.300	1
L1	2	0.701	0.271 *	[0.171–1.231]		
L2	6	0.529	0.160 **	[0.215–0.842]		

Note: k, the number of effect sizes; g, Hedges' g; SE, standard error; Q-value, Q value of the heterogeneity test between the subgroups; CI, confidence interval; \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

#### 3.4.1. Control Treatments

There was a no significant difference between the different control treatments ( $Q_B = 0.448$ ,  $p = 0.503$ ). The multimedia had a moderate effect size ( $g = 0.636$ ,  $p < 0.001$ ), while there was no significant effect size for traditional ( $g = 0.421$ ,  $p = 0.125$ ).

#### 3.4.2. Educational Levels

Table 6 showed that  $Q_B$  did not reach the 0.05 significance level ( $Q_B = 1.603$ ,  $p = 0.449$ ), indicating no significant difference among the different educational levels. College usage had a moderate-to-high effect size on affective gains ( $g = 0.719$ ,  $p < 0.001$ ), while the educational levels of elementary school ( $g = 0.339$ ,  $p = 0.376$ ) and middle school ( $g = 0.380$ ,  $p = 0.137$ ) did not obtain significant effect sizes. Djigunovic (2014) reported that, due to a lack of contact with the language and world knowledge, children relied on teachers' guidance, which may reduce children's motivation to learn L2 in a situational linguistic learning environment. The reduction in affective gain was influenced by self-assessment rather than language experience, which was similar to the findings of Wang et al. [21]. Therefore, college students' affective gain was significantly larger than elementary and middle school.

### 3.4.3. Hardware Types

There was a no significant difference between the categories of hardware types ( $Q_B = 2.513$ ,  $p = 0.113$ ). The no-immersive devices obtained a large effect size ( $g = 0.780$ ,  $p < 0.001$ ), followed by the immersive devices ( $g = 0.394$ ,  $p < 0.05$ ).

### 3.4.4. Target Languages

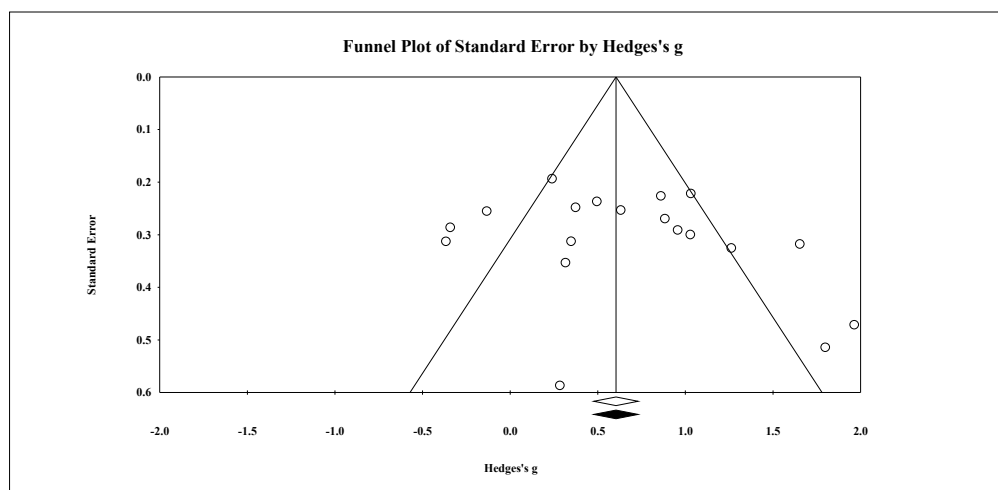
As shown in Table 6, English obtained a moderate effect size ( $g = 0.521$ ,  $p < 0.01$ ), while there was no significant effect size for Chinese ( $g = 0.834$ ,  $p = 0.057$ ), Spanish ( $g = 0.700$ ,  $p = 0.143$ ), and Korean ( $g = 0.570$ ,  $p = 0.191$ ). However,  $Q_B$  also did not reach the significance level ( $Q_B = 0.486$ ,  $p = 0.922$ ), which indicated that there was no significant difference.

