International Journal of Learning, Teaching and Educational Research
Vol. 25, No. 8, pp. 1-35, August 2026
https://doi.org/10.26803/ijlter.25.8.1
Received May 26, 2026; Revised Jul 10, 2026; Accepted Jul 17, 2026

VR-Enhanced EFL Instruction in Higher Education: A Systematic Review（2020-2024）

Ruimin Zou
Faculty of Education, Universiti Kebangsaan Malaysia,
Bangi, Selangor, Malaysia
School of International Studies, Guangzhou Xinhua University,
Guangzhou, Guangdong, PR China

Nur Ehsan Mohd Said* **and Nor Hafizah Adnan**
Faculty of Education, Universiti Kebangsaan Malaysia,
Bangi, Selangor, Malaysia

Abstract. This systematic review examines how Virtual Reality (VR) has been used to support English as a Foreign Language (EFL) instruction in higher education. It synthesises evidence on reported cognitive, affective, and behavioural outcomes, and analyses how learner characteristics, instructional design, and contextual conditions are associated with VR effectiveness. Following PRISMA 2020 guidelines, the review included 49 empirical studies published between 2020 and 2024. The findings show that VR was most consistently associated with effective and behavioural outcomes, including engagement, motivation, willingness to communicate, and active participation. Cognitive outcomes were less stable. Improvements were more often reported for meaning-focused performance than for linguistic accuracy or higher-order cognitive processing. More favourable results appeared in studies that used active, constructive tasks, instructor guidance, combined forms of scaffolding, and longer interventions. Outcomes also varied according to VR type and instructional setting. The review indicates that VR is more likely to support EFL learning when activities are matched to clear pedagogical goals, learners' readiness, and feasible classroom conditions. By taking a design-oriented perspective, this review shifts attention from whether VR works to the conditions under which it appears to support EFL learning in higher education contexts.

Keywords: Virtual Reality (VR); English as a Foreign Language (EFL); language learning; language education; higher education

Citation:
Zou, R., Said, N, E, M., & Adnan, N, H. (2026). VR-Enhanced EFL Instruction in Higher Education: A Systematic Review（2020-2024）. *International Journal of Learning, Teaching and Educational Research*, *25*(8), 1–35.
https://doi.org/10.26803/ijlter.25.8.1

*Corresponding author: Nur Ehsan Mohd Said; *nurehsan@ukm.edu.my*

1. Introduction

In recent years, there has been growing interest in integrating VR into language education. VR has made it possible for learners to engage in realistic, computer-generated contexts that support meaningful language use beyond traditional classrooms (Ou Yang et al., 2020; Parmaxi, 2023; Song et al., 2023; Utami et al., 2021; Wu et al., 2023; Yan et al., 2024; Yang et al., 2022). Immersive systems such as head-mounted displays (HMDs) can create a strong sense of presence, commonly understood as the feeling of being there, and this may help learners take part more actively in communicative tasks (Baxter & Hainey, 2024; Dhimolea et al., 2022).

At the same time, VR does not improve learning simply because it is immersive. Its educational value depends on how its affordances are organised through tasks, guidance, and classroom use (Dhimolea et al., 2022; Di Natale et al., 2020). The Cognitive Affective Model of Immersive Learning (CAMIL; Makransky & Petersen, 2021) offers a useful framework for this argument. In CAMIL, presence and agency affect learners' cognitive and affective processes, including interest, motivation, self-efficacy, and cognitive load. These processes influence learning outcomes, but they are also shaped by instructional, learner, and contextual conditions (Makransky & Petersen, 2021). This review therefore treats those conditions as possible factors in the effectiveness of VR-enhanced language learning.

Previous systematic reviews and meta-analyses have examined broad trends, technological affordances, and domain-specific uses of VR in language learning (Dhimolea et al., 2022; Di Natale et al., 2020; Ece et al., 2023; Parmaxi, 2023; Qiu et al., 2024; Zhang et al., 2025). They show that VR can provide immersive contexts, support engagement, and improve learners' experience. However, much of this work remains broad in scope, mainly descriptive, or focused on specific areas of language learning. It therefore says less about why outcomes vary from one study to another. Meta-analytic research has estimated overall effects and examined moderators such as educational level, region, VR type, and outcome domain (Qiu et al., 2024). In higher education EFL contexts, however, there is still a need to examine how instructional design, teacher support, scaffolding, task type, intervention duration, and learning environment relate to VR effectiveness.

This systematic review synthesises empirical studies published between 2020 and 2024 on VR-enhanced EFL instruction in higher education. It examines reported cognitive, affective, and behavioural outcomes, and analyses how learner characteristics, instructional design, and contextual conditions are associated with those outcomes. The aim is to move beyond the general question of whether VR works and to consider the conditions under which it appears to support EFL learning. The study addresses the following research questions:

1. What cognitive, affective, and behavioural outcomes have been reported in empirical studies on VR-enhanced EFL instruction in higher education?
2. What learner, instructional, and contextual factors are associated with differences in the outcomes of VR-enhanced EFL instruction in higher education?

2. Methodology

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021). It focused on empirical research that examined VR use in EFL instruction in higher education. Inclusion and exclusion criteria were defined before screening. Studies were included if they used quantitative, qualitative, or mixed-methods designs, involved tertiary-level participants, focused on EFL learning, and used VR as a central part of the instructional intervention. The review covered studies published between 2020 and 2024, a period chosen to reflect recent developments in VR-enhanced EFL instruction.

The review excluded non-empirical publications, studies outside higher education, research on ESL contexts or languages other than English, and studies in which VR was not central to the instructional design. For quantitative intervention studies, a control group or comparison condition was required, since studies without one provide limited evidence for changes in learning outcomes. Survey studies were retained when they addressed outcomes related to VR-enhanced EFL learning. Studies limited to usability or user perception were excluded if they did not connect those issues to EFL learning outcomes. Quantitative intervention studies with very small samples were also excluded when their evidence was not sufficient for identifying meaningful or comparable outcome patterns.

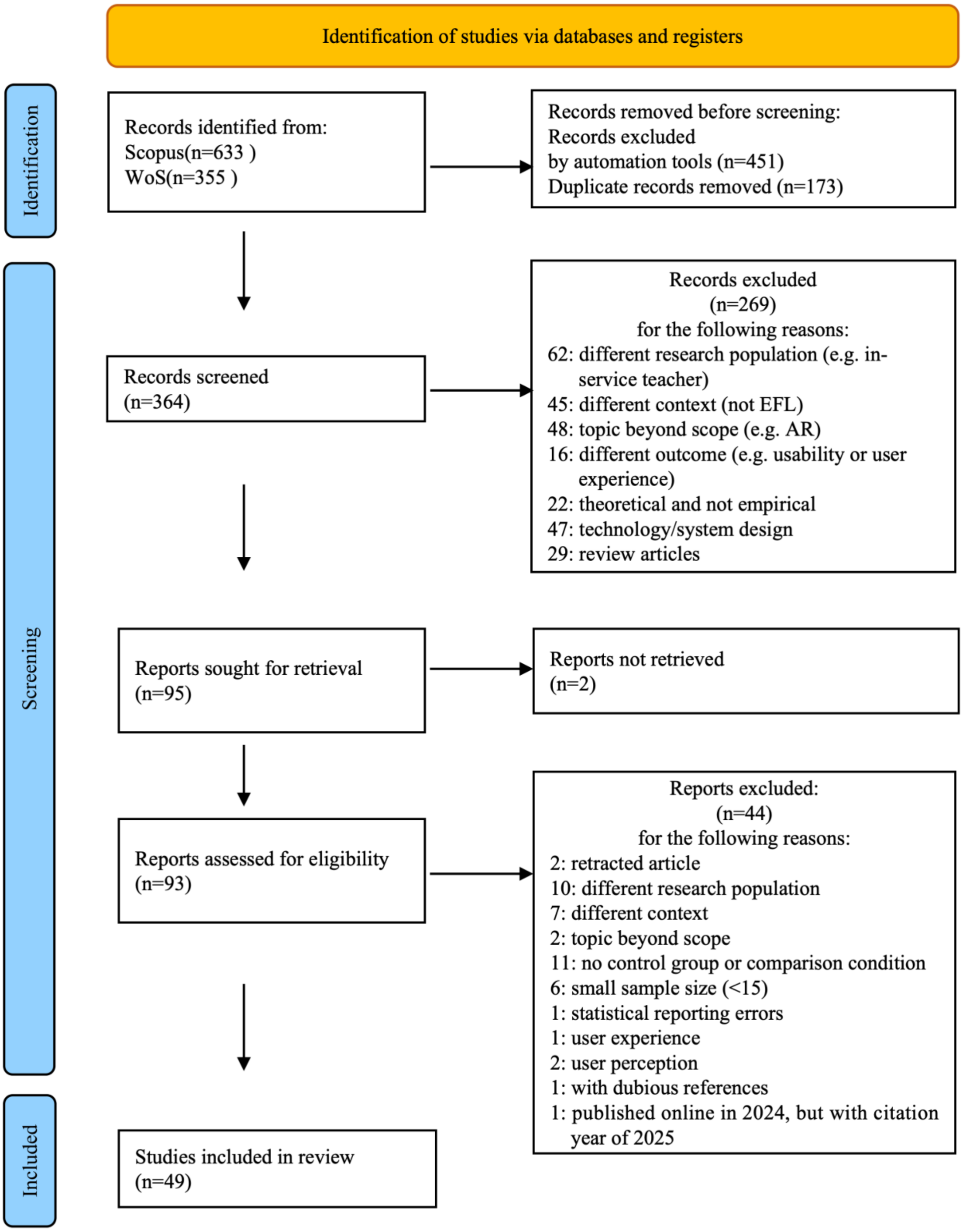

Figure 1: The PRISMA Flow Diagram (adapted from Page et al. 2021)

The search was conducted in March 2025 on the Web of Science and Scopus. These databases were used because they index peer-reviewed research across applied linguistics, educational technology, and language education. Search terms combined VR technologies, language learning, and higher education. The full search strings are reported in the Appendix Table A1. Records were filtered by publication year, document type, and language. Title and abstract screening were then conducted, followed by full-text assessment. The screening process resulted in 49 empirical studies for final inclusion.

Two researchers independently screened and coded all included studies. Inter-coder agreement was calculated using Cohen's Kappa and reached a high level ($\kappa = 0.83$; Landis & Koch, 1977). Disagreements were resolved through discussion, and a third researcher was available when consensus could not be reached. Data were extracted and analysed according to the research questions. The coding focused on cognitive, affective, and behavioural outcomes, together with learner, instructional, and contextual factors that may help explain differences in the outcomes of VR-enhanced EFL learning.

3. Results

3.1 Study Selection Results

The study selection process followed PRISMA 2020 statement. The database search identified 988 records, including 633 from Scopus and 355 from Web of Science. Before manual screening, 451 records were removed through filters for publication year (2020 to 2024), document type (journal articles and conference proceedings), and language (English). After filtering, 537 records remained. A further 173 duplicates were removed, leaving 364 records for the title and abstract screening. During this stage, 269 articles were excluded because they were non-empirical (e.g., reviews, theoretical papers), or because they did not match the population, context, or topic of the review. The remaining 95 articles were assessed in full-text assessment. At this stage, 44 articles were excluded, mainly because they did not meet the methodological criteria. These articles commonly lacked a control group or comparison condition, had an insufficient sample size (<15), or focused on user perception rather than learning outcomes. The final synthesis included 49 studies (see Figure 1).

