Alizadeh & Cowie *Research and Practice in Technology Enhanced Learning* (2027) 22:32
https://doi.org/10.58459/rptel.2027.22032

Research and Practice in Technology Enhanced Learning

RESEARCH Free and Open Access

Online flipped learning using video conferencing and virtual reality: A comparative study

Mehrasa Alizadeh* and Neil Cowie

*Correspondence:
m-alizadeh@otemon.ac.jp
Institute of General Education, Otemon Gakuin University, 1-1 OdaToshiba-cho, Ibaraki, Osaka 567-8620, Japan
Full list of author information is available at the end of the article

Abstract

Online flipped learning is widely used in foreign language education, including in fully online contexts. Past research has relied on video conferencing tools such as Zoom for facilitating online interaction, and alternative platforms such as virtual reality (VR) remain underexplored. This quasi-experimental study compared two online flipped learning environments, one using the video conferencing app Zoom and the other a WebVR platform. Volunteer Japanese university students participated in a multi-week video-based course on small talk skills for EFL learners. Each week, participants watched two videos and engaged in one-hour online sessions involving group discussions, mini-quizzes, and pair-based small talk practice. The Zoom group (N=30) and the VR group (N=25) received the same instructional content. VR participants received additional training and tutorial videos on platform navigation, avatar creation, and communication. Student interactions were recorded and rated at three points, and both groups completed engagement and anxiety surveys and focus group interviews. Findings showed significant improvement in small talk skills across both groups, with no major differences in learning gains, engagement or anxiety. While Zoom was familiar and user-friendly, student reactions to VR were mixed, highlighting both immersive benefits and usability challenges. The study contributes to the growing literature on synchronous online language learning by demonstrating that VR environments may support learning outcomes comparable to video conferencing platforms while also highlighting pedagogical and technological constraints.

Keywords: flipped learning, small talk, virtual reality (VR), mixed-methods research, learning outcomes

Introduction

Online language education has evolved over the past decades alongside technology advancements. Growing global access to the internet, increasing availability of digital tools,

and a rising demand for flexible, learner-centred education have all contributed to the expanding interest in online language learning (Tao & Gao, 2022). Among the various modes of online learning, flipped learning, also referred to as the flipped classroom model, has been recognised as a popular approach to language learning and teaching (Vitta & Al-Hoorie, 2023). This high interest in flipped learning is partly due to the fact that, in a flipped learning environment, students play a more active role by taking the lead in learning the content independently through watching videos or studying other materials usually provided by the instructor before taking a lesson. The commonly used format in flipped learning is for students to use online resources to learn new material and then use the classroom time to engage in higher-order learning skills that could involve activities such as discussion, analysis and synthesis of information, or creating a final product in groups (Baig & Yadegaridehkordi, 2023). However, with the advent of numerous learning platforms, it is now equally possible to adopt flipped learning in fully remote contexts (Jiang et al., 2021; Lee, 2023). When thoughtfully designed and implemented with gamified incentives to attract and motivate students, such an approach not only enhances learning outcomes but also fosters sustained engagement and the development of self-regulation skills, as highlighted in previous research (see Jia et al., 2023; Tutal & Yazar, 2021; Zou et al., 2022). Such research has looked at how fully remote flipped learning works using various platforms for the interactive portions of such classes; for example, through synchronous video chat or virtual reality (VR). However, as there is relatively little research that has examined the effectiveness of each, this study seeks to compare the usage of such platforms and examine the pedagogical differences between them.

Review of the literature

This section summarises the key features of flipped learning and provides an overview of how video conferencing and VR can be used for language learning. It also discusses how the use of VR may reduce foreign language anxiety. The section concludes by identifying research gaps and presenting the research questions.

Flipped learning in foreign language education

Flipped learning has emerged as a widely used instructional approach in online foreign language education. In flipped language learning contexts, learners typically engage with instructional input, such as video lectures or texts, prior to class, while synchronous sessions are devoted to applying the knowledge they gained before class to solve problems, interact with each other, practice their skills, and receive teacher and peer feedback (Zou et al., 2022). This approach has been widely adopted as a way to maximise limited classroom time and to promote learner-centred instruction (Leis, 2018; Pang, 2022).

Research suggests that flipped learning can be effective in online language learning settings when it is supported by careful task design and learner preparation (Wu et al.,

2020). Pre-class activities allow learners to process linguistic input at their own pace, which can help reduce cognitive load during live sessions and enable more meaningful participation in communicative tasks (Chen Hsieh et al., 2017; Xin & Zhang, 2024). By externalizing a portion of cognitive processing to the pre-class phase, learners may be better positioned to focus on interaction, fluency, and meaning negotiation during synchronous activities. At the same time, synchronous sessions play a crucial role in providing opportunities for interaction, negotiation of meaning, and instructor feedback, all of which are essential for language development.

Flipped learning can foster learner engagement by allowing students to regulate the pace and timing of their exposure to instructional content. Access to pre-class materials enables learners to review content multiple times and prepare more thoroughly for in-class activities, which may support deeper understanding and more confident participation during synchronous sessions. Moreover, the collaborative nature of flipped classrooms encourages knowledge co-construction through peer interaction and teacher guidance, contributing to more meaningful learning experiences (Bui et al., 2022; Divjak et al., 2022; Tang et al., 2020; Zainuddin et al., 2021).

Another reported benefit of flipped learning, closely related to learner engagement, is improved retention of learning content (Al-Zahrani, 2024). By engaging with instructional materials across multiple phases (before, during, and sometimes after class), learners encounter repeated opportunities to process and apply new knowledge in varied contexts. This cyclical engagement with content has been associated with enhanced consolidation of learning and more sustained engagement over time. As a result, the flipped classroom model has been increasingly adopted across educational settings as a means of optimizing instructional time and supporting active learning (Hwang et al., 2019).

Despite its pedagogical potential, flipped learning also presents several challenges that may limit its effectiveness, particularly in online contexts. A recurring concern in the literature is uneven learner engagement with pre-class materials. When students fail to complete preparatory activities or engage only superficially, the effectiveness of synchronous sessions may be compromised, as in-class tasks often assume a certain level of prior knowledge (Fischer & Yang, 2022). Learners who struggle to keep pace with pre-class content may experience cognitive overload, frustration, or reduced motivation, negatively affecting participation and learning outcomes (Lee & Choi, 2018).

Systematic reviews of flipped classroom research further highlight mixed findings and emphasise the context-dependent nature of flipped learning outcomes (Akçayır & Akçayır, 2018; Linling & Abdullah, 2023; Lo, 2024; Turan & Akdag-Cimen, 2019; Zainuddin & Halili, 2016). Identified challenges include pedagogical, technological, and learner-related dimensions, with inadequate student preparation and limited support for self-regulated learning emerging as particularly influential factors. In online flipped settings, additional constraints have been reported during synchronous sessions, including reduced participation, limited instructor visibility, and restrictions on hands-on or experiential activities (Lo, 2024).

While flipped learning provides a shared pedagogical structure across online modalities, the nature and quality of synchronous interaction may be shaped by the technological

platform through which it is implemented. In particular, features such as platform familiarity, cognitive load, social presence and embodiment are likely to influence how learners experience and engage in synchronous flipped activities, as well as how they perform speaking tasks and manage anxiety.

Video conferencing and language learning

Synchronous video conferencing using platforms such as Zoom, Microsoft Teams or Google Meet is a means where remote participants can interact with each other in a multimodal way through speech or text chat. Due to the COVID-19 pandemic, video conferencing became a ubiquitous way for educational institutions to provide online lessons for their students who could not meet face to face in a classroom. Prior to this, video conferencing was (and continues to be) a popular way for language learners to interact with students from other countries without having to travel long distances. Such virtual exchange or telecollaboration has provided fruitful research evidence that using video conferencing can help improve student speaking skills and intercultural competence (Alizadeh & Cowie, 2025; Spring et al., 2019; Tecedor & Kerber, 2025). On the other hand, there are a number of challenges that can hinder the effectiveness of video conferencing. One is that technological problems such as lost connections or distorted audio can disrupt lessons. Students can also suffer a lack of engagement and boredom from the challenges of studying on an online platform; there are also other symptoms of so-called 'Zoom fatigue' (Peper et al., 2021) such as feelings of social isolation and initial stress that are caused by using unfamiliar technology. However, overall, it is clear that when effective pedagogy can overcome these limitations (White, 2024), video conferencing clearly benefits learners. In addition, it has now become a very familiar and accustomed option for many students globally.

VR and language learning

VR has emerged as an immersive technology that allows learners to engage with different types of environments. In simple terms, VR refers to a computer-generated simulation that immerses users in an interactive virtual environment, which may be accessed through desktop-based (monoscopic) systems or head-mounted displays (stereoscopic), offering varying degrees of immersion. Early virtual worlds, such as Second Life, provided users with immersive, interactive environments designed primarily for social interaction and exploration. Although these platforms were not created specifically for educational purposes, researchers have noted that foreign language learners could make use of their affordances to engage in authentic communication, practice language skills, and explore language use in culturally rich virtual contexts (Chen, 2016). However, over time, there has been a gradual loss of interest in early virtual worlds, due to the emergence of newer and more advanced VR technologies, as well as the challenges associated with their user interface and stability (Parmaxi, 2020).