### 3.4.5. L1/L2

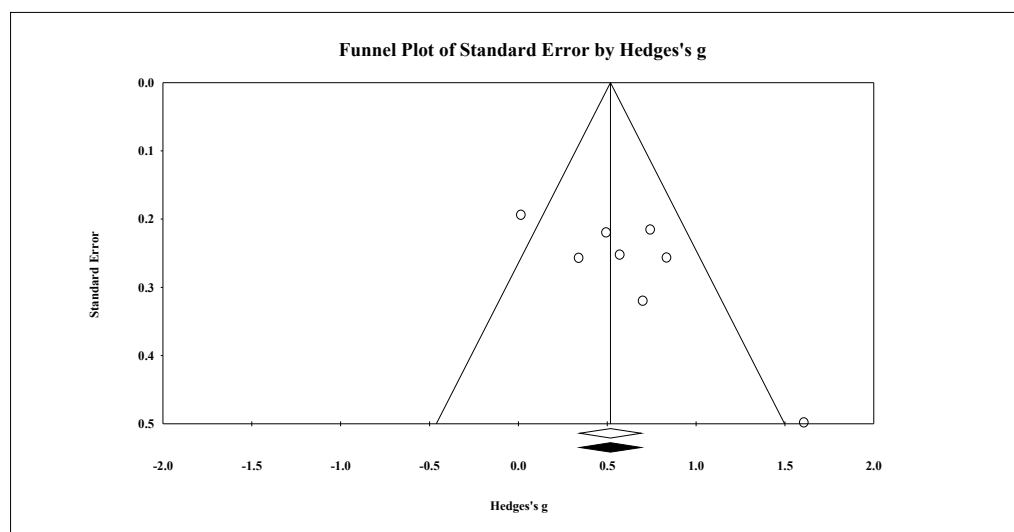
The  $Q_B$  statistics also did not reach the significance level, suggesting that there was no significant difference ( $Q_B = 0.300$ ,  $p = 0.584$ ). The effect size for L2 obtained a moderate effect size ( $g = 0.701$ ,  $p < 0.01$ ), and L1 also obtained a moderate effect size ( $g = 0.529$ ,  $p < 0.05$ ).

## 3.5. Evaluation of the Publication Bias

The funnel plot and classic fail-safe N were adopted in this study to examine whether the included studies were affected by publication bias. If there was no publication bias, the funnel plot was similar to a symmetrical inverted funnel [32]. The visual analysis in Figures 2 and 3 showed that the funnel chart was centralized and symmetrical, indicating no publication bias. To further confirm whether there was a publication bias, we computed the classic fail-safe N. The results of the classic fail-safe N showed that a total of 444 studies with invalid results were required to invalidate the effect size for linguistic gains, and a total of 77 studies were needed to invalidate the effect size for affective gains. To summarize, we concluded that publication bias was absent, which did not inflate the effect sizes.



**Figure 2.** Funnel plot assessing publication bias on linguistic gains.



**Figure 3.** Funnel plot assessing publication bias on affective gains.

#### 4. Discussion

Many studies have suggested the potential of VR-assisted language learning. However, there is little consensus on whether it can help improve language learning. Through integrating the finding of published empirical studies of VR-assisted language learning, our research provided concrete evidence on the overall effect sizes of VR-assisted language learning on students' linguistic and affective gain and how moderator variables influenced its effectiveness.

##### 4.1. The Effectiveness of VR-Assisted Language Learning

Based on the results of 21 articles ( $N = 1144$ ) in this meta-analysis, we find evidence for the overall effectiveness of VR applications on students' language learning achievement, and we also find an overall positive effect size of 0.662 and 0.570 on students' linguistic and affective gain, respectively, suggesting that VR can enhance students' language learning achievement compared to non-VR conditions. The overall effect size can be considered a medium effect [53], which is consistent with the findings of Wang et al. [21] that VR-assisted language learning can facilitate language knowledge acquisition and enhance affection. The positive findings related to VR-assisted language learning may be attributed to several features of virtual reality: (1) immersive learning can effectively promote language learning [49]; (2) improve language skills through interaction between VR and learners [5,54]; (3) effectively filter by authentic language learning settings (e.g., learning anxiety) [7,49]. Our meta-analysis reveals the potential of using VR applications in language learning and provides teachers with options to support their teaching.