3.2 General Characteristics of Included Studies

3.2.1 Publication Year Distribution

The 49 included studies were grouped by publication year to show how research output changed from 2020 to 2024. As shown in Figure 2, the number of empirical studies increased from 2020 to 2023. The highest number appeared in 2024, which accounted for about 41% of all included studies. This pattern indicates that empirical research on VR-supported EFL instruction in higher education became more active in the later years of the review period.

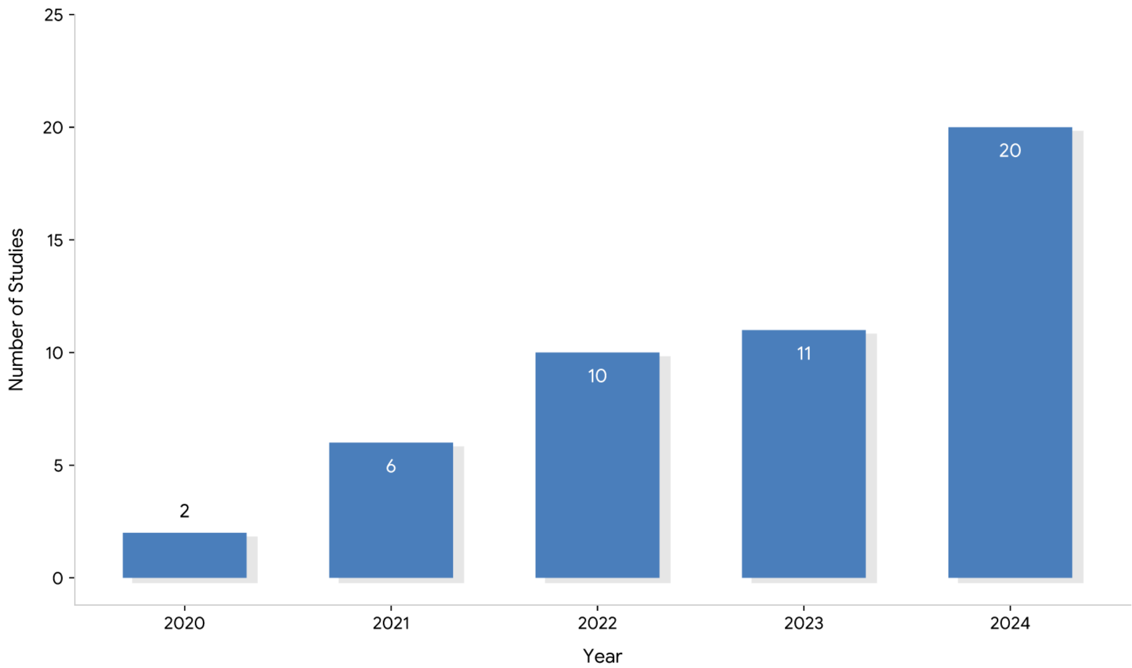

Figure 2: Distribution of Included Studies by Publication Year

3.2.2 Geographical Distribution

The included studies were concentrated in a small number of regions. As shown in Figure 3, 82% of the studies were conducted in Asia, with China contributing a large share. This concentration limits the generalizability of the findings to other educational contexts, although it also reflects the broader distribution of EFL learning contexts globally (Kachru, 1990; Bolton & Bacon-Shone, 2020). In China, teacher-centred instruction and high-stakes examination cultures remain common in EFL education (Cheng et al., 2023; Xin et al., 2025). These conditions may influence how VR is used and which outcomes researchers choose to measure. Government investment in educational technology infrastructure has also supported VR integration in China (Yan & Yang, 2021). Such conditions may be harder to reproduce in lower-resource EFL contexts.

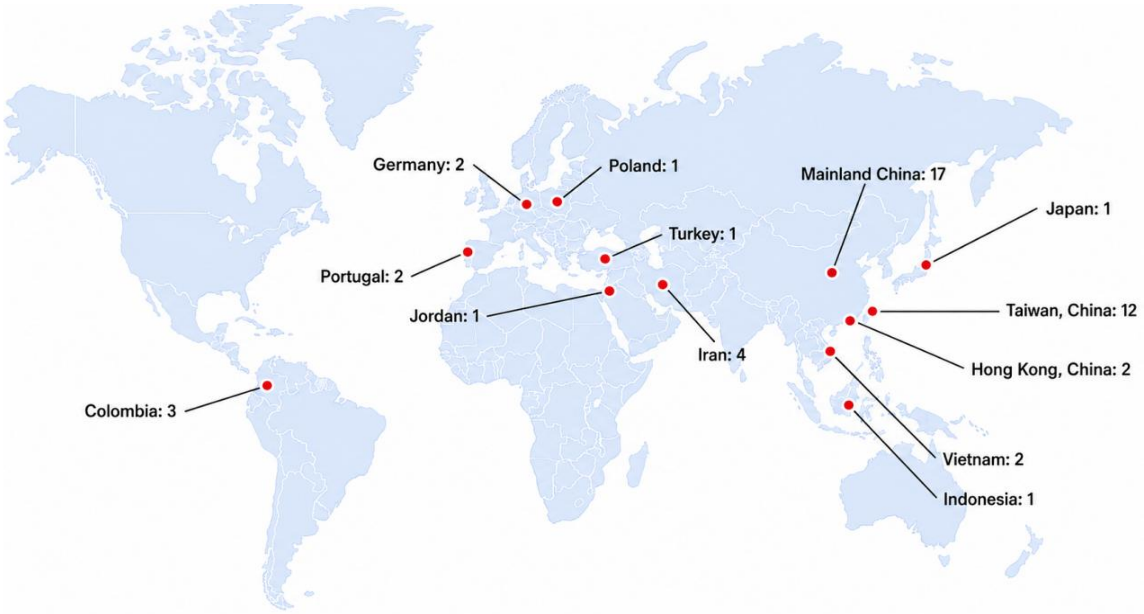

Figure 3: Geographical Distribution of Included Studies

3.2.3 Research Designs and Methodological Trends

Most included studies used quasi-experimental designs (N = 35), while 12 used true experimental designs. Two studies adopted non-experimental survey designs. The large number of quasi-experimental studies is consistent with research conducted in classroom settings, where random assignment and full experimental control are often difficult. Mixed-methods studies were also present in the dataset. Their use suggests that some researchers are moving beyond outcome measurement alone and are also examining learners' experiences and the processes through which VR may support EFL learning (see Figure 4).

3.2.4 Participant and Sample Size

Participants in all 49 studies were university students. Sample sizes ranged from 16 to 235 participants. Nearly half of the studies involved 21 to 50 participants, so small and medium samples were common. Many studies recruited English majors or students in related fields, such as interpreting and translation. Other studies included non-English majors from medicine, engineering, business, law, computer science, social sciences, and humanities. Several studies used mixed-discipline groups.

English proficiency was usually reported through Common European Framework of Reference for Languages (CEFR) levels, with participants ranging from A1 to B2 or above. Some studies used standardized tests, including TOEIC and CET-4, while others used placement tests or self-reports. A number of studies used general proficiency descriptors or did not report proficiency clearly. Prior VR experience was reported less consistently. 23 out of 49 studies either did not specify participants' prior VR exposure or gave only limited information. Among the studies that did report this information, most participants were new to VR. This missing baseline information matters because prior VR exposure may affect the novelty effect, where outcomes are influenced by learners' first experience with a new technology (Clark, 1983).

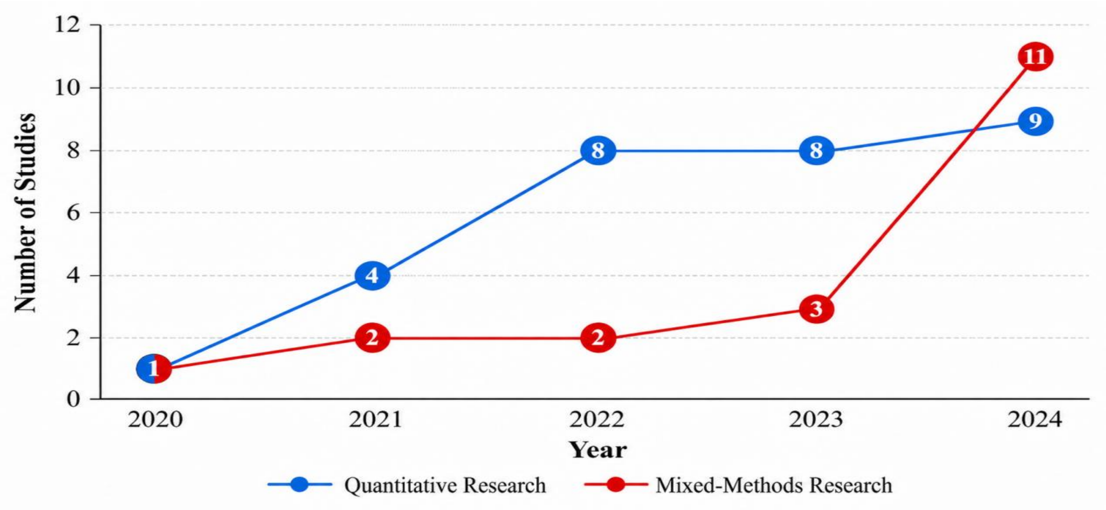

Figure 4: Trends in Research Methodologies by Year

3.2.5 VR Technology

As summarized in Table 1, the included studies employed several types of VR technology. HMD-based VR and mobile VR were the most common, each appearing in 19 studies. HMD-based VR delivers more enveloping virtual environments, richer visual realism, and more natural interaction through controllers, which can support spatial presence and experiential learning (Radianti et al., 2020). Mobile VR uses a smartphone with a lightweight viewer and is usually less costly and easier to access (Radianti et al., 2020; Tai, 2022). In terms of software, studies used custom-built environments, often developed with Unity, as well as existing educational applications such as EduVenture and VirtualSpeech. The reporting of technological specifications was uneven. Several studies described the VR system only in general terms or did not identity the hardware models. This limits replication and makes it harder to compare findings across studies that used different levels of immersion.

Table 1: Frequency of VR Hardware Types

VR Hardware Type	Representative Devices	N
HMD-based VR	HTC Vive, Oculus Quest, Pico Neo, Meta Quest 2	19
Mobile VR	Google Cardboard, Samsung Gear VR	19
Desktop VR	PC, laptop	4
Cross-Platform/Blended	Multiple devices	2
Not Specified	/	5
Total		49

3.2.6 Target Learning Domains

The included studies mainly targeted productive skills, especially speaking (see Appendix Table A2). This emphasis fits the way VR is often used: to create simulated, interactive situations for communicative practice. Receptive skills, including reading and listening, received less attention. Studies combining several language skills were also limited. Six studies focused on English for Specific or Academic Purposes (ESP/EAP), and five studies examined affective outcomes only, with motivation or anxiety reduction treated as the main outcome.

3.3 Learning Outcomes of VR-Enhanced EFL Instruction

This section reports the cognitive, affective, and behavioural outcomes identified in the reviewed studies in response to RQ1. Appendix Table A3 provides study-level details, including authorship, outcome types, and main findings.

3.3.1 Cognitive Outcomes

Of the 49 included studies, 41 examined cognitive outcomes. 25 of these studies (61%) reported mainly positive cognitive outcomes. The reported outcomes include lower-level and concrete learning, meaning-focused productive performance, linguistic accuracy, and higher-order cognitive processing. The most consistent gains were found in productive language skills, particularly speaking and writing. Across the studies, VR-supported instruction was associated with improvements in fluency, lexical resource, and content generation. The evidence was weaker for linguistic accuracy. Several studies reported limited or no improvement in grammar, pronunciation, or spelling (e.g., Ebadi & Ebadijalal, 2022; Hoang et al., 2023; Wang et al., 2022; Feng & Ng, 2023).