Current VR platforms can be accessed through head-mounted displays, PCs, or mobile devices, with some platforms being device agnostic. They encompass a wide array of experiences primarily designed for entertainment purposes, but there are also social VR apps available that hold potential for educational use and fostering interactive learning. VR environments have certain features that make them a suitable platform for interactive and collaborative learning. For instance, the sense of being immersed in the target culture through traveling virtually to locations one cannot easily access without long travel and financial costs is a huge benefit (Chun et al., 2022; Lan, 2020). Also, when in VR, learners are usually represented as avatars, and this can help some learners in overcoming inhibition caused by foreign language anxiety and develop their communication skills in a stress-free environment (Chen, 2022). In addition, learning by doing things using one's body, i.e., embodied cognition, can further enhance learning outcomes (Fuhrman et al., 2021; Legault et al., 2019; Li et al., 2020).

Language educators and practitioners have explored the impact of VR on a wide range of language learning skills and components, including but not limited to speaking, vocabulary, grammar, and pronunciation. Among these, vocabulary development and speaking skills have received considerable attention (Bahari, 2021; Deng & Yu, 2022; Lestiono & Lee, 2024), largely due to the interactive and immersive nature of VR environments, which closely simulate real-life communication contexts and provide learners with meaningful, contextualized language use within a low-risk environment. Following a situated learning approach, Yan et al. (2024) demonstrated that desktop-based VR can significantly improve students' fluency, vocabulary, pronunciation, and grammar through collaborative tasks in authentic VR contexts.

Several studies have demonstrated that immersive VR environments provide rich, contextualized exposure to target vocabulary and opportunities for learners to engage in meaningful interaction, which supports deeper semantic processing and retention (Dhimolea et al., 2022). For instance, Tai et al. (2020) found that learners who engaged in vocabulary learning through interactive conversations with virtual characters retained significantly more words compared to those watching videos. The authors attributed this gain to contextual clues and multimodal support as well as real-time interaction and feedback available in the VR environment.

Moreover, VR can be effective in fostering intercultural communicative competence. As Shadiev et al. (2023) observed, learners who created virtual tours of local attractions and shared them with international peers reported enhanced levels of intercultural competence. Similarly, Liaw (2019) found that learners who used an open VR application to interact with people from other countries benefited from the experience and demonstrated greater control in cross-cultural communication. These findings suggest that VR not only supports language development but also enhances learners' understanding of sociocultural norms and subtleties, which are essential for effective communication.

Foreign language anxiety

Speaking a foreign language can increase anxiety due, for example, to a learner's fear of being misunderstood or mocked because of their grammar mistakes or pronunciation issues (Macintyre, 2017). This can, understandably, have a strong demotivating effect on language learners who are vulnerable to such anxiety and may feel that speaking a foreign language is intimidating and face-threatening (Dörnyei, 2001). There have been a number of studies carried out to determine what kind of strategies can decrease foreign language anxiety (FLA) and improve learning (see Teimouri et al., 2019; Toyama & Yamazaki, 2021). There is some evidence that speaking a foreign language in a VR environment could help decrease FLA as an avatar provides a kind of protective barrier to a learner. As a result, they do not have to show their face and can remain anonymous (Yang et al., 2020). In addition, such anonymity may mean that learners can take on a fresh identity in a low-stakes and engaging 'third space' (Bhabha, 2012): such a new identity can decrease the fear of judgement and potentially decrease FLA. A third reason as to why VR may help decrease anxiety is the use of the technology to desensitise learners to fear, as has been successfully carried out in other fields such as paramedicine (Aiello et al., 2023). One example of such an approach in language learning is that of Gruber and Kaplan-Rakowski (2020) who used VR to support students making oral presentations. In a subsequent study, the same researchers (Kaplan-Rakowski & Gruber, 2023) found that VR outperformed Zoom in helping to alleviate levels of FLA. The main reason would appear to be that when using Zoom learners still need to interact more directly with other learners and that there is less anonymity when using such tools. As there are relatively few studies linking VR with FLA reduction and even fewer with Zoom, it would be fruitful, therefore, to compare the two platforms in the current study.

Research gaps in online flipped language learning

Research on online flipped learning environments has largely focused on the use of video conferencing tools such as Zoom. While these platforms enable synchronous interaction, they offer limited opportunities for immersive experiences. Few studies have investigated how more interactive and immersive platforms, such as VR or game-based environments, could enhance student engagement and learning outcomes in fully online flipped classroom models, and those that are available are either focused on disciplines other than language education or focus on language teacher education (e.g., Jong, 2023; Ronaghi, 2024).

Within this emerging area of research, however, the existing evidence remains unbalanced in pedagogical coverage, study scale and duration, and methodological focus. In a scoping review of recent studies published between 2020 and 2022 on VR in language learning, the researchers identified several gaps in the literature (Alizadeh & Cowie, 2022). First, most existing research on VR in L2 education has predominantly focused on vocabulary acquisition (e.g., Lai & Chen, 2021; Luan et al., 2024; Nicolaidou et al., 2021), while under-exploring communicative competencies such as interactional fluency and oral production skills in authentic or immersive contexts (Alizadeh & Cowie, 2024). As

language learning is inherently social and communicative, this lack of focus on interactional abilities represents a major limitation.

Second, the majority of studies in this area tend to rely on short-term, small-scale case studies, often involving a limited number of participants over brief intervention periods. As a result, aside from a few long-term studies such as Figueroa and Jung (2025), there is a lack of longitudinal and (quasi-)experimental research employing rigorous frameworks and designs that can evaluate the long-term impact of VR-based learning on larger student cohorts (Stracke et al., 2025; Zhao et al., 2022). Without such studies, it remains difficult to draw generalizable conclusions about the pedagogical value of VR for language learners.

Third, many existing VR studies are heavily reliant on subjective measures such as student self-reports and qualitative reflections. While these are valuable for capturing learner perceptions, they do not always offer robust evidence of measurable learning gains (Radianti et al., 2020). There is thus a need for more studies incorporating objective metrics, such as performance assessments or interaction analysis, to evaluate the effectiveness of VR-based instruction.

To address these research gaps, the present study examines the impact of VR on enhancing student engagement and learning outcomes, as well as its potential to reduce FLA within a fully remote flipped learning environment. While many of the studies cited in Alizadeh and Cowie (2022) touch on these factors, few maintain research rigor over a longitudinal period. To avoid such issues and enable meaningful comparison, this study included a control group using Zoom. It was assumed that because all the students had previous experience of video conferencing during the COVID-19 pandemic that the outcomes from a Zoom group could serve as a baseline, enabling the researchers to evaluate the added value of VR in enhancing student performance and engagement while reducing anxiety.

In doing so, the present study contributes to the growing literature on synchronous online language learning by directly comparing VR and Zoom-based flipped learning environments within the same pedagogical framework. Unlike many previous VR studies that have focused primarily on learner perceptions or isolated VR implementations, this study combines longitudinal performance data, engagement and anxiety measures, and qualitative feedback to provide a more comprehensive comparison of the two online learning environments. The findings therefore help address the limited empirical evidence regarding the relative pedagogical value of VR in comparison with more established synchronous video conferencing platforms.

Building on the findings outlined in the literature review, and in line with prior recommendations emphasizing the effectiveness of gamification in boosting student performance and engagement in online flipped learning environments (Phanwiriyarat et al., 2025), the present study introduces a novel approach that integrates VR into a flipped classroom model. The following research questions guided this study:

1. How do video conferencing and VR-based flipped learning environments impact student speaking performance?
2. How do students' engagement levels compare in video conferencing vs. VR environments?

3. How do students' FLA levels compare in video conferencing vs. VR environments?
4. How do students perceive the usability, benefits, and challenges of each platform?

The study

To answer the research questions, a quasi-experimental design was employed, involving two groups of volunteer students: a Zoom group and a VR group. Both groups participated in an online course focused on developing small talk skills. Small talk is defined as short initial conversations with other people that are important to develop new relationships. Previous studies have shown that without a set of established topics participants can find it difficult to keep conversations going (Spring et al., 2019); therefore, small talk was chosen as a theme for the study as none of the participants had ever met each other before and it seemed appropriate content for them to quickly discover common issues to discuss. The study was approved by the ethics committee at the first author's institution. Prior to the start of the study, all participants were informed about the purpose and procedures of the research and were assured that their participation was entirely voluntary. Written informed consent was obtained from all students, who were also advised that they could withdraw from the study at any time without penalty. Confidentiality was maintained throughout the research process, with all data anonymised during analysis and reporting to protect participants' identities.

The study was conducted in two phases, each corresponding to one of the participant groups. The project was first launched in 2022 with an initial study involving a comparison (Zoom) group, which ran for nine weeks from December 2022 to February 2023. This was followed by a second study involving an experimental (VR) group, which was conducted over eleven weeks from October 2023 to January 2024. It was not logistically possible for the two phases of research to be carried out during the same time period but there was no overlap of any of the participants. The following sections provide further details about the two studies, including information on participants, the learning context, and data collection procedures.

Participants

Two groups of participants took part in this project: a Zoom group and a VR group. The Zoom group initially comprised 37 undergraduate Japanese students enrolled at various universities in western Japan. Over the course of the study, some participants withdrew, resulting in a final sample of 30 students (19 females and 11 males; mean age = 20.07). The VR group began with 29 undergraduate Japanese students, four of whom withdrew during the study, leaving a final sample of 25 students (16 females and 9 males; mean age = 21.23).

Participants were recruited using criterion-based purposive sampling, whereby learners were intentionally selected from an open call because they met predefined criteria relevant to the study (Patton, 2015). These criteria included having Japanese as a first language, being an undergraduate university student, and possessing comparable English proficiency

levels, as indicated by TOEIC scores ranging from 400 to 600 or equivalent scores on other widely used proficiency tests. As they came from different universities and were largely unfamiliar with each other, this setting was particularly suitable for a course on small talk skills, enhancing the authenticity of the learning experience. The purpose of the study and the procedures were explained to all participants via Zoom, and informed consent was obtained prior to participation. Participants also agreed to receive compensation for their involvement at the end of the study period.