##### 4.2. The Moderator Analyses of VR-Assisted Language Learning

In this study, we identify six moderator variables and investigate how the design of the intervention influenced the effectiveness of VR-supported language learning. Although the overall results are positive, we should consider that individual studies' results may vary by factors, such as control treatment, education level, hardware type, target language, language skill, L1/L2. However, the findings of our moderator analysis indicate that only one moderator (hardware type) is significant at  $p < 0.01$ . Compared to immersive devices, non-immersive devices have the most significant effect on linguistic gains. The reason for this is that long-term use of immersive devices affects learners' senses, such as dizziness [10]. Furthermore, students might only focus on the interesting content rather than the learning content, failing to achieve learning goals [21]. It also shows that the effect of immersive and non-immersive devices on learners' language learning can be compared in future research.

Regarding the control treatment, we compared VR applications as a teaching resource with other types of teaching resources, including multimedia and traditional resources. We do not find significant differences among the control treatment in this study and suggest that VR applications promote language learners' linguistic and affective gain compared to multimedia and traditional resources. Moreover, higher-education students are the main research subjects, consistent with some research on VR-assisted language learning [10,21]. The reason for this may be that most of the researchers who support language learning with VR come from colleges and may find college students easier than collecting relevant data in K-12 schools. While VR has a large effect on the linguistic gains of middle education and a medium-to-large effect on the affective gains of higher education students, these differences were not statistically significant in the different educational levels. Overall, the language learning of the different educational levels was positively influenced by VR applications, which is consistent with previous research results [21].

The findings of our moderator analysis indicated a large effect size for speaking than other language skills in terms of language domains. However, these differences were not statistically significant for different language skills. It should be noted that research on reading is lacking, which should be addressed in future research. At the same time, the result related to the target language and L1/L2 revealed that most studies on VR-assisted language learning have focused on English as an L2. This may be because learning an L2 is far more challenging than learning L1 [17]. Our results regarding this moderator suggest that VR technology could be used in both L1 and L2 learning, indicating that these have a good potential as an educational resource.

#### *4.3. Limitations of the Study*

As a limitation of our meta-analysis, only journal articles published from 2010 to 2021 in the two major databases and major journals were included in this study. Although it was found that publication bias was unlikely to be an issue for this meta-analysis, the findings of linguistic and affective gains might be influenced by studies with insignificant results that were rejected due to the small sample of the study. Moreover, articles excluded by the eligibility criteria may also alter the meta-analysis results. More studies were needed to assess the actual value of large effect sizes with small samples. Therefore, future research should include journal articles, conference papers, doctoral theses, reviews, and articles in languages other than English should also be considered. Moreover, other databases (e.g., ProQuest or JSTOR) and longer time periods should be considered when searching for relevant articles.

#### **5. Conclusions**

Based on the above discussion, several suggestions for VR-assisted language learning are put forward as follows:

First, future research should enrich the diversity of applications of VR-assisted language learning. The essence of VR-assisted language learning is VR technology designed to improve the effectiveness of language teaching/learning methods. Existing research should not be limited to higher education but should focus more on kindergarten and K-12 education. For example, Cerezo, Calderón, and Romero [55] adopted VR-assist language learning to help preschool children practice the pronunciation of basic English vocabulary. The results showed that VR-assisted language learning significantly impacted the children's motivation and improved their performance compared to traditional methods. Furthermore, most of the studies focused on vocabulary-related knowledge (e.g., vocabulary and writing) and a lack of grammar knowledge and reading skills. Language learning programs mainly involve English as a foreign language, with few studies having focused on learning other languages as a second or foreign language, and there has been little research related to native language teaching. Thus, future research should consider different languages with different skills to expand the diversity of VR-assisted language learning.



Second, future research should examine the effects of different VR devices on language learning. The results of our moderator analysis show that only the different device types had significant differences in this meta-analysis. Non-immersive devices performed better for students' linguistic gain than immersive devices. The existing studies have mainly compared VR with traditional teaching on language learning, and there is a lack of research on the effect of different VR devices in language learning. Therefore, it is worthwhile to examine the effects of different VR devices on language learning in future research.