The pattern differed by type of cognitive outcome. VR was more often linked to lower-level and concrete outcomes, including vocabulary retention, information location, and basic comprehension (e.g., Kaplan-Rakowski & Gruber, 2024; Wang et al., 2021; Feng & Ng, 2024). Results were less consistent for higher-order processes, such as evaluation, reflection, and complex problem-solving. Some studies reported gains in creative or computational thinking (e.g., Mubarok et al., 2024), and others found no significant improvement in deeper analytical skills (e.g., Wang et al., 2021).

Studies on cognitive load and processing showed a similar degree of variation. Some reported enhanced cognitive absorption (e.g., Hsu, 2024) and more efficient processing when scaffolds were provided (e.g., Bacca-Acosta et al., 2022). Others found that immersive and multimodal environments could increase cognitive demands, divert attention, or reduce learners' focus on linguistic form (e.g., Gao et al., 2021; Peixoto et al., 2023). Individual differences also appeared to shape these outcomes. Higher-performing learners tended to process VR tasks more efficiently, while lower-performing learners were more likely to experience overload (e.g., Li et al., 2024; Chen & Hwang, 2022). Overall, the evidence suggests that VR supports meaning-focused performance and concrete learning outcomes more reliably than form-focused accuracy or higher-order cognitive processing.

3.3.2 Affective Outcomes

Affective outcomes were examined in 38 studies, and 63% reported significant positive effects. The most common finding was a reduction in language anxiety. Several studies found that VR gave learners a lower-pressure setting for practice, where they felt less exposed to peer or teacher judgment (e.g., Ebadi & Ebadijalal, 2022; Kaplan-Rakowski & Gruber, 2023; Wang et al., 2021). Practice through avatars or simulated environments was associated with lower anxiety than traditional face-to-face or video-mediated interaction (e.g., Chen, 2024; Kaplan-Rakowski & Gruber, 2023). In many studies, lower anxiety was reported alongside greater confidence, self-efficacy, and willingness to communicate.

VR interventions were also associated with positive motivational states, learner involvement, and emotional responses. These outcomes were often connected to the immersive and experiential qualities of VR. Studies reported that presence and perceived realism were related to stronger learner involvement and intrinsic motivation (e.g., Sun et al., 2020; Yuan et al., 2023; Wang et al., 2022). Learners also reported a heightened sense of involvement and interest when they had active roles in VR tasks. Activities involving creation or interaction within virtual environments fostered a sense of ownership and achievement, which contributed to positive emotional responses such as enjoyment and satisfaction (e.g., Hoang et al., 2023; Yuan et al., 2023).

However, these affective gains are not always sustained or transferable. Several studies reported a gap between learners' positive perceptions and measured outcomes. Higher engagement and enjoyment did not always correspond to improved performance or stronger self-efficacy (e.g., Lin & Wang, 2021; Hung et al., 2024). Similarly, reduced anxiety in VR did not always extend to real-world communication, where learners lacked the structured support and cues available in virtual settings (e.g., Ding, 2024). Some learners also experienced technology-related anxiety or negative interactions in open VR environments, including difficulty using unfamiliar tools or dealing with unregulated social behaviour (e.g., Huang, 2024; Le et al., 2024). These findings suggest that VR tends to enhance affective engagement, but its benefits depend on context and may not extend beyond the virtual environment.

3.3.3 Behavioural Outcomes

Behavioural outcomes were assessed in 22 studies through objective measures, such as eye-tracking and system logs, as well as self-reported data. Most of these studies reported higher levels of active participation and social interaction in VR environments (e.g., Chen et al., 2021; Yeh et al., 2022; Li et al., 2022). Learners often took more active roles in tasks involving content creation, exploration, and self-directed learning. These behaviours were reported together with sustained behavioural engagement and concentration (e.g., Yuan et al., 2023; Feng & Ng, 2024).

The studies also described more collaborative forms of task engagement. Learners engaged in negotiated interaction, including checking for peer understanding and providing deeper feedback during task completion (e.g., Chen et al., 2023; Shadiev et al., 2024). Some studies reported adaptive learning strategies, such as imitation of linguistic input and attention to contextual cues to complete communicative tasks (e.g., Hoang et al., 2023; Lin et al., 2023). Eye-tracking data added further evidence, showing that learners attended to embedded scaffolds and task-relevant information during VR activities (e.g., Li et al., 2024; Bacca-Acosta et al., 2022).

Behavioural changes were also seen in learner autonomy and participation beyond classroom requirements. Several studies reported that learners invested additional time in preparation and self-study, indicating more self-regulated learning behaviours (e.g., Hoang et al., 2023; Hsu et al., 2022). Studies of oral performance also reported improvements in non-verbal communication, including eye contact, vocal projection, and use of gestures use (e.g., Chen, 2024; Chorão et al., 2023; Kaplan-Rakowski & Gruber, 2023). These behavioural changes were not always matched by language gains. In some studies, behavioural engagement was accompanied by cognitive or linguistic improvements, including fewer errors, more efficient task completion across sessions, better vocabulary retention through repeated interaction, and greater use of target phrases or sentence patterns in output (e.g., Bacca-Acosta et al., 2022, 2023; Feng & Ng, 2024; Lin et al., 2023). Other studies reported null or mixed cognitive outcomes.

In some cases, immersion appeared to distract learners from peer language use, which suggests that behavioural engagement may reflect situational involvement rather than deeper language development (e.g., Ebadi & Ebadijalal, 2022; Gao et al., 2021). Across the reviewed studies, VR-enhanced EFL instruction was linked to more active, collaborative, and self-directed learning behaviours. These behaviours are important indicators of participation, but they should not be treated as evidence of language development unless they are supported by cognitive or linguistic outcome measures.

3.4 Factors Influencing the Effectiveness of VR-Enhanced EFL Instruction

To examine why outcomes varied across studies, this section addresses RQ2 by examining learner, instructional, and contextual factors. Outcome patterns were coded as positive, mixed, or null. Positive outcomes referred to consistent and statistically significant improvements. Mixed outcomes referred to cases where significant gains were found in some areas but not others, or where effects were inconsistent. Null outcomes referred to cases where no statistically significant improvement was reported.

3.4.1 Learner Factors

Two learner-related factors were associated with variation in the reviewed studies: English proficiency and academic major. These factors should not be read as direct causes of VR effectiveness, but they help describe where outcome patterns differed across learner groups.

a. English Proficiency

English proficiency was associated with differences in reported outcomes. Across proficiency levels, positive outcomes were more common than mixed or null outcomes, suggesting that VR can support learners at different levels of English. However, the distribution of mixed and null findings differed across groups (see Figure 5).

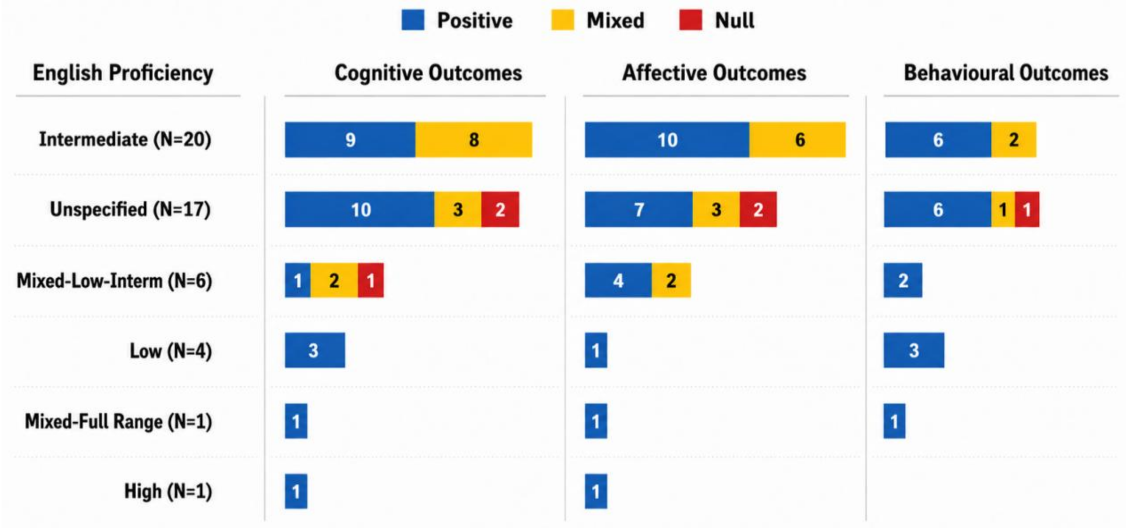

Figure 5: Analysis of Learners' English Proficiency and Cognitive, Affective and Behavioural Outcomes

Note: N = number of studies. Mixed-Low-to-Intermediate refers to studies in which learners' English proficiency ranged from low to intermediate. Mixed-Full Range refers to studies in which learners' English proficiency ranged from low, to intermediate and high.

Studies involving lower- and higher-proficiency learners were fewer in number, but they generally reported positive results. Lower-level learners benefited from support for basic vocabulary and grammar, and higher-level learners showed gains in domain-specific academic tasks. Intermediate learners were the most frequently studied group. Their outcomes were largely positive in the affective and behavioural domains, but cognitive findings were more variable, especially for complex skills such as reading comprehension and writing (e.g., Akay & Kessler, 2024; Ebadi & Ebadijalal, 2022; Khodabandeh, 2022). Studies that did not specify proficiency showed a similar pattern, with positive outcomes alongside mixed and null results. This variability may be associated with cognitive overload or difficulties adapting to VR environments (Gao et al., 2021; Huang, 2024).

b. Academic Major

Learners' academic major was also associated with differences in the effectiveness of VR-enhanced EFL instruction (see Figure 6). Learners from language-focused disciplines, including English and Translation/Interpreting, showed more consistent positive outcomes across cognitive, affective, and behavioural domains. Some variation was still found in behavioural performance (e.g., Chan, 2023; Feng & Ng, 2024).

Outcomes were more variable for learners in non-language disciplines. Some studies reported gains in discipline-specific communication (e.g., Chen & Hwang, 2022; Derakhshan et al., 2024), while others found limited improvements in vocabulary, cultural understanding, oral performance, or anxiety reduction (e.g., Chen, 2024; Gao et al., 2021; Hung et al., 2024). Studies involving mixed or unspecified majors showed a similar pattern. Overall, VR appeared to produce more consistent benefits for learners whose academic training was closely connected to language and communication. Outcomes were less consistent in other disciplinary contexts.

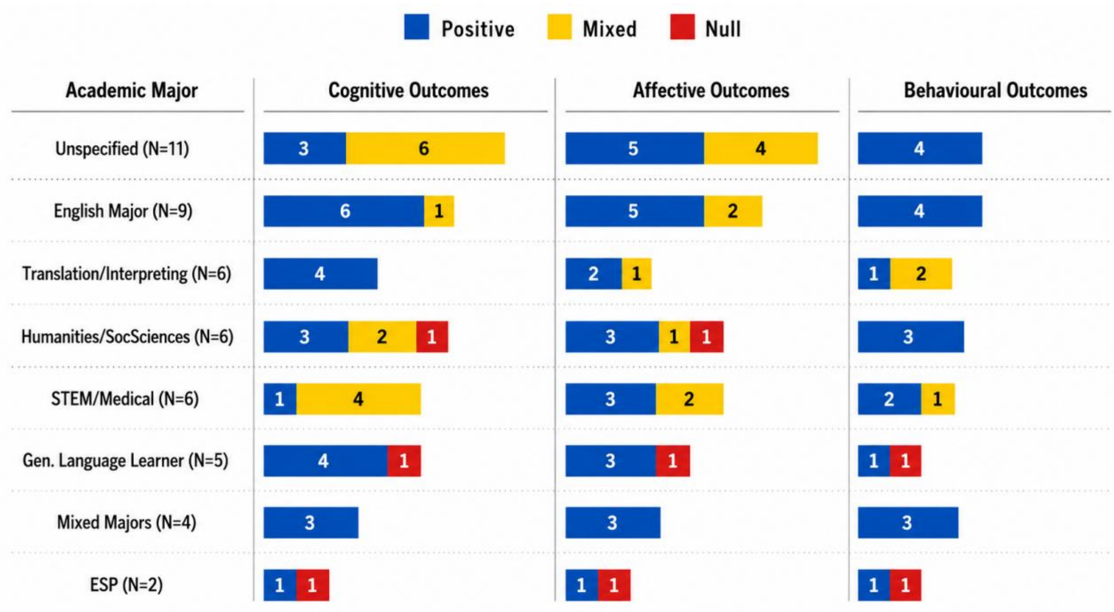

Figure 6: Analysis of Learners' Academic Major and Cognitive, Affective and Behavioural Outcomes

Note: N = number of studies. SocSciences refers to social sciences. STEM refers to Science, Technology, Engineering, and Mathematics. Gen. Language Learner refers to general language learner.