Learning context and platforms

The participants attended a remote online course hosted on Zoom and in VR for each group, respectively. The course followed a flipped learning approach, in which the students watched two videos before each weekly lesson on the topic of small talk in English and how to successfully create bonds with others. The videos were created by the second author and one co-creator and hosted on YouTube, and the links to individual videos were shared with the students on a Google Site created for the purpose of this study. As shown in Figure 1, during the one-hour synchronous online lessons, the students engaged in several rounds of group work. In addition, at the end of the lesson, they had two rounds of small talk practice in pairs in which they were asked to talk about any topic of their choice for five minutes.

For the Zoom group, breakout rooms with randomly assigned members were used to divide the students into groups/pairs. The session in the main room and up to 10 parallel breakout rooms were recorded locally through multiple computers running Zoom (Figure 2) and later trimmed using TechSmith Camtasia®. The students' channels of communication were audio and video. As a general rule, they were asked to keep their camera on all the time. As shown in Figure 1, quizzes to check understanding were given using the Zoom polling function.

The roles of the two researchers/teachers were quite distinct during the online lessons. The first author ran the administrative parts of the lesson right at the beginning and end of the sessions (items one and eight in Figure 1) and monitored the 10 parallel breakout rooms with the recording computers. She was also responsible for troubleshooting any problems that occurred during the lesson. The second author was responsible for facilitating the content of the lesson concerning small talk (items two to seven in Figure 1).

Figure 1

Sample lesson plan on small talk skills

Small Talk Lesson 5
January 18th, 2023, 10:00 – 11:00 am

Learning Goals:
- The students will learn how to make their small talk more interesting by using adjectives and adverbs.
- They will also learn how to be a good listener in small talk interactions using techniques such as echoing and back-channelling.

10:00 – 10-05 Welcome
The first five minutes are for greetings and house-keeping announcements if any. **1**

10:05 – 10:10 Easy Warm-up
Students chat with each other in small groups for five minutes. **2**

10:10 – 10:15 Video 1 Quiz
The students respond to some questions on the topic of the first video to check their understanding. The teacher adds some explanations to further deepen their understanding. **3**

10:15 – 10:25 Guided Group Discussion
The students exchange ideas in small groups in reference to the guiding question(s) provided by the teacher. **4**

10:25 – 10:30 Video 2 Quiz
Step 3 is repeated for the second video. **5**

10:30 – 10:40 Guided Group Discussion
Step 4 is repeated for the second video. **6**

10:40 – 10:55 Small Talk Practice in Pairs
The students practice small talk in pairs with two different partners randomly assigned in two rounds. **7**

10:55 – 11:00 Closing
The students are reminded of the videos for the next lesson and then leave Zoom. **8**

Figure 2

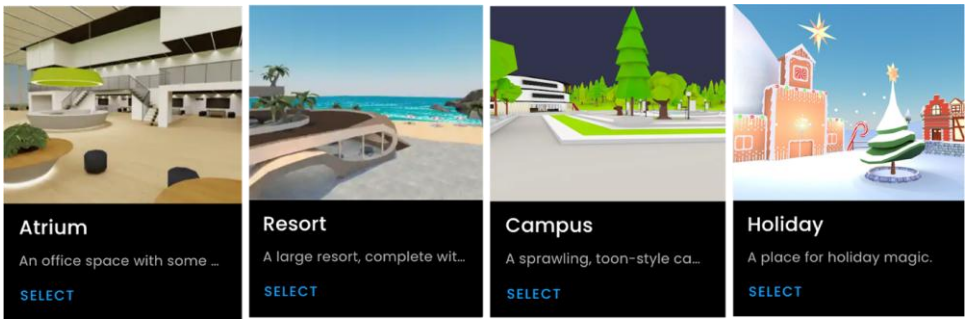
Technical setup with 10 PCs used for screen recordings

The VR group met on a WebVR platform called Frame (https://framevr.io/) that runs on a browser regardless of the device the students have, including PC, mobile phone and VR headset. Since a use-your-own-device policy was in place, students entered the VR environment using their own PCs or phones, devices that most of them already owned. Different environments were deployed for the VR group study, including an atrium, a beach resort, a campus, and a winter holiday scene (see Figure 3). Several factors had to be considered when selecting an environment on Frame, including the space's capacity to accommodate around 30 users, its simplicity (easy to navigate and not maze-like), and its performance rating (environments requiring excessive processing power and containing too many assets had to be avoided to ensure accessibility for all students). These design

choices were intended to ensure equitable access and to minimise technology-related confounds when comparing VR-based interaction with Zoom-mediated communication. The private voice zones on Frame were recorded from the perspective of dummy avatars, named TA1 to TA10, using TechSmith Snagit® version 2023 installed on multiple computers, similar to the setup used in the Zoom study.

Figure 3

Different Frame environments used in this study



After an initial orientation week, the participants were taught in week two how to enter Frame, navigate the environment, make their custom-created avatar and use chat, mic and view controls. All these instructions were also made available to the students on YouTube in the form of screencasts captioned in their first language, Japanese, to avoid any language barriers in understanding the basic functions of the VR environment.

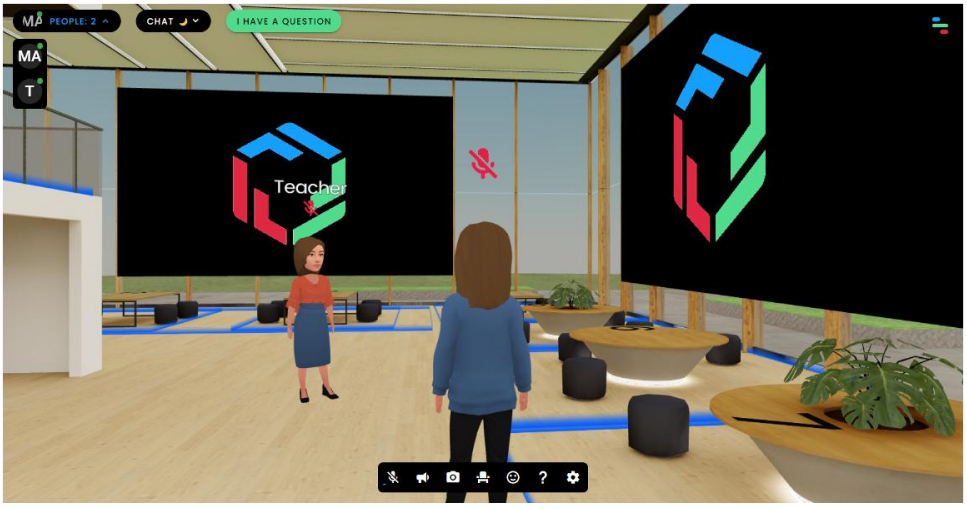
For group and pair work, private voice zones (similar to Zoom breakout rooms) were used. These private voice zones allow users to hold conversations with other users who are in the same zone but not hear avatars in other zones. Figure 4 displays the beach resort environment with 14 numbered private voice zones. These zones were customizable in size and location and could be toggled on or off by the instructor during whole-class and group/pair activities. Audio was the only channel of communication for the VR group, but they could see each other's avatars with the names appearing on top of each avatar head (refer to Figure 5).

Figure 4

Private voice zones in the beach resort environment

Figure 5

Two avatars in the atrium environment



Data collection procedures

To address the four research questions, data were collected from participants through performance assessments, surveys, and focus group interviews. First, to examine learning outcomes, student interactions in pairs were recorded as they engaged in small talk activities at the end of each lesson. To ensure objective evaluation, a holistic performance rubric was developed and shared with two external raters (EFL teacher colleagues from the second author's university). The rubric served as a structured scoring guide to support raters' overall judgments of learners' small talk speaking performance. The rubric featured a 5 × 5 grid comprising five performance criteria: (1) fluency and coherence, (2) lexical resources, (3) grammatical range and accuracy, (4) pronunciation, and (5) interaction and communication strategies (see Appendix 1). These criteria were not treated as independent outcome variables, but rather as reference dimensions to ensure scoring consistency and

objectivity across performances. Each performance received a single aggregated score, ranging from a minimum of five to a maximum of 25.

To further standardise the rating process, the two raters participated in an initial norming session with the researchers, during which the rubric items were thoroughly reviewed, and a sample dataset was shared for test rating. Following the norming process, the raters independently assessed three sets of student interactions at the beginning, middle, and end of the course to track changes in performance over time. These sets of numerical data were subsequently used for both within-group and between-group comparisons. Additionally, to establish inter-rater reliability, the raters convened after the rating process to discuss discrepancies and re-evaluate videos where there were significant differences in scoring. The final corrected scores resulting from this process were then used to calculate inter-rater reliability. Inter-rater reliability for the Zoom group's pre-, mid-, and post-test scores was calculated using intra-class correlation coefficients (ICC) in SPSS for Windows (Version 26; IBM Corp.) resulting in values of .87, .83, and .83, respectively. For the VR group, the ICC values were .86 for the pre-test, .91 for the mid-test, and .85 for the post-test. All these inter-rater reliability scores exceed .80, a threshold generally considered acceptable for educational research (Koo & Li, 2016).