Third, future research should not only focus on language knowledge acquisition and affective enhancement, but also focus on the development of higher-order thinking in the process of language acquisition. From these studies included in the meta-analysis, it can be found that the results of language gains are mainly obtained through post-intervention tests, and the analysis of emotional gain is mainly obtained through questionnaires. However, there is a lack of a deeper mechanism exploration of VR-assisted language learning, such as higher-order cognition or behavior evaluation. With the development of emerging technologies such as artificial intelligence and big data, more innovative approaches, such as learning behavior analysis, should be used to understand the nature of VR-supported language learning. For example, there has been evidence from neuroscience research to support immersion learning for L2 acquisition [56]. Future research should not only consider learning benefits, but also examine the development or changes in higher-order thinking or abilities.

**Author Contributions:** Conceptualization, B.C., Y.W. and L.W.; methodology, B.C., Y.W. and L.W.; software, B.C.; validation, B.C. and Y.W.; formal analysis, B.C., Y.W. and L.W.; investigation, B.C. and Y.W.; resources, B.C. and Y.W.; data curation, B.C. and Y.W.; writing—original draft preparation, B.C., Y.W. and L.W.; writing—review and editing, B.C., Y.W. and L.W.; visualization, B.C. and Y.W.; supervision, L.W.; project administration, B.C. and L.W.; funding acquisition, L.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Leading Talent Project of Philosophy and Social Science Planning of Zhejiang Province grant number 22YJRC02ZD-1YB.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** The coded papers.

Authors	Title	Journal	Reference
Acar and Cavas (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	[34]
Alfadil (2020)	Effectiveness of virtual reality game in foreign language vocabulary acquisition	Computers & Education	[35]
Chen and Hwang (2020)	Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles	Interactive Learning Environments	[36]
Chen and Liao (2021)	Effects of panoramic image virtual reality on the workplace English learning performance of vocational high school students	Journal of Educational Computing Research	[37]

Table A1. Cont.

Authors	Title	Journal	Reference
Chen et al. (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	[38]
Dolgunsöz et al. (2018)	The effect of virtual reality on EFL writing performance	Journal of Language and Linguistic Studies	[39]
Ebadi and Ebadijalal (2020)	The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency	Computer Assisted Language Learning	[40]
Huang et al. (2020)	Learning to be a writer: a spherical video-based virtual reality approach to supporting descriptive article writing in high school Chinese courses	British Journal of Educational Technology	[7]
Lan et al. (2018)	Real body versus 3D avatar: the effects of different embodied learning types on EFL listening comprehension	Educational Technology Research & Development	[41]
Lan et al. (2019)	Does a 3D immersive experience enhance Mandarin writing by CSL students?	Language Learning & Technology	[42]
Neville (2015)	The story in the mind: the effect of 3D gameplay on the structuring of written L2 narratives	ReCALL	[43]
Nicolaidou et al. (2021)	Comparing immersive virtual reality to mobile applications in foreign language learning in higher education: a quasi-experiment	Interactive Learning Environments	[44]
Tai et al. (2020)	The impact of a virtual reality app on adolescent EFL learners' vocabulary learning	Computer Assisted Language Learning	[45]
Tai and Chen (2021)	The impact of immersive virtual reality on EFL learners' listening comprehension	Journal of Educational Computing Research	[16]
Urun et al. (2017)	Supporting Foreign Language Vocabulary Learning Through Kinect-Based Gaming	Journal of Game-Based Learning	[46]
Wang et al. (2012)	Learning effects of an experimental EFL program in second life	Educational Technology Research & Development	[47]
Wehner et al. (2011)	The effects of Second Life on the motivation of undergraduate students learning a foreign language	Computer Assisted Language Learning	[15]
Xu et al. (2011)	A New Approach Toward Digital Storytelling: An Activity Focused on Writing Self-efficacy in a Virtual Learning Environment	Educational Technology & Society	[48]
Xie et al. (2019)	Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency	Computer Assisted Language Learning	[49]
Yang et al. (2010)	Integrating video-capture virtual reality technology into a physically interactive learning environment for English learning	Computers & Education	[50]
Yang et al. (2020)	From experiencing to expressing: a virtual reality approach to facilitating pupils' descriptive paper writing performance and learning behavior engagement	British Journal of Educational Technology	[51]