3.4.2 Instructional Factors

Instructional design choices were linked to differences in the reported outcomes of VR-enhanced EFL instruction. The reviewed studies were coded according to task type, teacher role, scaffolding type, and duration of VR exposure.

a. Learning Tasks in VR

Learning tasks were classified using the Interactive, Constructive, Active, and Passive framework (Chi & Wylie, 2014). Passive tasks involve receiving information, active tasks require physical manipulation of learning materials, constructive tasks involve generating new outputs, and interactive tasks involve co-construction of knowledge with peers (Chi & Wylie, 2014). These categories were used to describe the kinds of engagement built into VR-enhanced language learning.

Positive outcomes appeared across all task types (see Figure 7). Constructive and active tasks showed more stable patterns across cognitive, affective, and behavioural domains. Passive and interactive tasks exhibited greater variability. Constructive tasks, such as content creation and presentations, were linked to higher-order skills (e.g., Chen & Hwang, 2022; Mubarok et al., 2024). Active tasks, including vocabulary games and object manipulation appeared to support foundational learning (e.g., Bacca-Acosta et al., 2023; Budianto et al., 2023; Feng & Ng, 2024). Interactive tasks, such as role-plays and collaboration, were often connected with affective and behavioural outcomes, including communicative competence and reduced anxiety (Hsu et al., 2022; Yan et al., 2024). Passive tasks, such as watching or listening, were less consistent. Some studies reported positive findings, while others reported mixed cognitive outcomes.

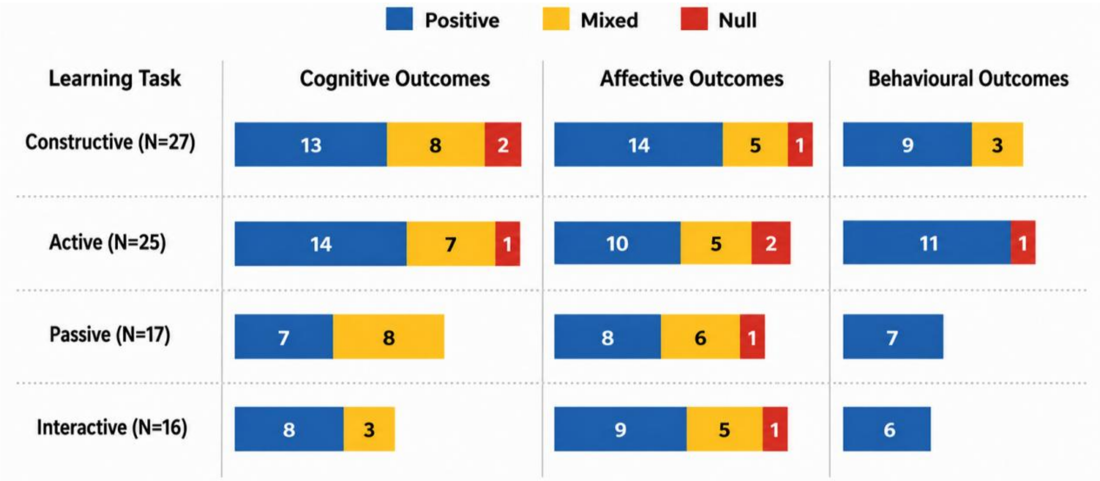

Figure 7: Analysis of Learning Task and Cognitive, Affective and Behavioural Outcomes

Note: N = number of studies. Categories are not mutually exclusive because some studies incorporated more than one type of learning task

b. Teacher Role in VR

Research in technology-enhanced language learning suggests that teacher involvement influences how learners engage with digital tools and tasks (Hoi & Mu, 2021; Wang, 2021). In this review, teacher involvement was coded into three roles: instructor, facilitator, and no teacher. The instructor role involved direct instruction, the facilitator role mainly involved technical support, and the no-teacher category referred to self-directed learning.

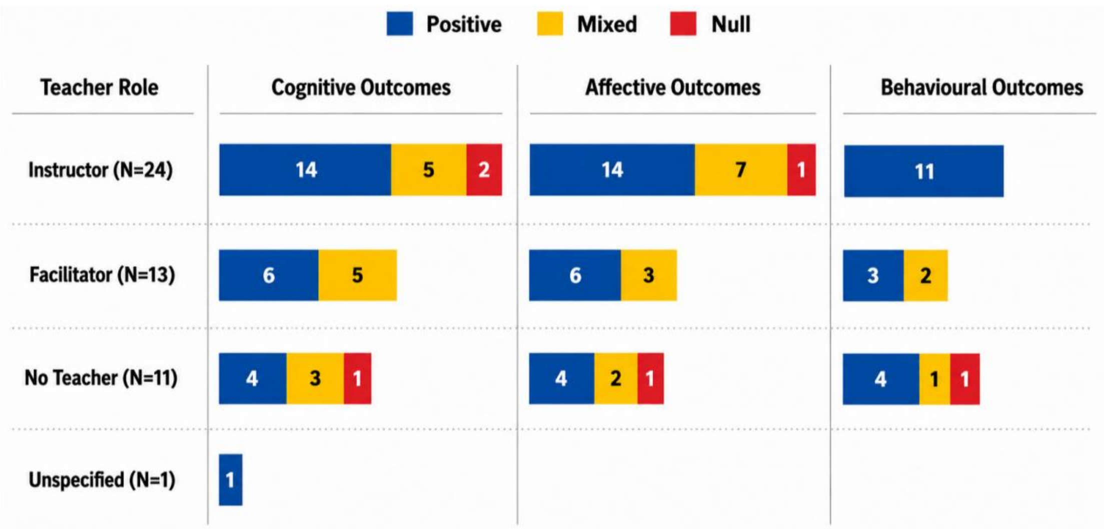

Figure 8: Analysis of Teacher Role and Cognitive, Affective and Behavioural Outcomes

Note: N = number of studies.

Instructor-led interventions showed the most consistent positive outcomes across all dimensions (see Figure 8). Behavioural outcomes were positive in all studies in this category. Cognitive and affective outcomes were also mostly positive (e.g., Ma, 2021; Porras et al., 2024; Wang et al., 2022; Yuan et al., 2023). Studies with teachers as facilitators showed a less consistent pattern. Affective results remained largely positive, but cognitive outcomes were more variable, with positive findings accounting for about half of the cases. This indicates that technical support may help learners stay engaged, but it does not always support cognitive gains.

No-teacher interventions presented greater variation. Cognitive outcomes included higher proportions of mixed and null results, indicating less consistent learning outcomes when teacher presence was absent (e.g., Lin et al., 2023; Liu & Hou, 2020). Behavioural outcomes remained relatively positive, suggesting that learners could stay engaged in VR tasks even without direct external support (e.g., Bacca-Acosta et al., 2023; Ebadijalal & Yousofi, 2024).

c. Scaffolding in VR

Scaffolding was coded using Hannafin et al.'s (2013) framework. Conceptual scaffolding guides learners toward key ideas, metacognitive scaffolding supports monitoring and reflection, strategic scaffolding offers alternative approaches, and procedural scaffolding helps learners navigate systems and tasks.

As presented in Figure 9, conceptual scaffolding was frequently reported and generally linked to positive affective outcomes. In several studies, it included concept lists or collaborative construction to support vocabulary and grammar learning (e.g., Li et al., 2022; Yan et al., 2024). Procedural supports, such as tutorials, orientation, embedded prompts, helped learners overcome technical barriers (e.g., Bacca-Acosta et al., 2023; Sun et al., 2020). Strategic supports, including peer feedback and modelling, were linked to higher-order thinking and positive behavioural outcomes (e.g., Chan, 2024; Mubarok et al., 2024). Metacognitive scaffolding was less common, but the studies that used it connected reflection, monitoring, and feedback with learner autonomy (e.g., Chen, 2024; Huang, 2024).

Among the studies, 11 used only one type of scaffolding, and 8 of them (72.7%) reported mixed or null outcomes. By comparison, 38 studies used two or more scaffolding types, and half of these studies reported positive outcomes. These findings suggest that combined scaffolding may offer more stable support than a single form of support.

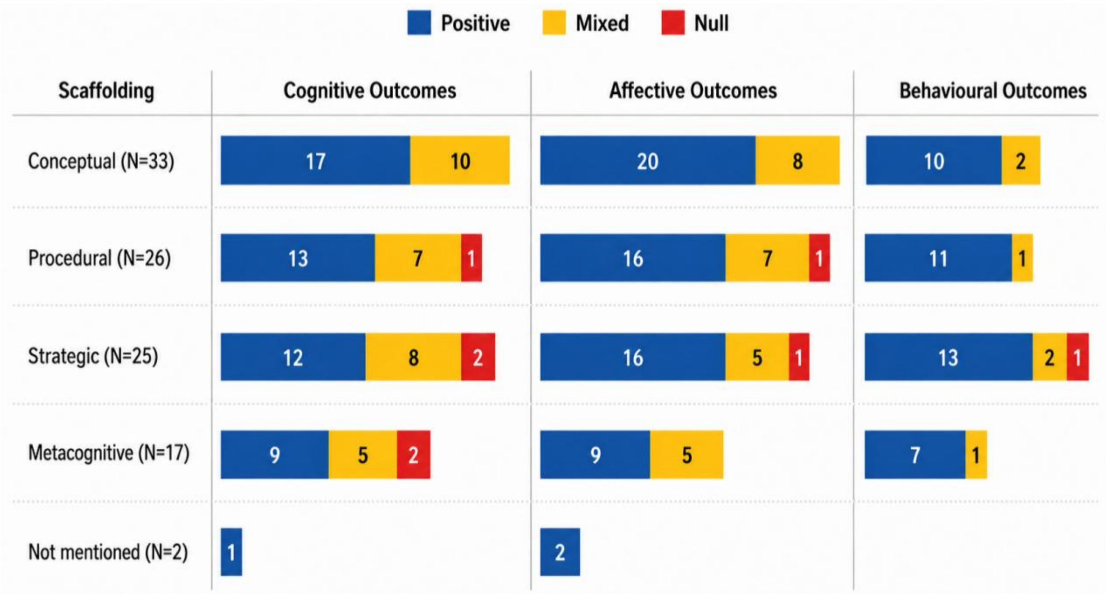

Figure 9: Analysis of Scaffolding Type and Cognitive, Affective and Behavioural Outcomes

Note: N = number of studies. Categories are not mutually exclusive because some studies included more than one type of scaffolding.

d. VR Instructional Duration

The duration of VR exposure was also related to outcomes patterns (see **Error! Reference source not found.**). Longer interventions, defined as 8 weeks or more, reported more positive results. Studies lasting 10 weeks reported improvements in speaking, writing, lower anxiety, and greater learner autonomy (e.g., Hoang et al., 2023; Hsu et al., 2022; Yan et al., 2024). Short-term interventions, lasting less than 8 weeks, also reported positive findings, especially in motivation and participation, but cognitive outcomes were less consistent.