In addition to the performance data, survey data were collected at the end of each study to measure and compare students' engagement and anxiety levels across groups. To assess engagement, an adapted version of the engagement scale developed by Sun and Rueda (2012) was utilised. Based on Fredricks et al.'s (2004) multidimensional model of engagement, the scale included 19 items divided into three subcategories: behavioural engagement (five items), emotional engagement (six items), and cognitive engagement (eight items). The original five-point Likert scale structure (ranging from one = strongly disagree to five = strongly agree) was retained, although some items were slightly modified to better align with the context of the current study.

To ensure clarity and minimise language barriers, the engagement survey was translated into Japanese by one author, then double-checked by the other and a native Japanese speaker, also a language teaching researcher. A key modification in the Japanese version involved changing the tense of the survey items from present to past. This adjustment was made to avoid confusion among participants, who might otherwise interpret present-tense items as referring to any online course rather than the specific online flipped course under investigation. The engagement survey was administered online via Microsoft Forms.

The study also employed the Japanese version of the Foreign Language Classroom Anxiety Scale (FLCAS), originally developed by Horwitz et al. (1986) and translated by Yashima et al. (2009). Like the engagement survey, the anxiety scale was administered online through Microsoft Forms at the end of each study.

Finally, following a mixed methods research design (Riazi & Candlin, 2014), the researchers conducted semi-structured focus group interviews to triangulate the findings. Students were interviewed in groups of three to five. All interviews were conducted in Japanese, recorded via Zoom by the first author, and each lasted approximately one hour. Participants were asked about various aspects of the course, organised into five pre-defined categories: overall impressions of the course, the impact of video materials, the

effectiveness of group and pair work, perceived learning outcomes, and technical challenges. Students in the VR group were also asked additional questions related specifically to the VR learning environment. The interviews were transcribed using Sonix (https://sonix.ai/). The qualitative analysis followed a deductive thematic approach, in which the pre-defined interview categories were used as the initial analytical framework. The authors reviewed the transcripts several times and coded excerpts according to recurring themes within each category. Where possible, percentages and representative quotes are included to support the findings.

Findings

RQ1: How do video conferencing and VR-based flipped learning environments impact student speaking performance?

Table 1 presents the descriptive statistics (means and standard deviations) for the two treatment groups (Zoom and VR) across the three measurement points (pre-test, mid-test, and post-test). Both groups showed a gradual increase in mean scores over time, with similar patterns of improvement.

Table 1
Means and standard deviations for Zoom and VR groups

Measure	Zoom Group M (SD)	VR Group M (SD)
Pre-test	17.88 (2.54)	17.50 (1.60)
Mid-test	18.46 (2.72)	18.54 (2.05)
Post-test	20.02 (1.80)	19.83 (1.40)

The null hypothesis stated that there would be no significant difference in speaking skills between students completing the flipped online course via Zoom and those completing it via VR over time. To test this hypothesis, a repeated measures analysis of variance (ANOVA) with Bonferroni adjustment was conducted using SPSS to examine the effects of Time (pre-test, mid-test, post-test) as a within-subjects factor and Treatment (Zoom vs. VR) as a between-subjects factor on participants' speaking scores.

As shown in Table 2, the results of the repeated measures ANOVA revealed a significant main effect of Time, $F(2, 106) = 45.02$, $p < .001$, indicating that students' speaking skills improved significantly over the duration of the study. Effect sizes were interpreted following the recommendations of Norouzian and Plonsky (2018). The main effect of Time was large, accounting for approximately 46% of the total variance in scores ($\eta^2 = .457$; partial $\eta^2 = .459$), reflecting substantial overall improvement over time.

The main effect of Treatment was not significant, $F(1, 53) = 0.11$, $p = .74$, with a negligible effect size ($\eta^2 = .002$; partial $\eta^2 = .002$), indicating no meaningful overall difference between the Zoom and VR groups when scores were averaged across time. The Time × Treatment interaction was also not significant, $F(2, 106) = 0.46$, $p = .63$. This

interaction explained less than 1% of the total variance ($\eta^2 = .005$; partial $\eta^2 = .009$), suggesting that the pattern of change over time did not differ meaningfully between groups.

Bonferroni-adjusted pairwise comparisons further indicated significant improvement in small talk speaking skills across all time points. Participants showed significant gains from the pre-test to the mid-test, from the pre-test to the post-test, and from the mid-test to the post-test (see Table 3). These findings provide additional support for a strong and consistent effect of Time.

Table 2

Repeated measures ANOVA results

Effect	F	Df	p	η^2	Partial η^2
Time	45.02	2, 106	<.001	.457	.459
Treatment	0.11	1, 53	.74	.002	.002
Time × Treatment	0.46	2, 106	.63	.005	.009

Table 3

Pairwise comparisons of time points with Bonferroni adjustment for multiple comparisons

Comparison	Mean Difference	Sig.	95% Confidence Interval for Difference [Lower Bound, Upper Bound]
Pre-test – Mid-test	−0.81	.006	[−1.42, −0.20]
Pre-test – Post-test	−2.23	< .001	[−2.74, −1.71]
Mid-test – Post-test	−1.42	< .001	[−2.05, −0.79]

Note. Negative mean differences indicate improvement over time.

RQ2: How do students' engagement levels compare in video conferencing vs. VR environments?

The students completed an engagement survey at the end of the study period, and the results are presented in Figure 6. A visual inspection of the data suggests that, aside from a few items such as CE3 and CE5, both the Zoom and VR groups were highly engaged in the online courses. Items CE3 and CE5 asked students whether they took the initiative to search for additional materials, such as other YouTube videos, related to the course content. Given the nature of the course and the voluntary participation of the students, it was not surprising that these items received relatively low endorsement.

The internal consistency of the engagement scale was assessed after re-coding the negatively worded items (BE2, BE3, and EE6) using Cronbach's alpha, calculated separately for each group. Alpha coefficients were .81 for the Zoom group and .87 for the VR group, indicating good reliability (Pallant, 2016).

In addition, to examine whether there were statistically significant differences between the two groups, the 19 items of the engagement scale were recomputed into three subscales representing behavioural, emotional, and cognitive engagement, in accordance with the original scale design. Mann–Whitney U tests were then conducted using the mean subscale scores. The results indicated no statistically significant differences between the groups

across any of the three engagement dimensions. This suggests that both groups were equally engaged.

RQ3: How do students' foreign language anxiety levels compare in video conferencing vs. VR environments?

In addition to engagement, students' anxiety levels in both groups were measured using the Japanese version of the Foreign Language Classroom Anxiety Scale. Figure 7 presents the mean ratings for each item across the two groups. Similar to the engagement scale, the internal consistency of the anxiety scale was assessed after re-coding the negatively worded items (Items 2, 5, 8, 11, 14, 18, 22, 28, and 32) using Cronbach's alpha, calculated separately for each group. The analysis yielded alpha coefficients of .93 for the Zoom group and .90 for the VR group, indicating very good reliability (Pallant, 2016).

The anxiety questionnaire items were aggregated into a single composite score, following the original scale design. Although the FLCAS was originally conceptualised around communication apprehension, test anxiety, and fear of negative evaluation (Horwitz et al., 1986), it was validated as a unidimensional measure. The anxiety scale was therefore treated as a single construct rather than reporting sub-scale scores. Scores were interpreted such that lower values indicated lower levels of anxiety. A Mann–Whitney U analysis of the mean composite scores did not reveal any statistically significant differences between the two groups, suggesting that students in the Zoom and VR groups experienced comparable levels of anxiety.

Figure 6

Mean values of the Zoom and VR group engagement rankings

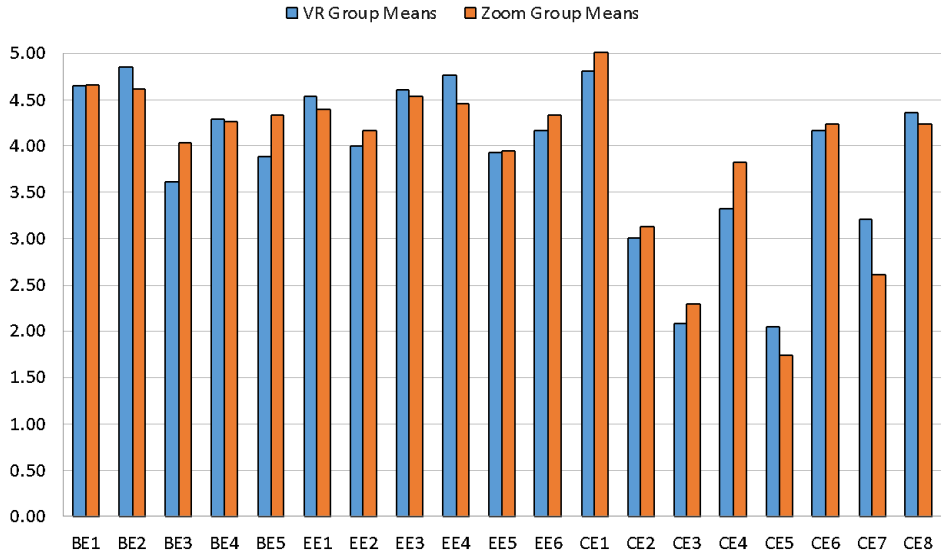
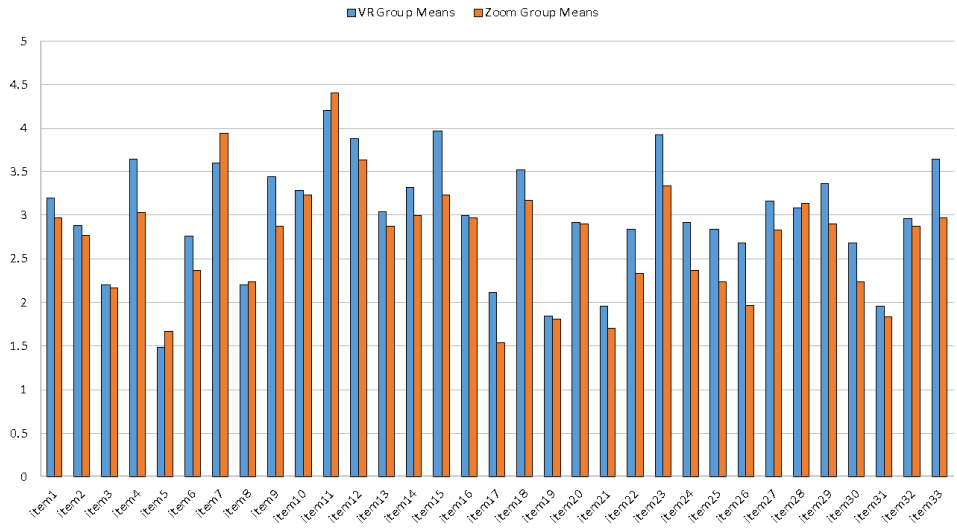


Figure 7

Mean values of the Zoom and VR group anxiety rankings



RQ4: How do students perceive the usability, benefits, and challenges of each platform?