## References

1. McKay, S.L. English as an international language: What it is and what it means for pedagogy. *RELJ* **2018**, *49*, 9–23. [[CrossRef](#)]
2. Moeller, A.J.; Catalano, T. *Foreign Language Teaching and Learning*. *International Encyclopedia of the Social & Behavioral Sciences*, 2nd ed.; Pergamon Press: Oxford, UK, 2015.

3. Golonka, E.M.; Bowles, A.R.; Frank, V.M.; Richardson, D.L.; Freynik, S. Technologies for foreign language learning: A review of technology types and their effectiveness. *Comput. Assist. Lang. Learn.* **2014**, *27*, 70–105. [\[CrossRef\]](#)
4. Hwang, G.J.; Fu, Q.K. Trends in the research design and application of mobile language learning: A review of 2007–2016 publications in selected SSCI journals. *Interact. Learn. Environ.* **2018**, *27*, 567–581. [\[CrossRef\]](#)
5. Chen, Y.L. The effects of virtual reality learning environment on student cognitive and linguistic development. *Asia-Pac. Educ. Res.* **2016**, *25*, 637–646. [\[CrossRef\]](#)
6. Genz, F.; Fuchs, N.; Kolb, D.; Müller, S.; Kranzlmüller, D. Evaluation of Proprietary Social VR Platforms for Use in Distance Learning. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2021; Volume 12980. [\[CrossRef\]](#)
7. Huang, H.L.; Hwang, G.J.; Chang, C.Y. Learning to be a writer: A spherical video-based virtual reality approach to supporting descriptive article writing in high school chinese courses. *Br. J. Educ. Technol.* **2020**, *51*, 1386–1405. [\[CrossRef\]](#)
8. Hay, K.E. Educational application of virtual reality: A rational and case studies of 3D visualization and world building. In *Proceedings of the Indiana University Virtual Reality Conference*, Bloomington, IN, USA, 30 September 1997; Volume 58. [\[CrossRef\]](#)
9. Rene, G. Revolutionizing Education: The promise of virtual reality. *Child. Educ.* **2018**, *94*, 40–43. [\[CrossRef\]](#)
10. Qiu, X.Y.; Chiu, C.K.; Zhao, L.L.; Sun, C.F.; Chen, S.J. Trends in VR/AR technology-supporting language learning from 2008 to 2019: A research perspective. *Interact. Learn. Environ.* **2021**, 1–24. [\[CrossRef\]](#)
11. Hwang, W.Y.; Hu, S.S. Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving. *Comput. Educ.* **2013**, *62*, 308–319. [\[CrossRef\]](#)
12. Parong, J.; Mayer, R.E. Learning science in immersive virtual reality. *J. Educ. Psychol.* **2018**, *110*, 785–797. [\[CrossRef\]](#)
13. Vapenstad, C.; Buzink, S.N. Procedural Virtual Reality simulation in minimally invasive surgery. *Surg. Endosc.* **2013**, *27*, 364–377. [\[CrossRef\]](#)
14. Pinho, M.S.; Bowman, D.A.; Freitas, C.M.D.S. Cooperative object manipulation in collaborative virtual environments. *J. Braz. Comput. Soc.* **2008**, *14*, 53–67. [\[CrossRef\]](#)
15. Wehner, A.K.; Gump, A.W.; Downey, S. The effects of Second Life on the motivation of undergraduate students learning a foreign language. *Comput. Assist. Lang. Learn.* **2011**, *24*, 277–289. [\[CrossRef\]](#)
16. Tai, T.Y.; Chen, H.J. The impact of immersive virtual reality on EFL learners' listening comprehension. *J. Educ. Comput. Res.* **2021**, *59*, 1272–1293. [\[CrossRef\]](#)