Brief interventions of 3 to 4 weeks improved reading comprehension in some studies (e.g., Liu et al., 2024; Wang et al., 2022), but findings for higher-order skills such as problem-solving, evaluation, and reflection were not significant. One-time sessions produced the most variable results. In the affective domain, mixed and null findings were more common than positives findings. Although some studies reported cognitive or behavioural gains, many found only marginal or no improvements in proficiency, motivation, or engagement (e.g., Gao et al., 2021; Kaplan-Rakowski & Gruber, 2024; Li et al., 2022).

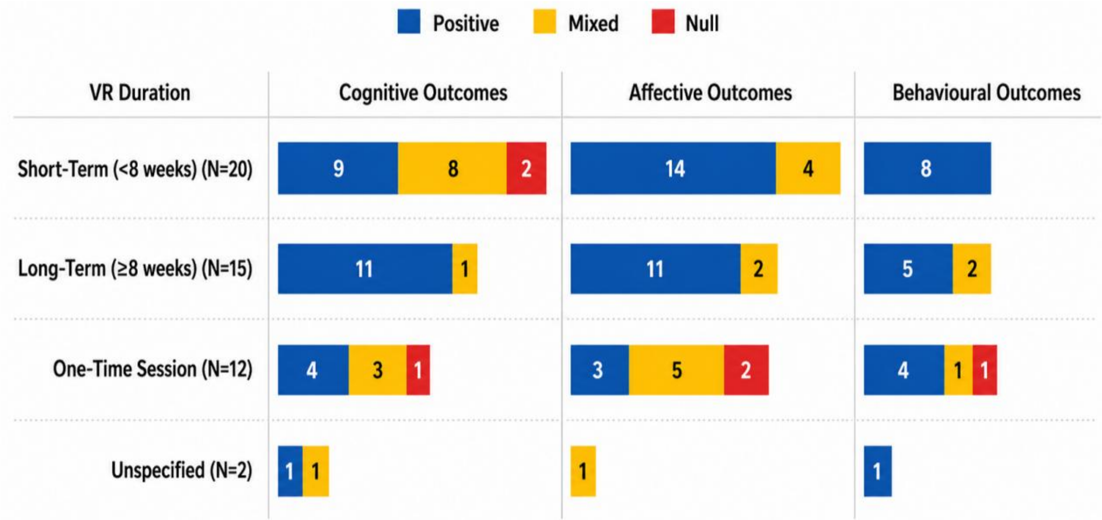

Figure 10: Analysis of VR Duration and Cognitive, Affective and Behavioural Outcomes

Note: N = number of studies.

3.4.3 Contextual Factors

Contextual factors were related to how VR-enhanced instruction was implemented and experienced. This review coded two such factors: VR type and instructional setting.

a. VR Type

The reported outcomes varied by VR type (see Figure 11). Mobile VR studies reported more positive findings across cognitive, affective, and behavioural domains, including improvements in oral performance, motivation, and intercultural awareness (e.g., Derakhshan et al., 2024; Hoang et al., 2023; Hsu et al., 2022). HMD-based VR showed a less consistent pattern. Some studies reported selective gains, such as improvement in lower-level reading without deeper reflection, or improvement in writing while spelling errors remained.

Other studies found gaps between learners' positive perceptions and measured performance, particularly in anxiety reduction (e.g., Ding, 2024; Feng & Ng, 2023; Wang et al., 2021; York et al., 2021). Desktop and blended VR were less studied, but the available evidence reported positive outcomes in speaking performance and self-efficacy (e.g., Le et al., 2024; Li et al., 2024; Xie et al., 2022; Yan et al., 2024). Notably, intervention duration differed across VR types. Most HMD-based studies were short-term or one-off. Mobile VR was more often used in longer-term implementations. Differences in task design, teacher support, and exposure time may therefore partly explain the outcome patterns.

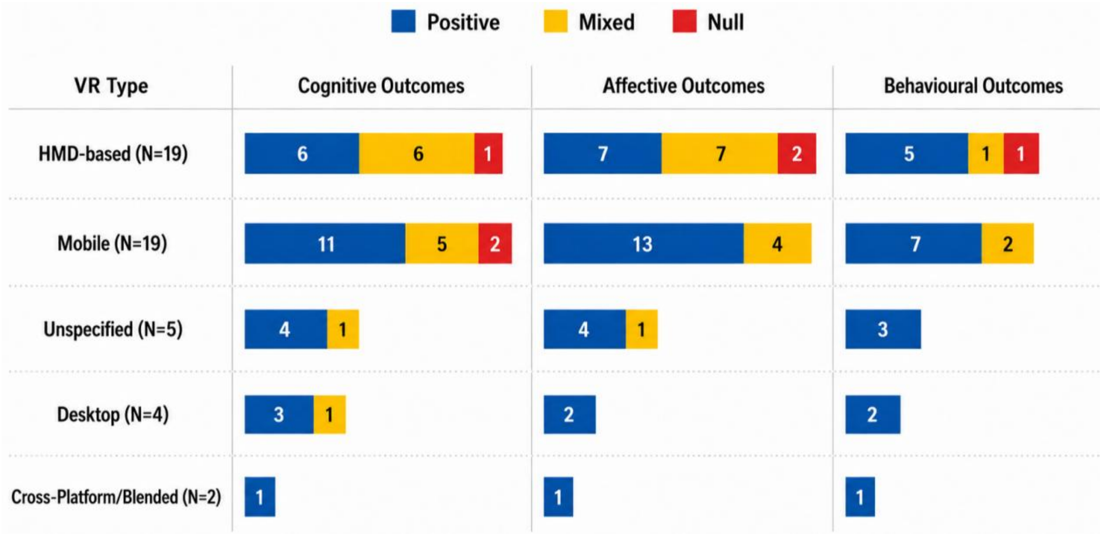

Figure 11: Analysis of VR Type and Cognitive, Affective and Behavioural Outcomes

Note: N = number of studies. Cross-platform/Blended refers to instruction integrated VR with additional digital platforms.

b. Instructional Setting

Outcomes also varied across instructional settings (see Figure 12). Classroom-based studies reported the most consistent positive outcomes, particularly for behavioural measures, with few null findings. Lab-based studies showed more mixed cognitive outcomes, while affective and behavioural outcomes were mostly positive (e.g., Feng & Ng, 2024; Kaplan-Rakowski & Gruber, 2023; Nai & Hassan, 2022). Findings from other settings were generally positive but based on fewer studies. Online VR studies reported positive cognitive outcomes, with more variable affective results. Blended and self-directed implementations also reported positive outcomes across domains (e.g., Chan, 2024; Xie et al., 2022). One summer workshop reported positive findings, but its narrow scope limits how far the result can be generalised.

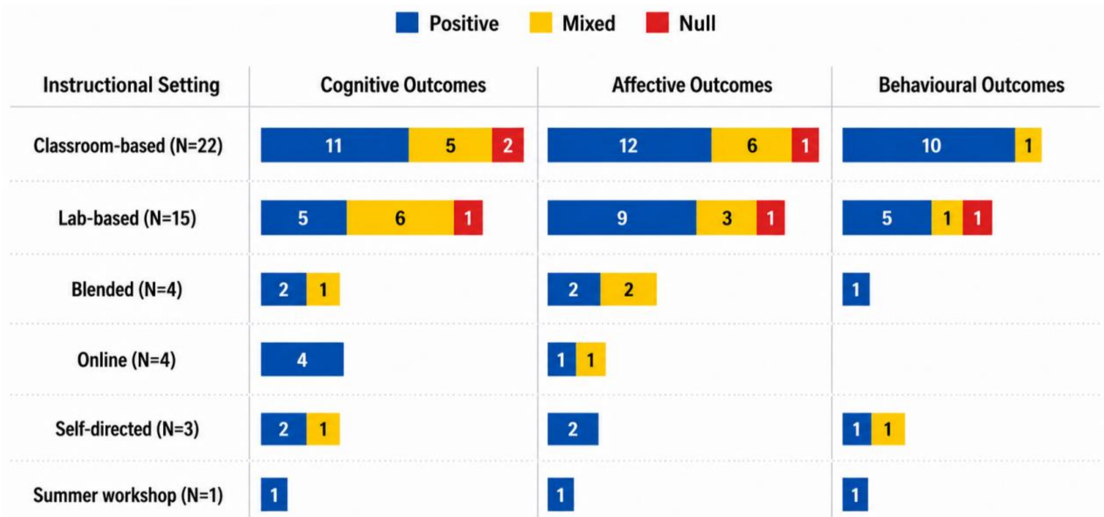

Figure 12: Analysis of Instructional Setting and Cognitive, Affective and Behavioural Outcomes

Note: N = number of studies.

4. Discussion

4.1 What VR Changes: Cognitive, Affective, and Behavioural Outcomes

The reviewed findings show a tension between immersive engagement and cognitive processing demands. Across the studies, VR was more consistently associated with engagement, motivation, and meaning-focused performance than with linguistic accuracy, pronunciation, or higher-order cognitive processes. This pattern suggests that VR is better suited to contextualized language use than to tasks that require sustained attention to linguistic form or deeper analysis. This pattern is broadly consistent with CAMIL, in which presence and agency shape cognitive and affective processes during immersive learning. In the reviewed studies, VR seemed to create conditions in which learners were more actively involved in purposeful language use.

The uneven cognitive findings can also be interpreted from a cognitive load perspective. Visual and spatial cues in VR may support lower-level processing, such as vocabulary retention, information location, and basic comprehension. At the same time, multimodal input, navigation, and interaction demands may place additional pressure on learners' cognitive resources. This can leave fewer resources for form monitoring, error correction, and higher-order reflection (Skulmowski & Xu, 2022). Such a reading helps explain why VR was more often linked to meaning-focused performance than to linguistic accuracy or deeper cognitive processing. However, only a small number of studies directly measured cognitive load or examined its role in language-learning outcomes.

Affective outcomes also need careful interpretation. Many studies reported reduced anxiety, increased motivation, enjoyment, confidence, self-efficacy, or willingness to communicate in VR environments. These positive affective responses can support learning, but they do not automatically lead to improved performance or transfer to real-world communication contexts. Their value depends on whether they are paired with clear task structure, teacher guidance, and cognitive demands that learners can manage.

The reviewed studies also reported changes in learner behaviour, including increased participation, collaboration, and self-regulated learning. These behaviours often reflect learners' sense of agency as they explore virtual spaces, interact with peers, make choices, and produce language in response to contextual prompts. Presence may also intensify learners' involvement in the immediate learning situation. Still, behavioural engagement did not consistently correspond to language gains. It should therefore be treated as a condition that may support learning, not as evidence of language achievement by itself.

4.2 When and for Whom Is VR Most Effective: Key Factors

The effectiveness of VR-enhanced language learning appears to depend on how learner characteristics, instructional design, and contextual conditions work together. These factors affect how learners allocate cognitive resources in immersive environments and whether engagement turns into measurable learning.

Learner differences, especially proficiency level and disciplinary background, were associated with different outcome patterns. Lower-proficiency learners appeared to benefit from contextual support that reduced reliance on abstract processing. Intermediate learners seemed more susceptible to cognitive overload. Advanced learners were better positioned to engage in domain-specific and higher-level tasks. Learners in language-related disciplines also showed more consistent benefits. For learners in other fields, VR tasks seemed to work better when they were closely connected to the communicative situations of the discipline.