This section presents findings from the focus group interviews. Participants were asked a range of questions, which were broadly categorised into the following areas: overall impressions of the course, the video materials, group and pair work, perceived learning outcomes, and technical challenges. In addition, students in the VR group responded to specific questions related to their sense of immersion and presence, as well as their experiences using avatars. A summary of their responses is presented below.

Regarding the course overall, most students reported high levels of interest and motivation (VR: 72%; Zoom: 80%), appreciating the rare opportunity to study small talk in an all-English course with peers from different universities and backgrounds. Students frequently contrasted this experience with more exam-focused classes (e.g., "*I usually study grammar or reading, but here I could actually talk.*"). Reports of boredom were rare, with participants instead highlighting the interactive structure and frequent speaking opportunities as engaging aspects of the course. The participants in both groups also demonstrated cognitive engagement through high quiz participation (VR: 92%; Zoom: 77%), reporting quizzes as easy to manageable when preparatory videos were watched in advance. Many students in the VR group emphasised the novelty of the learning environment, noting that this was their first experience attending classes in a VR setting and described avatar-based interaction as engaging and enjoyable (e.g., "*Using an avatar and talking to different people in VR was fun.*").

In terms of the video materials, most participants across both conditions reported regular but not perfectly consistent viewing before class (over 80% for both groups), often watching videos shortly before the lesson as preparation. The videos were generally evaluated as useful and easy to follow, particularly for understanding what small talk is,

learning practical expressions, and preparing for in-class interaction (e.g., "*It was interesting because I'd never learned small talk strategies like that.*"; "*If I watched the video, I could follow the class more easily.*"). Learner strategies varied when encountering unfamiliar vocabulary or concepts. Some paused to look up items immediately, while others relied on context or checked later, reflecting different approaches to managing cognitive load. Playback speed adjustment was common, with most watching at normal speed and some increasing speed for efficiency.

Many students stated that group and pair activities were successful (approximately 90% and 85% for the Zoom and VR groups, respectively). However, participation levels varied, with many participants reporting alternating roles between active contributors and more passive listeners who engaged through back-channelling, especially in larger groups. Across both conditions, pair work was consistently described as more balanced and easier to manage than group work, particularly for turn-taking and sustaining conversation. Leadership was largely situational rather than fixed, typically emerging in response to silence or confusion (e.g., "*At first no one spoke, so I decided to start asking questions.*"). In the VR group, however, the absence of facial expressions and recurring technical issues (e.g., audio delays, entry problems) occasionally made it difficult to judge partners' willingness to speak, leading to uncertainty about whether silence reflected interactional reluctance or technical failure.

With regard to perceived learning outcomes, participants in both the Zoom and VR groups emphasised the course's role in improving their small talk skills. Most participants reported that one of the most noteworthy outcomes was discovering what small talk is and how it functions, including how to initiate conversations and sustain interaction through follow-up questions. As one participant noted, "*I realised small talk isn't just chatting. It helps build rapport and leads into deeper talk.*" Many participants perceived moderate improvement in small talk ability, particularly in reduced anxiety, quicker responses, and greater willingness to speak, even when vocabulary or grammar gains were described as limited (e.g., "*I'm still not perfect, but I'm not afraid to start talking anymore.*").

As noted earlier, VR participants were asked additional questions about the VR environment, including their sense of immersion, presence, and the use of avatars for communication. Responses generally fell into two categories: VR enthusiasts and VR sceptics. VR enthusiasts appreciated the gamified learning experience and felt a sense of shared space with their peers (e.g., "*Because I had to move and operate in VR, I concentrated more than in Zoom.*"; "*It felt like we were sharing the same space.*"). They also valued the ability to personalise their identity through avatars. Several noted that the lack of visible faces reduced self-consciousness and encouraged speaking, with avatars lowering anxiety and making participation feel easier (e.g., "*Because my face wasn't shown, I could speak more casually without fear.*"). Sceptics, on the other hand, were less enthusiastic, citing concerns about the technology's current limitations and lack of user-friendliness. They emphasised interactional constraints, particularly the absence of facial expressions, which made it harder to judge turn-taking, attentiveness, or listener reactions (e.g., "*Without facial expressions, it was hard to know if the other person was listening.*").

Finally, in terms of technical challenges, both groups occasionally faced network issues when accessing online lessons from home or other locations. Some VR group members also experienced device incompatibility, for example, being unable to perform certain functions on their phones that were available on PCs and expressed concerns about audio disruptions. As one participant explained, "*Sometimes we could hear other groups, and it was confusing who was talking*," while another noted that "*there were sometimes audio issues, so we had to re-enter, and our speaking time got shorter*." In fact, for approximately three weeks during the course, serious audio interference was reported, although this issue was eventually resolved.

Discussion

To summarise the findings, a significant main effect was observed for time, indicating that participants' small talk skills improved over the duration of the course. However, no significant main effect was found for treatment; in other words, the platform used (Zoom vs. VR) did not lead to meaningful differences in learning outcomes. Participants in both groups were equally engaged in general, with no significant differences in engagement scores. A similar pattern was observed in reported levels of FLA. These results suggest that well-designed courses delivered via either Zoom or VR can effectively enhance learning outcomes, sustain student engagement, and manage anxiety levels.

From a theoretical perspective, these findings are noteworthy given that Zoom and VR differ substantially in terms of learners' familiarity with the platform, cognitive load, social presence and embodiment, factors that have been widely assumed to influence speaking performance, engagement, and anxiety in online language learning. On the one hand, Zoom represented a familiar environment for most learners and likely imposed lower cognitive demands due to its relatively intuitive interface and students' prior experience using video conferencing during the COVID-19 pandemic. On the other hand, VR offered a more immersive and embodied learning experience that may have enhanced learners' sense of social presence and engagement. However, its unfamiliar interface, steeper learning curve, and occasional technical difficulties may also have increased cognitive load and reduced some of these potential advantages. The largely similar outcomes observed across the two conditions therefore suggest that the respective advantages and disadvantages of each platform may have partially offset one another. While VR may provide increased immersion and embodiment, these benefits may have been counterbalanced by the additional cognitive and technical demands associated with navigating an unfamiliar environment.

The present study contributes to the growing literature on synchronous online language learning by addressing a relatively underexplored area of research, i.e., the direct comparison of VR and video conferencing-based flipped learning environments within the same pedagogical framework. While previous studies have often examined VR in isolation or focused primarily on learner perceptions, the current study combined longitudinal speaking performance measures with engagement, anxiety, and qualitative data to provide a more comprehensive comparison of the two platforms. The findings suggest that

immersive VR environments are not necessarily superior to commonly used video conferencing tools such as Zoom in promoting language learning gains. At the same time, the study indicates that VR may offer a distinct learning experience shaped by factors such as embodiment and immersion, particularly in contexts where these affordances meaningfully transform interaction and communication. These findings therefore contribute to a more balanced understanding of both the pedagogical potential and practical limitations of VR in online language education.

This study aligns with the findings reported by Dhimolea et al. (2022), who reviewed eight studies examining the impact of VR on oral communication skills. Of these, only one study (Yang et al., 2020) reported statistically significant improvements in the computer-generated VR group compared to the control group. Dhimolea et al. further concluded that, overall, there was limited evidence to suggest that VR use led to substantial language learning gains, with the notable exception of short-term vocabulary retention. Instead, their review found more consistent support for the idea that students perceived VR as fun, enjoyable, and motivating, although these positive emotional responses did not consistently lead to measurable improvements in language learning.

In the current study, it is noteworthy that students in the VR group demonstrated continuous improvement in measured small talk speaking performance across all testing points, with no decline in scores over time, indicating no negative learning outcomes despite variation in students' attitudes toward the VR environment. However, when designing remote learning courses, educators should consider whether the immersive features of VR are essential for achieving learning objectives. If they are not, more conventional platforms such as Zoom may offer comparable effectiveness at a lower cost and with fewer technical demands.