17. Lin, T.J.; Lan, Y.J. Language Learning in Virtual Reality Environments: Past, Present, and Future. *Educ. Technol. Soc.* **2015**, *18*, 486–497.
18. Solak, E.; Erdem, G. A content analysis of virtual reality studies in foreign language education. *Participat. Educ. Res.* **2015**, *15*, 21–26. [\[CrossRef\]](#)
19. Parmaxi, A. Virtual reality in language learning: A systematic review and implications for research and practice. *Interact. Learn. Environ.* **2020**, 1–13. [\[CrossRef\]](#)
20. Huang, X.; Zou, D.; Cheng, G.; Xie, H. A systematic review of AR and VR enhanced language learning. *Sustainability* **2021**, *13*, 4639. [\[CrossRef\]](#)
21. Wang, C.P.; Lan, Y.J.; Tseng, W.T.; Lin, Y.T.R.; Gupta, C.L. On the effects of 3d virtual worlds in language learning—A meta-analysis. *Comput. Assist. Lang. Learn.* **2020**, *33*, 891–915. [\[CrossRef\]](#)
22. Higgins, J.; Green, S. *Cochrane Handbook for Systematic Reviews of Interventions*; Version 5.1.0; Wiley: Hoboken, NJ, USA, 2011.
23. Glass, G.V.; McGaw, B.; Smith, M.L. *Meta-Analysis in Social Research*; Sage Publications: Thousand Oaks, CA, USA, 1981.
24. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *J. Clin. Epidemiol.* **2009**, *62*, 1006–1012. [\[CrossRef\]](#)
25. Cooper, H. *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*; Sage Publications: London, UK, 2010.
26. Merchant, Z.; Goetz, E.T.; Cifuentes, L.; Keeney-Kennicutt, W.; Davis, T.J. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Comput. Educ.* **2014**, *70*, 29–40. [\[CrossRef\]](#)
27. Garzón, J.; Acevedo, J. Meta-analysis of the impact of Augmented Reality on students' learning gains. *Educ. Res. Rev.* **2019**, *27*, 244–260. [\[CrossRef\]](#)
28. Sung, Y.T.; Chang, K.E.; Yang, J.M. How effective are mobile devices for language learning? A meta-analysis. *Educ. Res. Rev.* **2015**, *16*, 68–84. [\[CrossRef\]](#)
29. Borenstein, M.; Hedges, L.V.; Higgins, J.P.T.; Rothstein, H.R. *Introduction to Meta-Analysis*; Wiley: Chichester, UK, 2009.
30. Lipsey, M.W.; Wilson, D.B. *Practical Meta-Analysis*; Sage: Thousand Oaks, CA, USA, 2001.
31. Higgins, J.P.; Thompson, S.G.; Deeks, J.J.; Altman, D.G. Measuring inconsistency in meta-analyses. *Br. Med. J.* **2003**, *327*, 557–560. [\[CrossRef\]](#)
32. Egger, M.; Smith, G.D.; Schneider, M.; Minder, C. Bias in meta-analysis detected by a simple, graphical test. *Br. Med. J.* **1997**, *315*, 629–634. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Rosenthal, R. The file drawer problem and tolerance for null results. *Psychol. Bull.* **1979**, *86*, 638–641. [\[CrossRef\]](#)
34. Acar, A.; Cavas, B. The Effect of Virtual Reality Enhanced Learning Environment on the 7th-Grade Students' Reading and Writing Skills in English. *MOJES Malays. Online J. Educ. Sci.* **2020**, *8*, 22–33.
35. Alfadil, M. Effectiveness of virtual reality game in foreign language vocabulary acquisition. *Comput. Educ.* **2020**, *153*, 103893. [\[CrossRef\]](#)
36. Chen, M.R.A.; Hwang, G.J. Effects of experiencing authentic contexts on english speaking performances, anxiety and motivation of EFL students with different cognitive styles. *Interact. Learn. Environ.* **2020**, 1–21. [\[CrossRef\]](#)