Instructional design whether VR activities supported learning or only produced engagement. Active and constructive tasks tended to promote deeper processing because learners had to generate, organise, and use language. Passive tasks gave learners fewer opportunities to do this. Teacher presence also mattered. Instructor-led approaches were more consistently associated with positive outcomes because they combined procedural, conceptual, and strategic support. Technical facilitation alone was less likely to support conceptual understanding, even when learners responded positively to the VR experience. Duration was another factor. Longer interventions gave learners time to become familiar with the VR environment, which may have reduced cognitive demands and allowed more attention to language learning. Short-term or one-off interventions are more likely to capture initial interest or novelty, but they were less reliable in producing sustained cognitive gains.

Contextual factors also shaped the reported outcomes. More accessible technologies, such as mobile VR, appeared to support more consistent outcomes, possibly because they reduce interaction costs. More immersive systems, especially HMD-based VR, may have added cognitive demands through navigation, multimodal input, and interaction. Instructional settings showed a similar pattern. Structured learning environments appeared to support more stable outcomes because learners had external guidance from an instructor. Less structured contexts, including self-directed learning contexts, required greater self-regulation and may produce more variable outcomes. These findings suggest that VR-supported language learning depends on fit rather than immersion alone. VR is more likely to support EFL learning when the task, learner group, technology, and teaching arrangement are aligned.

4.3 Pedagogical Implications

VR should be planned as part of language pedagogy rather than added as a technological feature. The choice and sequence of VR tasks need to reflect learners' proficiency levels and disciplinary backgrounds. Teachers should therefore avoid assigning the same VR task to all learners without adjustment. Task complexity, scaffolding, and communicative focus can be adapted to different learner groups. For learners in non-language disciplines, VR tasks should be tied to communicative situations they are likely to encounter in their fields, so that the virtual environment supports purposeful language use rather than general exploration.

Task design and teacher involvement are central to effective VR use. VR appears most useful when it supports active and constructive tasks in which learners generate, organise, and communicate meaning. These tasks still require instructional guidance. Without it, learners may spend too much attention on navigation, interface use, or interaction management, leaving less attention for language learning. Scaffolding should therefore include procedural support for using the system, conceptual support for understanding language content, and strategic support for completing the task. This is especially important in the early stages of VR use. Longer implementation is also preferable where possible, because repeated exposure gives learners time to become familiar with the environment and focus more fully on the language task.

Implementation decisions should be made according to local teaching conditions. The choice of VR technology needs to take account of learner readiness, instructional goals, available infrastructure, teacher readiness, and institutional support. Highly immersive systems may enhance engagement, but they may also add cognitive, physical, and logistical demands. In settings where high-end systems are difficult to maintain, mobile VR or other accessible platforms may offer a more practical starting point. VR is pedagogically useful when the technology, task, support, and teaching context fit together. Immersion by itself is not enough.

5. Conclusion

This systematic review synthesised 49 empirical studies on VR-enhanced EFL instruction in higher education. It examined reported cognitive, affective, and behavioural outcomes, as well as the conditions under which VR appeared to support language learning. The findings indicate that VR was more consistently associated with behavioural engagement and affective outcomes than with cognitive gains. Cognitive outcomes were uneven and more often favoured meaning-focused performance over linguistic accuracy. VR effectiveness therefore depends less on immersion itself than on how well instructional design, learner characteristics, and contextual conditions fit together.

More favourable outcomes were reported in studies involving active and constructive task designs, strong teacher presence, combined scaffolding strategies, and longer intervention duration. These conclusions, however, should be read with caution. The included studies were concentrated in Asian EFL contexts, many used small-scale quasi-experimental designs, and several relied on self-reported measures. Reporting was also inconsistent for participant characteristics and technological specifications. Although this review did not conduct a formal quality appraisal, the eligibility criteria included some quality-related thresholds, such as excluding quantitative intervention studies without a control or comparison condition or with insufficient sample sizes. The positive, mixed, and null classification was descriptive and could not fully account for evidence strength, effect size, or overlapping study characteristics. The review was also limited to English-language publications, which may have introduced language bias.

Future research should use longitudinal and large-scale experimental designs to examine whether VR-related learning gains persist beyond initial novelty effects. Studies also need to report learner characteristics, prior VR experience, technological specifications, instructional design, intervention procedures, and outcome measures more consistently. A minimal reporting checklist for VR-based EFL interventions would improve transparency and make comparisons across studies more meaningful. Broader inclusion of different educational contexts is also needed, especially from settings outside the regions most represented in the current literature.

Overall, this review suggests that VR has potential value for EFL instruction when immersive affordances are aligned with pedagogical design, learner needs, and contextual conditions. As VR technologies continue to evolve, their educational value will depend on how purposefully they are integrated to support learners' cognitive and communicative development.

Conflict of interest statement

The authors declare no conflict of interest.

7. Funding

The authors declare that financial support was received for the research and publication of this article. This research was supported by the Faculty of Education, Universiti Kebangsaan Malaysia. (Grant Number: GG-2024-082).

8. Acknowledgments

The authors would like to acknowledge the use of ChatGPT for language editing and improving the clarity of the manuscript. This paper remains an accurate representation of the authors' work and intellectual contributions.

9. References:

Akay, M., & Kessler, M. (2024). The impact of virtual reality (VR) pedagogy on L2 English learners' oral communication and pragmatic competence. *Language Learning & Technology*, *28*(1), 1–11. https://hdl.handle.net/10125/73568

Bacca-Acosta, J., Tejada, J., Fabregat, R., Kinshuk, & Guevara, J. (2022). Scaffolding in immersive virtual reality environments for learning English: an eye tracking study. *Education Technology Research Development* 70, 339–362. https://doi.org/10.1007/s11423-021-10068-7

Bacca-Acosta, J., Tejada, J., Lampropoulos, G., & Avila, C. (2023). A Comparative Eye-Tracking Study Between a Virtual Reality System and a Desktop Environment for Learning the Prepositions of Place in English. *CALICO Journal*, *40*(1), 68–90. https://doi.org/10.1558/cj.22803

Baxter, G., & Hainey, T. (2024). Using immersive technologies to enhance the student learning experience. *Interactive Technology and Smart Education*, *21*(3), 403-425. https://doi.org/10.1108/ITSE-05-2023-0078

Bolton, K., & Bacon-Shone, J. (2020). The Statistics of English across Asia. In K. Bolton, W. Botha, & A. Kirkpatrick (Eds.), *The Handbook of Asian Englishes* (1st ed., pp. 49–80).

Wiley. https://doi.org/10.1002/9781118791882.ch3

Budianto, L., Suparmi, S., Biyanto, B., & Azkiyah, S. N. (2023). Virtual Reality Games in EFL Class: Examining Learners' Vocabulary Learning. *The Journal of AsiaTEFL*, *20*(4), 942–951. https://doi.org/10.18823/asiatefl.2023.20.4.14.942

Chan, V. (2023). Investigating the impact of a virtual reality mobile application on learners' interpreting competence. *Journal of Computer Assisted Learning*, *39*(4), 1242–1258. https://doi.org/10.1111/jcal.12796

Chan, V. (2024). Investigating students' perceptions of the use of a multifunctional mobile-based virtual reality application for English/Chinese interpreting learning. *International Journal of Mobile Learning and Organisation*, *1*(1), 304–322. https://doi.org/10.1504/IJMLO.2024.10056650

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

Chen, C.-Y., Chang, S.-C., Hwang, G.-J., & Zou, D. (2023). Facilitating EFL learners' active behaviors in speaking: A progressive question prompt-based peer-tutoring approach with VR contexts. *Interactive Learning Environments*, *31*(4), 2268–2287. https://doi.org/10.1080/10494820.2021.1878232

Chen, M.-R. A., & Hwang, G.-J. (2022). Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles. *Interactive Learning Environments*, *30*(9), 1619–1639. https://doi.org/10.1080/10494820.2020.1734626

Chen, Y. (2024). 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

Cheng, H.-Y., Tian, R.-M., & Mei, X.-X. (2023). Comparing the behavioural differences of teachers and students in learner-centred classrooms between China and the USA. *European Journal of Psychology of Education*, *38*(1), 63–82. https://doi.org/10.1007/s10212-021-00592-8

Chi, M., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Chorão, G., Tavares, C., & Duarte, P. (2023). Enhancing Speaking Skills in a Foreign Language Using Virtual Reality. A Study in a Portuguese Higher Education Institution. In A. Mesquita, A. Abreu, J. V. Carvalho, C. Santana, & C. H. P. De Mello (Eds.), *Perspectives and Trends in Education and Technology* (Vol. 366, pp. 701–712). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-5414-8_64

Clark, R. E. (1983). Reconsidering Research on Learning from Media. *Review of Educational Research*, 53(4), 445–459. https://doi.org/10.3102/00346543053004445

Derakhshan, A., Teo, T., & Khazaie, S. (2024). Is game-based language learning general or specific-oriented? Exploring the applicability of mobile virtual realities to medical English education in the middle east. *Computers & Education*, *213*, 105013. https://doi.org/10.1016/j.compedu.2024.105013

Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A Systematic Review of Research on High-Immersion Virtual Reality for Language Learning. *TechTrends*, 66(5), 810–824. https://doi.org/10.1007/s11528-022-00717-w

Di Natale, A. F., Repetto, C., Riva, G., & Villani, D. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. *British Journal of Educational Technology*, 51(6), 2006–2033. https://doi.org/10.1111/bjet.13030

Ding, M. (2024). The impact of high-immersion virtual reality on EFL learners' foreign language speaking anxiety: A mixed-method approach. *ReCALL*, *36*(3), 287–305.