Because this study employed a comparative research design, care was taken to control for intervening variables. To this end, instructional activities provided to the Zoom group were closely replicated in the VR environment to ensure consistency. This meant that in practice there were few clear differences between the Zoom and VR teaching approaches. Nevertheless, future applications of VR in language education may be more justified when the technology can go beyond "substitution" or "augmentation" and could instead "modify" or "redefine" the learning experience, terms derived from Puentedura's (2006) SAMR model. In such cases, VR may offer unique affordances that could transform the language learning experience and justify the investment of time, resources, and training. Such affordances could include one-to-one interactions with AI-generated avatars that will be able to give a personalised experience and individual feedback to learners on their speaking performance (Robert et al., 2025).

With respect to learner engagement, the results indicate that students in both the Zoom and VR groups were highly engaged overall, with no statistically significant differences observed between the two platforms across behavioural, emotional, or cognitive engagement dimensions. This finding reinforces the view that engagement in online language learning is driven primarily by pedagogical factors, such as task relevance, opportunities for interaction, and clarity of course structure, rather than by technological affordances alone. This interpretation is consistent with Dixson's (2015) conceptualization

of online engagement, which emphasises social presence, community, and meaningful interaction as key determinants of engagement.

Regarding FLA, the results similarly showed no statistically significant differences between the Zoom and VR groups in overall anxiety levels, as measured by the composite Foreign Language Classroom Anxiety Scale score. These findings indicate that participation via VR did not lead to systematically higher or lower levels of language anxiety compared to Zoom, despite differences in platform characteristics such as embodiment, anonymity, and familiarity.

The lack of between-group differences in anxiety suggests that potential anxiety-reducing features of VR such as the use of avatars may have been counterbalanced by other factors, including increased cognitive load, unfamiliar controls, or uncertainty during peer interaction. Conversely, while Zoom offers a familiar and low-barrier environment, the direct visibility of oneself and others may sustain evaluative pressure for some learners. The comparable anxiety levels observed across platforms, therefore, point to a complex interaction between technological affordances and learner perceptions, rather than a straightforward advantage of one platform over the other. These findings are in line with Spring et al. (2019), who argued that affective outcomes in video-synchronous interaction depend largely on pedagogical decisions such as appropriate partner pairing, explicit training in interactional strategies, and sustained instructional scaffolding.

This interpretation is consistent with prior research showing mixed effects of VR on anxiety. Although prior research has often found that VR can reduce anxiety in performance-based situations such as public speaking (Chen, 2022; Wu et al., 2021), evidence is not uniformly positive. For example, Wu and Hung (2022) reported significantly higher anxiety levels in a VR group when anxiety was measured as a sub-construct of willingness to communicate. This suggests that while VR may lower anxiety in some contexts, it can also exacerbate it in others, particularly when users are navigating new communicative demands in an unfamiliar environment. Another possible explanation is that studies showing VR reduces anxiety in foreign language learning are often based on participants working autonomously in the VR environment. In this study participants were interacting with each other which may explain the discrepancy.

In sum, while VR shows strong potential as an engaging alternative to traditional video conferencing tools like Zoom, its effects on language anxiety appear to be more complex. For some learners, the immersive nature of VR may enhance emotional engagement, but for others, it may introduce or exacerbate anxiety. This was also noted in the focus group interviews, where some students expressed discontent with the absence of facial expressions in avatars. Combined with technical issues such as internet disconnections, this left them unsure about the status of their partners.

Regarding scaffolding and training in VR environments, the researchers noticed that some students did not properly place their avatars in a way that would allow them to face their partners. Although this may seem intuitive in video conferencing, learners may need more training than originally anticipated in controlling their avatars in VR environments. In addition, such training can help VR sceptics (those uncertain about the benefits of VR) recognise its potential as a suitable platform for their learning and better leverage its

advantages. These findings highlight the importance of providing sufficient orientation, support, and scaffolding when implementing VR in language learning contexts (Cowie & Alizadeh, 2022).

This study had several limitations. First, the participants were recruited through purposive volunteer sampling rather than random selection, which limits the generalisability of the findings. Students who volunteered to participate may have already held more positive attitudes toward technology, online learning, or VR-based interaction than the broader student population. In particular, learners who were more open to experimenting with immersive technologies may have been more willing to join the VR condition, creating the possibility of self-selection bias. As a result, the findings may not fully represent students with lower technological confidence, lower motivation, or more negative perceptions of VR environments. Future studies should therefore attempt to recruit larger and more diverse samples across multiple institutions and, where possible, implement random assignment procedures.

Second, technical issues may also have influenced the findings, particularly in the VR condition. Issues such as device incompatibility, unstable internet connections, and inconsistent audio quality occasionally disrupted participation and communication. Furthermore, during the initial three weeks of the project, the VR platform suffered from particularly poor reception and performance problems, which may have negatively affected student engagement, motivation, and interaction patterns. These technical difficulties may therefore have influenced both the learning experience and participants' perceptions of the VR environment itself. Future studies could address these issues by selecting less resource-intensive platforms that function reliably across a wider range of devices and network conditions. Ensuring stable and uninterrupted audio communication would also be especially important in VR-based language learning environments where spoken interaction is the main mode of communication.

Finally, it should be acknowledged that as the Zoom and VR groups were taught in different academic years the claims this study makes about Zoom versus VR should be interpreted cautiously, as cohort differences and contextual factors unrelated to the treatment itself may have influenced the results. Nevertheless, the claims that there are clear within group improvements over time can be accepted with some confidence.

Conclusion

This study examined and compared the learning outcomes, levels of engagement, and foreign language anxiety experienced by students in two flipped learning courses conducted via Zoom and in VR. Overall, the results indicated that students in both groups demonstrated improvement in their small-talk speaking skills over time. However, no statistically significant differences in learning gains were observed between the two groups. Furthermore, participants from both learning environments reported similar levels of engagement and anxiety, suggesting that both Zoom and VR were effective in sustaining student interest and engagement while helping to manage foreign language anxiety. These findings suggest that both Zoom and VR can support student engagement while helping to

manage foreign language anxiety, although the results should be interpreted with caution because the two groups were studied in different academic years.

In the future, the researchers aim to build on these findings through an international collaborative project. This initiative will involve Japanese students working with peers from overseas institutions, with the goal of further enhancing intercultural communication skills through technology-mediated learning environments.

Appendices

Appendix 1

Performance assessment rubric

	1- Not able to perform	2- Inadequate	3- Needs improvement	4- Meets expectations	5- Exceeds expectations
Fluency and coherence	Speaks very little	Speaks with a lot of hesitation which interferes with communication	Speaks with occasional hesitation which can interfere with communication	Speaks with some hesitation but it does not interfere with communication	Speaks very smoothly
Lexical resources	Very limited vocabulary	Limited vocabulary	Can talk about familiar topics but has limited flexibility	Good range of vocabulary and can paraphrase well. The occasional inappropriate choice	Wide range of vocabulary. Good use of idiomatic language
Grammatical range and accuracy	Limited success at basic sentences	Basic structures and frequent errors	Basic structures and frequent errors	Good variety of structures and occasional errors	Excellent variety of structures and almost no errors
Pronunciation	Unable to be understood	Listener frequently has difficulty understanding	Listener occasionally has difficulty understanding	Listener seldom has difficulty understanding	Listener has no difficulty understanding
Interaction and communication strategies	Cannot interact with partner	Poor at interacting	Tries to interact but needs a lot of help communicating	Responds appropriately most of the time and tries to interact positively	Interacts very smoothly, initiating, asking follow-up questions

Appendix 2

Engagement scale items adapted from Sun and Rueda (2012)

BE1. I followed the rules of the online class.
BE2. I sometimes skipped the online lessons for no good reason.*
BE3. When I was in the online class, I pretended I understood what to do.*
BE4. I was able to consistently pay attention when I was taking the online class.
BE5. I watched the pre-lesson videos on time.
EE1. I liked taking the online class.
EE2. I felt excited at the online class.
EE3. The online classroom was a fun place to be.
EE4. I was interested in the learning activities of this online class.
EE5. I felt happy when taking this online class.
EE6. I felt bored by the online class.*
CE1. I answered the quizzes during the online lessons.
CE2. I studied at home.
CE3. I tried to look for some course-related information or materials on other resources such as YouTube or the Internet.
CE4. After watching the videos, I answered the quiz questions on the class website to make sure I understood what they were about.
CE5. I looked for and studied extra materials to learn more about things we did in the online class.
CE6. If I did not know about something when I was learning in the online class, I would do something to figure it out.
CE7. If I did not understand what I learned online, I would go back to watch the video and learn it again.
CE8. I found the online course so interesting that I want to recommend it to friends or others who may want to learn about small talk in English if possible.
* = negatively worded items

Abbreviations

FLA: Foreign Language Anxiety; FLCAS: Foreign Language Classroom Anxiety Scale; ICC: intra-class correlation coefficients; SAMR: Substitution, Augmentation, Modification, Redefinition; VR: Virtual Reality.

Authors' contributions

Mehrasa Alizadeh: Conceptualisation, Investigation, Formal analysis, Writing - original draft, review and editing.
Neil Cowie: Investigation, Formal analysis, Writing - review and editing.

Authors' information

MA is Associate Professor at Otemon Gakuin University in Ibaraki, Osaka, Japan, where she teaches English as a foreign language to Japanese students. Her research interests include mobile learning, immersive learning, and the use of virtual reality in language education. She has published and presented extensively on immersive learning in international conferences and journals.
NC was an EFL teacher in Japan for many years before moving to New Zealand. He has published mainly in the areas of language teaching motivation and online learning. He is now working in the English Language Academy at the University of Auckland.

Funding

This research is funded by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (Project Numbers: 22K13756 & 25K16379).