37. Chen, H.L.; Liao, Y.C. Effects of panoramic image virtual reality on the workplace english learning performance of vocational high school students. *J. Educ. Comput. Res.* **2021**, *8*, 1601–1622. [\[CrossRef\]](#)
38. Chen, C.H.; Hung, H.T.; Yeh, H.C. Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of english language learners. *J. Comput. Assist. Learn.* **2021**, *37*, 851–860. [\[CrossRef\]](#)
39. Dolgunsöz, E.; Yildirim, G.; Yildirim, S. The effect of virtual reality on EFL writing performance. *J. Lang. Linguist. Stud.* **2018**, *14*, 278–292.
40. Ebadi, S.; Ebadijalal, M. The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency. *Comput. Assist. Lang. Learn.* **2020**, 1–25. [\[CrossRef\]](#)
41. Lan, Y.J.; Fang, W.C.; Hsiao, I.; Chen, N.S. Real body versus 3D avatar: The effects of different embodied learning types on EFL listening comprehension. *Educ. Technol. Res. Dev.* **2018**, *66*, 709–731. [\[CrossRef\]](#)
42. Lan, Y.J.; Lyu, B.N.; Chin, C.K. Does a 3D immersive experience enhance Mandarin writing by CSL students? *Lang. Learn. Technol.* **2019**, *23*, 125–144.
43. Neville, O.D. The story in the mind: The effect of 3D gameplay on the structuring of written L2 narratives. *ReCALL* **2015**, *27*, 21–37. [\[CrossRef\]](#)
44. Nicolaidou, I.; Pissas, P.; Boglou, D. Comparing immersive virtual reality to mobile applications in foreign language learning in higher education: A quasi-experiment. *Interact. Learn. Environ.* **2021**, 1–15. [\[CrossRef\]](#)
45. Tai, T.Y.; Chen, H.J.; Todd, G. The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Comput. Assist. Lang. Learn.* **2020**, 1–26. [\[CrossRef\]](#)
46. Urun, M.F.; Aksoy, H.; Comez, R. Supporting Foreign Language Vocabulary Learning Through Kinect-Based Gaming. *Int. J. Game-Based Learn.* **2019**, *7*, 20–35. [\[CrossRef\]](#)
47. Wang, C.X.; Calandra, B.; Hibbard, S.T.; Lefaiver, M.D. Learning effects of an experimental EFL program in second life. *Educ. Technol. Res. Dev.* **2012**, *60*, 943–961. [\[CrossRef\]](#)
48. Xu, Y.; Park, H.; Baek, Y. A New Approach Toward Digital Storytelling: An Activity Focused on Writing Selfefficacy in a Virtual Learning Environment. *Educ. Technol. Soc.* **2011**, *14*, 181–191.
49. Xie, Y.; Chen, Y.; Ryder, L.H. Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Comput. Assist. Lang. Learn.* **2019**, *34*, 225–245. [\[CrossRef\]](#)
50. Yang, J.C.; Chen, C.H.; Jeng, C.M. Integrating video-capture virtual reality technology into a physically interactive learning environment for english learning. *Comput. Educ.* **2010**, *55*, 1346–1356. [\[CrossRef\]](#)
51. Yang, G.; Chen, Y.T.; Zheng, X.L.; Hwang, G.J. From experiencing to expressing: A virtual reality approach to facilitating pupils' descriptive paper writing performance and learning behavior engagement. *Br. J. Educ. Technol.* **2020**, *52*, 807–823. [\[CrossRef\]](#)
52. Borenstein, M.; Hedges, L.V.; Higgins, J.; Rothstein, H.R. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res. Synth. Methods* **2010**, *1*, 97–111. [\[CrossRef\]](#)
53. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Erlbaum: Hillsdale, NJ, USA, 1988.
54. Liaw, M.L. EFL learners' intercultural communication in an open social virtual environment. *J. Educ. Technol. Soc.* **2019**, *22*, 38–55.
55. Cerezo, R.; Calderón, V.; Romero, C. A holographic mobile-based application for practicing pronunciation of basic english vocabulary for spanish speaking children. *Int. J. Hum.-Comput. Stud.* **2019**, *124*, 13–25. [\[CrossRef\]](#)
56. Linck, J.A.; Kroll, J.F.; Sunderman, G. Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychol. Sci.* **2009**, *20*, 1507–1515. [\[CrossRef\]](#)