https://doi.org/10.1017/S0958344024000156
Ebadi, S., & Ebadijalal, M. (2022). The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency. *Computer Assisted Language Learning*, *35*(8), 1975–2000. https://doi.org/10.1080/09588221.2020.1854311
Ebadijalal, M., & Yousofi, N. (2024). 'Take me to a virtual trip if you want me to write better!': The impact of Google Expeditions on EFL learners' writing motivation and performance. *Computer Assisted Language Learning*, *37*(7), 1806–1828. https://doi.org/10.1080/09588221.2022.2123001
Ece, Ç., Balkan, Ö., & Dinçer, N. (2023). Immersive Virtual Reality in Tertiary Level EFL Education: A Systematic Review of Recent Applications. *Innovational Research in ELT*, 4(2), 59–69. https://doi.org/10.29329/irelt.2023.623.6
Feng, B., & Ng, L. L. (2023). Facilitating writing performance of EFL learners via virtual reality: Immersion, presence, embodiment. *Frontiers in Psychology*, *14*, 1134242. https://doi.org/10.3389/fpsyg.2023.1134242
Feng, B., & Ng, L.-L. (2024). The Spatial Influence on Vocabulary Acquisition in an Immersive Virtual Reality-Mediated Learning Environment: *International Journal of Computer-Assisted Language Learning and Teaching*, *14*(1), 1–17. https://doi.org/10.4018/IJCALLT.339903
Gao, L., Wan, B., Liu, G., Xie, G., Huang, J., & Meng, G. (2021). Investigating the Effectiveness of Virtual Reality for Culture Learning. *International Journal of Human–Computer Interaction*, *37*(18), 1771–1781. https://doi.org/10.1080/10447318.2021.1913858
Hannafin, M., Land, S., & Oliver, K. (2013). Open learning environments: Foundations, methods, and models. In C. M. Reigeluth (Ed.), *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory* (Vol. 2, pp. 115–140). Routledge.
Hoang, D. T. N., McAlinden, M., & Johnson, N. F. (2023). Extending a learning ecology with virtual reality mobile technology: Oral proficiency outcomes and students' perceptions. *Innovation in Language Learning and Teaching*, *17*(3), 491–504. https://doi.org/10.1080/17501229.2022.2070626
Hoi, V. N., & Mu, G. M. (2021). Perceived teacher support and students' acceptance of mobile-assisted language learning: Evidence from Vietnamese higher education context. *British Journal of Educational Technology*, 52(2), 879–898. https://doi.org/10.1111/bjet.13044
Hsu, C.-C., Chen, Y.-L., Lin, C.-Y., & Lien, W. (2022). Cognitive development, self-efficacy, and wearable technology use in a virtual reality language learning environment: A structural equation modeling analysis. *Current Psychology*, *41*(3), 1618–1632. https://doi.org/10.1007/s12144-021-02252-y
Hsu, L. (2024). Exploring EFL learners' acceptance and cognitive absorption at VR-Based language learning: A survey and experimental study. *Heliyon*, *10*(3), e24863. https://doi.org/10.1016/j.heliyon.2024.e24863
Huang, H. L. (2024). Exploring the Impact of VR Scaffolding on EFL Teaching and Learning: Anxiety Reduction, Perceptions, and Influencing Factors. *Multimodal Technologies and Interaction*, *8*(10), 85. https://doi.org/10.3390/mti8100085
Hung, S.-T. A., Chen, W.-J., & Chien, S.-Y. (2024). Virtual reality is not always a cure-all: Evidence from a quasi-experiment of EFL business speaking courses. *Interactive Learning Environments*, *32*(8), 4426–4442. https://doi.org/10.1080/10494820.2023.2200811
Jwaiʿed, A. M., Masri, A. A., Hijazi, D., & Smadi, M. (2024). Utilizing Virtual Reality (VR) and Augmented Reality (AR) Technologies in EFL Classrooms: A Novel Approach to Improve Vocabulary Learning and Retention. *Journal of Ecohumanism*, *3*(7), 22–33. https://doi.org/10.62754/joe.v3i7.4171
Kachru, B. B. (1990). World Englishes and applied linguistics. *World Englishes*, 9(1), 3–20.

https://doi.org/10.1111/j.1467-971X.1990.tb00683.x
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
Kaplan-Rakowski, R., & Gruber, A. (2024). An experimental study on reading in high-immersion virtual reality. *British Journal of Educational Technology*, *55*(2), 541–559. https://doi.org/10.1111/bjet.13392
Khodabandeh, F. (2022). Exploring the applicability of virtual reality- enhanced education on extrovert and introvert EFL learners' paragraph writing. *International Journal of Educational Technology in Higher Education*, *19*(1), 27. https://doi.org/10.1186/s41239-022-00334-w
Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159. https://doi.org/10.2307/2529310
Le, V. H. H., Prabjandee, D., & Kewara, P. (2024). Virtual Reality: A Virtual World to Realize English as a Lingua Franca and Enhance Students' Self–Efficacy. *Computer Assisted Language Learning Electronic Journal*, *25*(2), 152–172. https://callej.org/index.php/journal/article/view/449
Li, X., Lan, Y., Pi, Z., Qi, G. Y., Grant, S., & Sun, J. (2024). Pedagogical agent positioning in external videos improves English academic presentation proficiency in desktop virtual reality settings. *British Journal of Educational Technology*, 13531. https://doi.org/10.1111/bjet.13531
Li, Y., Ying, S., Chen, Q., & Guan, J. (2022). An Experiential Learning-Based Virtual Reality Approach to Foster Students' Vocabulary Acquisition and Learning Engagement in English for Geography. *Sustainability*, *14*(22), 15359. https://doi.org/10.3390/su142215359
Lin, V., Barrett, N. E., Liu, G.-Z., Chen, N.-S., & Jong, M. S.-Y. (2023). Supporting dyadic learning of English for tourism purposes with scenery-based virtual reality. *Computer Assisted Language Learning*, *36*(5–6), 906–942. https://doi.org/10.1080/09588221.2021.1954663
Lin, Y.-J., & Wang, H. (2021). Using virtual reality to facilitate learners' creative self-efficacy and intrinsic motivation in an EFL classroom. *Education and Information Technologies*, *26*(4), 4487–4505. https://doi.org/10.1007/s10639-021-10472-9
Liu, C.-C., Guo, Y., Hwang, G.-J., Tu, Y.-F. & Wang, Z. (2024). Effects of an article-structure strategy-based spherical video-based virtual reality approach on EFL learners' English reading comprehension and learning conceptions. *Interactive Learning Environments* 32(6): 2647–2664. https://doi.org/10.1080/10494820.2022.2155840
Liu, X., & Hou, S. (2020). Effects of VR on Users' Speaking Performance and Cognitive Load: Design of VR-assisted English-Speaking Application on Mobile Phones. *2020 International Conference on Intelligent Design (ICID)*, 128–131. https://doi.org/10.1109/ICID52250.2020.00035
Ma, L. (2021). An Immersive Context Teaching Method for College English Based on Artificial Intelligence and Machine Learning in Virtual Reality Technology. *Mobile Information Systems*, *2021*, 1–7. https://doi.org/10.1155/2021/2637439
Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): A Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychology Review*, 33(3), 937–958. https://doi.org/10.1007/s10648-020-09586-2
Mubarok, H., Lin, C.-J., & Hwang, G.-J. (2024). A virtual reality-based collaborative argument mapping approach in the EFL classroom. *Interactive Learning Environments*, *32*(9), 4978–4996. https://doi.org/10.1080/10494820.2023.2207197
Nai, R., & Hassan, H. (2022). Multi-modal Simultaneous Interpreting Teaching: Based on Situated Learning in Virtual Reality. (2022). *Lecture Notes in Electrical Engineering*, 494–501. https://doi.org/10.1007/978-981-19-4132-0_61

Nowak, S., & Kaplan-Rakowski, R. (2024). The Impact of Relaxational Virtual Reality and Video Learning Activities on Subjective Vitality in a Blended Learning Environment. *International Journal on E-Learning*, *23*(4), 415–428. https://doi.org/10.70725/377570tszdgt

Ou Yang, F.-C., Lo, F.-Y.R., Hsieh, J.C. & Wu, W.-C.V. (2020). Facilitating Communicative Ability of EFL Learners via High-Immersion Virtual Reality. *Educational Technology & Society* 23(1): 30–49. https://doi.org/10.30191/ETS.202001_23(1).0003

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. https://doi.org/10.1136/bmj.n71

Parmaxi, A. (2023). 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

Peixoto, B., Bessa, L. C. P., Gonçalves, G., Bessa, M., & Melo, M. (2023). Teaching EFL With Immersive Virtual Reality Technologies: A Comparison with the Conventional Listening Method. *IEEE Access*, *11*, 21498–21507. https://doi.org/10.1109/ACCESS.2023.3249578

Porras, V. D. C. A., Olaya, C. I. B., & Cuervo, S. A. C. (2024). Exploring the impact of the metaverse on English language teaching: Perspectives of students from the University of Boyacá. *Journal of Infrastructure, Policy and Development*, *8*(7), 4779. https://doi.org/10.24294/jipd.v8i7.4779

Qiu, X. B., Shan, C., Yao, J., & Fu, Q. K. (2024). The effects of virtual reality on EFL learning: A meta-analysis. *Education and Information Technologies*, *29*(2), 1379–1405. https://doi.org/10.1007/s10639-023-11738-0

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

*Shadiev, R., Liu, J., & Yang, L. (2024). The effect of evaluation and feedback in 360-degree video-based VR environments on language learning outcomes. *Computer Assisted Language Learning*, 1–31. https://doi.org/10.1080/09588221.2023.2300990

Skulmowski, A., & Xu, K. M. (2022). Understanding Cognitive Load in Digital and Online Learning: A New Perspective on Extraneous Cognitive Load. *Educational Psychology Review*, 34(1), 171–196. https://doi.org/10.1007/s10648-021-09624-7

Song, C., Shin, S. Y., & Shin, K. S. (2023). Optimizing foreign language learning in virtual reality: a comprehensive theoretical framework based on constructivism and cognitive load theory (VR-CCL). *Applied Sciences*, *13*(23), 12557. https://doi.org/10.3390/app132312557

Sun, C., Yao, Y., & Wang, R. (2020). A Study on the Influence of Scene Reality of VR Environment on English Learners' Learning Engagement and Learning Effectiveness. *2020 IEEE 2nd International Conference on Computer Science and Educational Informatization (CSEI)*, 181–185. https://doi.org/10.1109/CSEI50228.2020.9142520

Tai, T.-Y. (2022). Impact of mobile virtual reality on EFL learners' listening comprehension. *Language Learning & Technology*, 26(1), 1–23. https://doi.org/10.64152/10125/73491

Utami, L. P. R. A., Suwastini, N. K. A., Dantes, G. R., Suprihatin, C. T., & Adnyani, K. E. K. (2021). Virtual reality for supporting authentic learning in 21st century language classroom. *Jurnal Pendidikan Teknologi dan Kejuruan*, 18(1), 132-141.

https://doi.org/10.23887/jptk-undiksha.v18i1.32376

Wang, Y., Luo, X., Liu, C.-C., Tu, Y.-F., & Wang, N. (2022). An Integrated Automatic Writing Evaluation and SVVR Approach to Improve Students' EFL Writing Performance. *Sustainability*, 14(18), 11586. https://doi.org/10.3390/su141811586

Wang, Z. (2021). On Computer Assisted Language Learning (CALL) and Change of Teachers' Role. *IOP Conference Series: Earth and Environmental Science*, 632(5), 052049. https://doi.org/10.1088/1755-1315/632/5/052049

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

Wu, B., Chang, X., & Hu, Y. (2023). A meta-analysis of the effects of spherical video-based virtual reality on cognitive and non-cognitive learning outcomes. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2023.2184389

Xie, Y., Liu, Y., Zhang, F., & Zhou, P. (2022). Virtual Reality-Integrated Immersion-Based Teaching to English Language Learning Outcome. *Frontiers in Psychology*, 12, 767363. https://doi.org/10.3389/fpsyg.2021.767363

Xin, Q., Alibakhshi, G., & Javaheri, R. (2025). A phenomenographic study on Chinese EFL teachers' cognitions of positive and negative educational, social, and psychological consequences of high-stake tests. *Scientific Reports*, *15*(1), 18829. https://doi.org/10.1038/s41598-025-03753-7

Yan, S., & Yang, Y. (2021). Education Informatization 2.0 in China: Motivation, Framework, and Vision. *ECNU Review of Education*, *4*(2), 410–428. https://doi.org/10.1177/2096531120944929

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, H., Tsung, L., & Cao, L. (2022). The Use of Communication Strategies by Second Language Learners of Chinese in a Virtual Reality Learning Environment. *Sage Open*, 12(4), 21582440221141877. https://doi.org/10.1177/21582440221141877

Yeh, H.-C., Tseng, S.-S., & Heng, L. (2022). Enhancing EFL students' intracultural learning through virtual reality. *Interactive Learning Environments*, *30*(9), 1609–1618. https://doi.org/10.1080/10494820.2020.1734625

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

Yuan, J., Liu, Y., Han, X., Li, A., & Zhao, L. (2023). Educational metaverse: An exploration and practice of VR wisdom teaching model in Chinese Open University English course. *Interactive Technology and Smart Education*, 20(3), 403–421. https://doi.org/10.1108/ITSE-10-2022-0140

Zhang, M. M., Hashim, H., & Yunus, M. M. (2025). Analyzing and comparing augmented reality and virtual reality assisted vocabulary learning: A systematic review. *Frontiers in Virtual Reality*, 6, 1522380. https://doi.org/10.3389/frvir.2025.1522380