Availability of data and materials
Not applicable

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
Mehrasa Alizadeh: Institute of General Education, Otemon Gakuin University, Ibaraki, Osaka, Japan
Neil Cowie: English Language Academy, University of Auckland, New Zealand

Received: 11 October 2025 Accepted: 21 August 2026
Published online: 1 January 2027 (Online First: 8 September 2026)

References

Aiello, S., Cochrane, T. & Sevigny, C. (2023). The affordances of clinical simulation immersive technology within healthcare education: A scoping review. *Virtual Reality, 14*, 1–19. https://doi.org/10.1007/s10055-022-00745-0

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education, 126*, 334–345. https://doi.org/10.1016/j.compedu.2018.07.021

Allcoat, D., & von Mühlenen, A. (2018). Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology*, *26*, 2140. https://doi.org/10.25304/rlt.v26.2140

Al-Zahrani, A. M. (2024). Enhancing postgraduate students' learning outcomes through flipped mobile-based microlearning. *Research in Learning Technology*, *32*. https://doi.org/10.25304/rlt.v32.3110

Alizadeh, M., & Cowie, N. (2022). Language learning and virtual reality: A scoping review. In S. Wilson, N. Arthars, D. Wardak, P. Yeoman, E. Kalman, & D.Y.T. Liu (Eds.), *Reconnecting relationships through technology.* Proceedings of the 39th International Conference on Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education, ASCILITE 2022 in Sydney: e22258. https://doi.org/10.14742/apubs.2022.258

Alizadeh, M., & Cowie, N. (2024). Investigating the impact of online learning platforms on student engagement and learning outcomes: Comparing Zoom with VR. In T. Cochrane, V. Narayan, E. Bone, C. Deneen, M. Saligari, K. Tregloan & R. Vanderburg (Eds.), *Navigating the Terrain: Emerging frontiers in learning spaces, pedagogies, and technologies.* Proceedings ASCILITE 2024. Melbourne (pp. 437–441). https://doi.org/10.14742/apubs.2024.1180

Alizadeh, M., & Cowie, N. (2025). Technology-enhanced intercultural exchange in language learning and teaching: A scoping review. In S. Barker, S. Kelly, R. McInnes & S. Dinmore (Eds.), *Future Focussed. Educating in an era of continuous change*. Proceedings ASCILITE 2025. Adelaide (pp. 487–492). https://doi.org/10.65106/apubs.2025.2698

Bahari, A. (2021). Affordances and challenges of teaching language skills by virtual reality: A systematic review (2010–2020). *E-Learning and Digital Media*, *19*(2), 163–188. https://doi.org/10.1177/20427530211036583

Baig, M. I., & Yadegaridehkordi, E. (2023). Flipped classroom in higher education: A systematic literature review and research challenges. *International Journal of Educational Technology in Higher Education*, *20*, 61. https://doi.org/10.1186/s41239-023-00430-5

Bhabha, H. K. (2012). *The Location of Culture* (2nd ed.). Routledge. https://doi.org/10.4324/9780203820551

Bui, T. T. H., Kaur, A., & Hung, T. V. (2022). A flipped classroom to personalise learning for engineering students. *International Journal of Learning Technology*, *17*(2), 97–114. https://doi.org/10.32664/ijitgeb.v6i2.155

Chen, J. C. (2016). The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. *Computers & Education*, *102*, 152–171. https://doi.org/10.1016/j.compedu.2016.08.004

Chen, Y. C. (2022). Effects of technology-enhanced language learning on reducing EFL learners' public speaking anxiety. *Computer Assisted Language Learning*, *37*(4), 789–813. https://doi.org/10.1080/09588221.2022.2055083

Chen Hsieh, J. S., Wu, W. C. V., & Marek, M. W. (2017). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, *30*(1–2), 1–21. https://doi.org/10.1080/09588221.2015.1111910

Chun, D. M., Karimi, H., & Sañosa, D.J. (2022). Traveling by headset: Immersive VR for language learning. *CALICO Journal*, *39*(2), 129–149. http://dx.doi.org/10.1558/cj.21306

Cowie, N., & Alizadeh, M. (2022). The affordances and challenges of Virtual Reality for language teaching. *International Journal of TESOL Studies, 4*(3) 50–65. https://doi.org/10.46451/ijts.2022.03.05

Deng, X. & Yu, Z. (2022). A systematic review on the influence of virtual reality on language learning outcomes. *International Journal of Online Pedagogy and Course Design*, *12*(1), 1–18. https://doi.org/10.4018/IJOPCD.302083

Dhimolea, T.K., Kaplan-Rakowski, R. & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends*, *66*, 810–824. https://doi.org/10.1007/s11528-022-00717-w

Divjak, B., Rienties, B., Iniesto, F., Vondra, P., & Žižak, M. (2022). Flipped classrooms in higher education during the COVID-19 pandemic: Findings and future research recommendations. *International Journal of Educational Technology in Higher Education, 19*(1). https://doi.org/10.1186/s41239-021-00316-4

Dixson, M. D. (2015). Measuring student engagement in the online course: The Online Student Engagement Scale (OSE). *Online Learning*, *19*(4). https://doi.org/10.24059/olj.v19i4.561

Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667343

Fischer, I. D., & Yang, J. C. (2022). Flipping the flipped class: Using online collaboration to enhance EFL students' oral learning skills. *International Journal of Educational Technology in Higher Education, 19*(1). https://doi.org/10.1186/s41239-022-00320-2

Figueroa, R. J. B., & Jung, I. (2025). Integrating VR tours in online language learning: A design-based research study. *Australasian Journal of Educational Technology*, *41*(1), 60–87. https://doi.org/10.14742/ajet.9498

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*, 59–109. https://doi.org/10.3102/00346543074001059

Fuhrman, O., Eckerling, A., Friedmann, N., Tarrasch, R., & Raz, G. (2021). The moving learner: Object manipulation in virtual reality improves vocabulary learning. *Journal of Computer Assisted Learning*, *37*(3), 672–683. https://doi.org/10.1111/jcal.12515

Gruber, A., & Kaplan-Rakowski, R. (2020). User experience of public speaking practice in virtual reality. In R. Z. Zheng (Ed.), *Cognitive and affective perspectives on immersive technology in education* (pp. 235–249). IGI Global. https://doi.org/10.4018/978-1-7998-3250-8.ch012

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Hwang, G.-J., Yin, C., & Chu, H.-C. (2019). The era of flipped learning: Promoting active learning and higher order thinking with innovative flipped learning strategies and supporting systems. *Interactive Learning Environments*, *27*(8), 991–994. https://doi.org/10.1080/10494820.2019.1667150

Jia, C., Hew, K. F., Jiahui, D., & Liuyufeng, L. (2023). Towards a fully online flipped classroom model to support student learning outcomes and engagement: A 2-year design-based study. *The Internet and Higher Education*, *56*, 100878. https://doi.org/10.1016/j.iheduc.2022.100878

Jiang, L., Meng, H., & Zhou, N. (2021). English learners' readiness for online flipped learning: Interrelationships with motivation and engagement, attitude, and support. *Language Teaching Research*, *28*(5), 2026–2051. https://doi.org/10.1177/13621688211027459

Jong, M. S. Y. (2023). Flipped classroom: Motivational affordances of spherical video-based immersive virtual reality in support of pre-lecture individual learning in pre-service teacher education. *Journal of Computing in Higher Education*, *35*(1), 144–165. https://doi.org/10.1007/s12528-022-09334-1

Kaplan-Rakowski, R., & Gruber, A. (2023). The impact of high-immersion virtual reality on foreign language anxiety. *Smart Learning Environments*, *10*(1), 46. https://doi.org/10.1186/s40561-023-00263-9

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lai, K. W. K., & Chen, H. J. H. (2021). A comparative study on the effects of a VR and PC visual novel game on vocabulary learning. *Computer Assisted Language Learning*, *36*(3), 312–345. https://doi.org/10.1080/09588221.2021.1928226

Lan, Y. J. (2020). Immersion, interaction and experience-oriented learning: Bringing virtual reality into FL learning. *Language Learning & Technology*, *24*(1), 1–15. http://hdl.handle.net/10125/44704

Lee, J. (2023). Flipped learning. In O. Zawacki-Richter & I. Jung (Eds.), *Handbook of open, distance and digital education* (pp. 1179–1196). Springer. https://doi.org/10.1007/978-981-19-2080-6_69

Lee, J., & Choi, H. (2018). Rethinking the flipped learning pre-class: Its influence on the success of flipped learning and related factors. *British Journal of Educational Technology, 50*(2), 934–945. https://doi.org/10.1111/bjet.12618

Leis, A. (2018). Content-based language teaching and the flipped classroom: A case study in the Japanese EFL environment. In J. Mehring & A. Leis (Eds.), *Innovations in flipping the language classroom* (pp. 221–230). Springer. http://dx.doi.org/10.1007/978-981-10-6968-0_15

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages*, *4*(1), 13. https://doi.org/10.3390/languages4010013

Lestiono, R., & Lee, S.-M. (2024). Virtual reality exposure therapy for reducing students' public speaking anxiety: A systematic review. *The JALT CALL Journal*, *20*(2), 1–24. https://doi.org/10.29140/jaltcall.v20n2.1515

Li, P., Legault, J., Klippel, A., & Zhao, J. (2020). Virtual reality for student learning: Understanding individual differences. *Human Behaviour and Brain*, *1*(1), 28-36. http://dx.doi.org/10.37716%2Fhbab.2020010105

Liaw, M.-L. (2019). EFL learners' intercultural communication in an open social virtual environment. *Educational Technology & Society*, *22*(2), 38–55. https://www.jstor.org/stable/26819616

Linling, Z., & Abdullah, R. (2023). The impact of COVID-19 pandemic on flipped classroom for EFL courses: A systematic literature review. *SAGE Open, 13*(1). https://doi.org/10.1177/21582440221148149

Lo, C. K. (2024). Strategies for enhancing online flipped learning: A systematic review of empirical studies during the COVID-19 pandemic. *Interactive Learning Environments*, *32*(7), 3517–3545. https://doi.org/10.1080/10494820.2023.2184392

Luan, L., Hwang, G.-J., Yi, Y., Lu, Z., & Jing, B. (2024). The effects of a self-developed virtual reality environment on college EFL learners' vocabulary learning. *Interactive Learning Environments*, *33*(1), 335–346. https://doi.org/10.1080/10494820.2024.2344056

MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J. M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11–30). Multilingual Matters. https://doi.org/10.21832/9781783097722-003

Nicolaidou, I., Pissas, P., & Boglou, D. (2021). Comparing immersive virtual reality to mobile applications in foreign language learning in higher education: A quasi-experiment. *Interactive Learning Environments*, *31*(4), 2001–2015. https://doi.org/10.1080/10494820.2020.1870504

Norouzian, R., & Plonsky, L. (2018). Eta-and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*, *34*(2), 257–271. https://doi.org/10.1177/0267658316684904

Pallant, J. (2016). *SPSS survival manual* (6th ed.). Open University Press.

Pang, Y. (2022). The role of web-based flipped learning in EFL learners' critical thinking and learner engagement. *Frontiers in Psychology*, *13*, 1008257. https://doi.org/10.3389/fpsyg.2022.1008257

Parmaxi, A. (2020). Virtual reality in language learning: A systematic review and implications for research and practice. *Interactive Learning Environments*, *31*(1), 172–184. https://doi.org/10.1080/10494820.2020.1765392

Patton, M. Q. (2015). *Qualitative evaluation and research methods* (4th ed.). Sage.

Peper, E., Wilson, V., Martin, M., Rosegard, E., & Harvey, R. (2021). Avoid Zoom fatigue, be present and learn. *NeuroRegulation*, *8*(1), 47-47. https://doi.org/10.15540/nr.8.1.47

Phanwiriyarat, K., Anggoro, K. J., & Chaowanakritsanakul, T. (2025). Exploring AI-powered gamified flipped classroom in an English-speaking course: A case of Duolingo. *Cogent Education*, *12*(1). https://doi.org/10.1080/2331186X.2025.2488545

Puentedura, R. (2006, August 18). Transformation, technology, and education. Retrieved from http://hippasus.com/resources/tte/

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778. https://doi.org/10.1016/j.compedu.2019.103778

Riazi, A. M., & Candlin, C. N. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues and challenges. *Language Teaching*, *47*(2), 135–173. https://doi.org/10.1017/s0261444813000505

Robert, J., Muscanell, N., McCormack, M., Pelletier, K., Arnold, K., Arbino, N., Young, K. & Reeves, J. (2025). *2025 EDUCAUSE Horizon Report®: Teaching and Learning Edition*. EDUCAUSE. https://www.educause.edu/horizon-report-teaching-and-learning-2025

Ronaghi, M. H. (2024). A new approach to business ethics education: Virtual reality-based flipped learning. *The International Journal of Management Education*, *22*(3), 101076. https://doi.org/10.1016/j.ijme.2024.101076

Shadiev, R., Dang, C., Sintawati, W., Yi, S., & Huang, Y.-M. (2023). Facilitating information literacy and intercultural competence development through the VR tour production learning activity. *Educational Technology Research and Development*, *71*, 2507–2537. https://doi.org/10.1007/s11423-023-10283-4

Spring, R., Kato, F., & Mori, C. (2019). Factors associated with improvement in oral fluency when using

video-synchronous mediated communication with native speakers. *Foreign Language Annals*, 52(1), 87–100. https://doi.org/10.1111/flan.12381

Stracke, C. M., Bothe, P., Adler, S., Heller, E. S., Deuchler, J., Pomino, J., & Wölfel, M. (2025). Immersive virtual reality in higher education: A systematic review of the scientific literature. *Virtual Reality*, *29*, 64. https://doi.org/10.1007/s10055-025-01136-x

Sun, J. C.-Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology*, *43*(2), 191–204. https://doi.org/10.1111/j.1467-8535.2010.01157.x

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

Tang, T., Abuhmaid, A. M., Olaimat, M., Oudat, D. M., Aldhaeebi, M., & Bamanger, E. (2020). Efficiency of flipped classroom with online-based teaching under COVID-19. *Interactive Learning Environments, 31*(2), 1077–1088. https://doi.org/10.1080/10494820.2020.1817761

Tao, J., & Gao, X. A. (2022). Teaching and learning languages online: Challenges and responses. *System*, *107*, 102819. https://doi.org/10.1016/j.system.2022.102819

Tecedor, M. & Kerber, N. (2025). Videoconferencing in second language learning. In L. McCallum & D. Tafazoli (Eds.) *The Palgrave encyclopedia of computer-assisted language learning*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-51447-0_26-1

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition, 41*(2), 363–387. https://doi.org/10.1017/S0272263118000311

Toyama, M., & Yamazaki, Y. (2021). Classroom interventions and foreign language anxiety: A systematic review with narrative approach. *Frontiers in Psychology*, *12*, 614184. https://doi.org/10.3389/fpsyg.2021.614184

Turan, Z., & Akdag-Cimen, B. (2019) Flipped classroom in English language teaching: A systematic review. *Computer Assisted Language Learning, 33*(5–6), 590–606. https://doi.org/10.1080/09588221.2019.1584117

Tutal, Ö., Yazar, T. (2021). Flipped classroom improves academic achievement, learning retention and attitude towards course: A meta-analysis. *Asia Pacific Education Review*, *22*, 655–673. https://doi.org/10.1007/s12564-021-09706-9

Vitta, J. P., & Al-Hoorie, A. H. (2023). The flipped classroom in second language learning: A meta-analysis. *Language Teaching Research*, *27*(5), 1268–1292. https://doi.org/10.1177/1362168820981403

White, J. R. (2024). Language learning through Zoom. *International Journal of Online Pedagogy and Course Design, 14* (1). https://doi.org/10.4018/IJOPCD.356273

Wu, J. G., Miller, L., & Wang, Q. (2021). Learning with immersive virtual reality: An exploratory study of Chinese college nursing students. *RELC Journal*, *54*(3), 697–713. https://doi.org/10.1177/00336882211044860

Wu, W. C. V., Yang, J. C., Chen Hsieh, J., S., & Yamamoto, T. (2020). Free from demotivation in EFL writing: The use of online flipped writing instruction. *Computer Assisted Language Learning*, *33*(4), 353–387. https://doi.org/10.1080/09588221.2019.1567556

Wu, Y.-H. S., & Hung, S.-T., A. (2022). The effects of virtual reality infused instruction on elementary school students' English-speaking performance, willingness to communicate, and learning autonomy. *Journal of Educational Computing Research*, *60*(6), 1558–1587. https://doi.org/10.1177/07356331211068207

Xin, X., & Zhang, M. (2024). Effects of flipped English learning designs on learning outcomes and cognitive load: Workload of out-of-class activities versus during-class activities. *Journal of Computer Assisted Learning*, *40*(4), 1745–1765. https://doi.org/10.1111/jcal.12978

Yan, W., Lowell, V. L., & Yang, L. (2024). Developing English language learners' speaking skills through applying a situated learning approach in VR-enhanced learning experiences. *Virtual Reality*, *28*, 167. https://doi.org/10.1007/s10055-024-01061-5

Yang, F.-C. O., Lo, F.-Y. R., Hsieh, J. C., & Wu, W.-C. V. (2020). Facilitating communicative ability of EFL learners via high-immersion virtual reality. *Journal of Educational Technology & Society*, *23*(1), 30–49. https://www.jstor.org/stable/26915405

Yashima, T., Noels, K. A., Shizuka, T., Takeuchi, O., Yamane, S., & Yoshizawa, K. (2009). The interplay of classroom anxiety, intrinsic motivation, and gender in the Japanese EFL context. *Journal of Foreign Language Education and Research*, *17*, 41–64. http://hdl.handle.net/10112/768

Zainuddin, Z., Farida, R., Keumala, C. M., Kurniawan, R., & Iskandar, H. (2021). Synchronous online flip learning with formative gamification quiz: Instruction during COVID-19. *Interactive Technology and Smart Education, 19*(2), 236–259. https://doi.org/10.1108/ITSE-01-2021-0002

Zainuddin, Z., & Halili, S. H. (2016). Flipped classroom research and trends from different fields of study. *International Review of Research in Open and Distributed Learning, 17*(3). https://doi.org/10.19173/irrodl.v17i3.2274

Zhao, J., Wallgrün, J. O., Sajjadi, P., LaFemina, P., Lim, K. Y. T., Springer, J. P., & Klippel, A. (2022). Longitudinal effects in the effectiveness of educational virtual field trips. *Journal of Educational Computing Research*, *60*(4), 1008–1034. https://doi.org/10.1177/07356331211062925

Zou, D., Luo, S., Xie, H., & Hwang, G.-J. (2022): A systematic review of research on flipped language classrooms: Theoretical foundations, learning activities, tools, research topics and findings. *Computer Assisted Language Learning*, *35*(8), 1811–1837. https://doi.org/10.1080/09588221.2020.1839502

Publisher's Note