Appendix

Table A1: Search Strings

Database	Search string
WoS	TS=((VR OR "virtual reality" OR "immersive technology" OR "mixed reality" OR MR OR "extended reality" OR XR OR "3D virtual environment" OR metaverse) AND (EFL OR "English as a foreign language" OR "English language teaching" OR "language learning" OR "foreign language teaching" OR "foreign language learning" OR "L2 learning" OR "second language acquisition") AND ("higher education" OR college OR university OR "tertiary education" OR undergraduate* OR student* OR learner*))
Scopus	TITLE-ABS-KEY((VR OR "virtual reality" OR "immersive technology" OR "mixed reality" OR MR OR "extended reality" OR XR OR "3D virtual environment" OR metaverse) AND (EFL OR "English as a foreign language" OR "English language teaching" OR "language learning" OR "foreign language teaching" OR "foreign language learning" OR "l2 learning" OR "second language acquisition") AND ("higher education" OR college OR university OR "tertiary education" OR undergraduate* OR student* OR learner*))

Table A2: Target Learning Domains

Domains	Specific domains	Studies	N
Foundational skills	Vocabulary	Budianto et al., 2023; Chen et al., 2021; Feng & Ng, 2024; Jwai‹ed et al., 2024; Li et al., 2022; Nowak & Kaplan-Rakowski, 2024	6
	Grammar	Bacca-Acosta et al., 2022, 2023	2
Productive skills	Speaking	Akay & Kessler, 2024; Chen, 2024; Chen et al., 2023; Chen & Hwang, 2022; Chorão et al., 2023; Ebadi & Ebadijalal, 2022; Hoang et al., 2023; Huang, 2024; Liu & Hou, 2020; Mubarok et al., 2024; Shadiev et al., 2024; Xie et al., 2022; Yan et al., 2024; Yuan et al., 2023	14
	Writing	Ebadijalal & Yousofi, 2024; Feng & Ng, 2023; Khodabandeh, 2022; Wang et al., 2022	4
Receptive skills	Reading	Kaplan-Rakowski & Gruber, 2024; Liu et al., 2024; Wang et al., 2021	3
	Listening	Peixoto et al., 2023	1
Advanced/ Specialized Skills	ESP/EAP	Derakhshan et al., 2024; Lin et al., 2023; Hsu et al., 2022; Hung et al., 2024; Hsu, 2024; Li et al., 2024	6
	Interpreting	Chan, 2023, 2024; Nai & Hassan, 2022	3
	Cultural competence	Gao et al., 2021; Yeh et al., 2022	2
Integrated Skills	Multiple skills	Ma, 2021; Porras et al., 2024; Sun et al., 2020	3
Affective gains	Anxiety/ Efficacy/ Motivation	Ding, 2024; Kaplan-Rakowski & Gruber, 2023; Le et al., 2024; Lin & Wang, 2021; York et al., 2021	5
Total			49

Notes: N = number of studies. Each study is included in only one category according to the main target learning domain. Affective gains include only studies without any language skills measurement.

Table A3: Summary of Studies Reporting Cognitive, Affective, and Behavioural Outcomes

Study ID	Authors (Year)	Cognitive Outcome	Affective Outcome	Behavioural Outcome	Main Finding
1	Bacca-Acosta et al. (2023)	√		√	VR improved learning prepositions; greater improvement in second session.
2	Sun et al. (2020)		√		High-reality VR increased engagement; linked to learning effectiveness.
3	Mubarok et al. (2024)	√	√		VR-CAM improved oral, computational, creative thinking; boosted cultural interest.
4	Li et al. (2022)	√	√	√	VR improved vocabulary and theoretical understanding; better behavioural and social engagement. No significant difference in the emotional engagement.
5	Kaplan-Rakowski & Gruber (2024)	√	√		VR improved reading comprehension; increased motivation and presence. Cognitive load was not significantly different.
6	Ma (2021)	√	√		VR group had higher English scores, especially in listening and speaking; greater satisfaction.
7	Wang et al. (2022)	√	√	√	SVVR-AWE improved writing quality, motivation, self-efficacy; reduced anxiety; fostered independent writing.
8	Hsu et al. (2022)	√	√	√	English proficiency improved in all sub-skills (lexical, semantics, reading, syntax); advanced learners had higher self-efficacy and behavioural intention.
9	Yan et al. (2024)	√	√	√	Speaking performance improved in fluency, vocabulary, pronunciation, grammar; increased confidence, collaboration and reflective learning.
10	Yuan et al. (2023)	√	√	√	Higher oral English scores; more interest and presence; consistent weekly preparation time.
11	York et al.		√		All modalities (voice, video,

	(2021)				VR) reduced foreign language anxiety but there was no significant difference in the mean scores between the three modalities. VR was perceived as more fun, effective, easier for communication.
12	Liu et al. (2024)	√	√		VR improved reading comprehension, metacognitive awareness in global and support reading strategies but there was no significant difference in cognitive load or problem-solving strategies. VR stimulated extrinsic motivation with no significant difference for overall motivation, intrinsic motivation, or sense of presence.
13	Chen & Hwang (2022)	√	√		Compared with field-dependent learners, field-independent learners achieved higher oral presentation scores; fostered more professional presentation structures; reduced speaking anxiety; increased learning motivation.
14	Chen (2024)	√	√		No significant differences in the final oral performances. Only VR group had significant reduction in public speaking anxiety.
15	Liu & Hou (2020)	√	√		VR improved content summary, with no significant difference in fluency or vocabulary; lowered intrinsic cognitive load; higher germane load; accepted by learners.
16	Yeh et al. (2022)	√	√		VR improved intracultural awareness; presence boosted motivation to learn cultural sites.
17	Chorão et al. (2023)			√	VR improved eye contact, vocal clarity, pacing; but with some gaze imbalance.
18	Hsu (2024)	√	√		VR increased cognitive absorption, linked to

					comprehension; joy and control predicted immersion and curiosity did not predict immersion.
19	Khodabandeh (2022)	√			VR improved the paragraph writing performance of both extroverts and introverts, with no significant difference between the two personality types.
20	Porras et al. (2024)	√	√		Metaverse improved vocabulary, grammar, pronunciation, communication skills, fluency and real-world applicability; boosted motivation, confidence, and reduced anxiety.
21	Huang (2024)	√	√	√	VR improved vocabulary, contextual use; reduced foreign language anxiety; more peer interaction and English use.
22	Hoang et al. (2023)	√	√		VR significantly improved overall oral proficiency, with notable gains in fluency and coherence, lexical resource, and pronunciation, and a marginal improvement in grammatical range and accuracy; VR enhanced motivation and confidence.
23	Chen et al. (2023)	√	√	√	VR group had better speaking performance, self-efficacy; more active and interactive peer tutoring behaviours; no significant differences in learning motivation or speaking-related anxiety.
24	Feng & Ng (2023)	√			IVR improved writing quality with more target words, higher lexical density and greater distribution richness; more spelling mistakes in Task 2.
25	Chan (2024)	√	√	√	VR improved interpreting competence, motivation, confidence; fostered independent learning; with low peer collaboration.
26	Gao et al.	√	√	√	No significant differences in

	(2021)				cultural knowledge, vocabulary retention, attitudes, or gift exchange behaviours.
27	Chan (2023)	√			Improved interpreting performance, especially in harder tasks (consecutive interpreting and English-Chinese interpreting); increased declarative knowledge.
28	Derakhshan et al. (2024)	√			VR improved English for Medical and Academic Purposes speaking and writing skills.
29	Nai & Hassan (2022)	√	√	√	VR improved interpreting scores; reduced anxiety; increased interest in interpreting learning; increased vocabulary preparation and practice enthusiasm.
30	Li et al. (2024)	√		√	Out-of-video Pedagogical Agent in VR improved English for Academic Purposes scores, especially speech presentation; more active content-related behaviors.
31	Bacca-Acosta et al. (2022)	√		√	Dynamic IVR improved learning; students used more scaffolds and improved efficiency.
32	Nowak & Kaplan-Rakowski (2024)		√		No significant subjective vitality difference: VR increased affective engagement.
33	Lin et al. (2023)	√	√	√	Speaking improved in language accuracy, destination promotion, knowledge retention, and replication of English for Tourism Purposes language. Writing improved in sophisticated vocabulary use, knowledge retention, and replication of ETP language, but showed no significant difference in promoting destinations. Higher self-efficacy; mimicked tour guide language.

34	Ebadijalal & Yousofi (2024)	√	√	√	Improved writing in task achievement, lexical resource, and grammatical range and accuracy but not in coherence and cohesion; increased motivation, confidence; encouraged autonomous learning; reduced anxiety.
35	Peixoto et al. (2023)	√	√		Interactive iVR had higher knowledge test scores than passive iVR; no significant difference in knowledge retention compared with the conventional method; high satisfaction and motivation.
36	Wang et al. (2021)	√	√		VR group had higher reading scores and better lower-level reading skills (information location, text comprehension); no significant improvement in higher-level skills (evaluation, reflection); increased motivation; reduced anxiety.
37	Shadiev et al. (2024)	√	√		360-degree video technology improved speaking; increased motivation and engagement, but some physical discomfort.
38	Ebadi & Ebadijalal (2022)	√	√	√	VR improved oral performance in fluency, coherence and lexical resource but not in grammatical range and accuracy or pronunciation, increased willingness and motivation. Students prepared less due to ready-made sentences, allowing more practice time, and added more descriptions in presentations, but some focused on VR and ignored what peers were doing and saying.
39	Kaplan-Rakowski & Gruber (2023)		√		VR reduced foreign language anxiety more than Zoom; helped overcome stage fright.
40	Ding (2024)		√		No significant Foreign

					Language Speaking Anxiety change: VR perceived as engaging and useful; transfer to real-life limited.
41	Akay & Kessler (2024)	√	√		VR slightly improved pragmatic competence: enjoyable but technical challenges noted.
42	Feng & Ng (2024)	√	√	√	VR improved vocabulary acquisition and retention; authentic context boosted motivation. Words in frequently visited locations and at natural eye level had better retention. Learners interacted briefly but revisited words often, showing exploratory and incidental learning behaviours.
43	Lin & Wang (2021)		√	√	Creative thinking efficacy improved significantly, but no gains in creative production or persistence under negative feedback. Students with high creative self-efficacy reported higher intrinsic motivation and less pressure. VR allowed free creation; enriched learning experience.
44	Jwai‘ed et al. (2024)	√			VR and AR improved vocabulary acquisition and retention; VR highest in delayed post-test.
45	Budianto et al. (2023)	√		√	Interactive VR game improved vocabulary: learners used multiple learning strategies (associating, predicting and outsourcing).
46	Chen et al. (2021)	√	√	√	VR improved specialized vocabulary but with no significant gain in problem-solving performance; boosted motivation in terms of promotion-instrumentality and attitude but no significance in ideal L2 self, ought-to L2 self, prevention-instrumentality and English anxiety; encouraged active

					knowledge creation.
47	Hung et al. (2024)	√	√		No significant oral proficiency. No significant differences in speaking anxiety or emotions, though interviews suggested VR's safe environment could help reduce anxiety.
48	Le et al. (2024)		√		VR improved general and listening self-efficacy.
49	Xie et al. (2022)	√		√	VR improved learning efficiency and speed; enhanced listening, speaking, reading, and writing, with AI speech recognition aiding pronunciation correction; deepened topic memory and comprehension more than text or images; enabled active, autonomous, applied learning.

Note: "√" indicates that the study reported at least one finding in the corresponding outcome category; blank cells indicate that no findings were reported for that category.