



Contents lists available at [Elora Center](#)

Lentera Negeri

Journal homepage: <https://lentera.eloracenter.org/lentera>



Effectiveness of immersive virtual reality interventions for l2 english oral communication: a systematic review

Kristian Burhan

Universitas Negeri Padang, Indonesia

Article Info

Article history:

Received Sep 06th, 2026

Revised Sep 06th, 2026

Accepted Sep 07th, 2026

Keyword:

Immersive virtual reality,
English speaking,
Willingness to communicate,
Foreign language speaking
anxiety,
Computer-assisted language
learning

ABSTRACT

Speaking is one of the most demanding skills for English learners, who often lack authentic, low-anxiety opportunities to practise oral communication. This systematic review synthesises empirical evidence on immersive virtual reality (VR) for L2 English speaking and oral communication. Following PRISMA 2020, searches of Scopus, PubMed, and ScienceDirect identified 948 records; after 11 duplicates were removed, 937 records were screened, 162 full texts were assessed, and 24 empirical studies published between 2021 and 2025 were included. Eligible studies examined English-speaking or oral-communication outcomes in EFL/ESL/ESP learners using immersive VR and were appraised with the Mixed Methods Appraisal Tool (MMAT 2018). Because the included studies differed substantially in design, outcomes, instruments, and reporting, findings were synthesised narratively and thematically rather than pooled statistically. The corpus was dominated by mixed-methods and quasi-experimental studies and was concentrated in Asian higher-education contexts. Across the 24 studies, confidence, willingness to communicate, and engagement showed the most consistently positive descriptive pattern, whereas effects on measured oral proficiency and speaking anxiety were mixed. The evidence therefore supports task-based, scaffolded, and socially supported VR use, while cautioning against equating positive perceptions with demonstrated proficiency gains. Formal inter-rater agreement statistics were not recorded in the original review and are identified as a methodological limitation.



© 2026 The Authors.

This is an open access article under the CC BY-NC-SA license
(<https://creativecommons.org/licenses/by-nc-sa/4.0>)

Corresponding Author:

Kristian Burhan,

✉ kristianburhan@unp.ac.id

Introduction

Oral communication in English has become a decisive competence in globalised education, employment, and mobility, yet it is widely regarded as the hardest skill for foreign-language learners to master. In many English as a foreign language (EFL) and English as a second language (ESL) contexts, learners study grammar and vocabulary extensively but rarely encounter authentic situations in which they must speak spontaneously. Digital technologies have long been mobilised to close this gap, and recent syntheses document a rapid expansion of immersive and intelligent tools in language education (Nuryana et al., 2025; Yi et al., 2025). Against this backdrop, immersive virtual reality (VR) has emerged as a particularly promising medium because it can place learners inside believable, interactive scenes where speaking is purposeful rather than performative.

Immersive VR refers to computer-generated environments-delivered through head-mounted displays, 360° video, social VR, metaverse platforms, or desktop simulations-that surround users with multimodal,

spatially situated stimuli and allow embodied interaction. In this review, presence refers to the learner's experienced sense of being located and able to act within the mediated environment, whereas authenticity refers to the extent to which speaking tasks simulate recognisable communicative goals, roles, and contexts. The hypothesised pathway is that greater presence and contextual authenticity can reduce perceived social-evaluative risk and increase opportunities for purposeful oral production; repeated supported production may then contribute to speaking proficiency. However, this review does not treat presence or authenticity as established mediators unless an included study explicitly measured them. In studies where these constructs were inferred only from intervention design, they are interpreted as pedagogical affordances rather than demonstrated causal mechanisms. This distinction is important because oral communication depends heavily on context, motivation, affect, and opportunities for meaningful practice.

A substantial body of empirical work has examined immersive VR across the four skills and their affective correlates. Reviews of the broader field report growth in studies addressing vocabulary, reading, listening, and writing, alongside psychological constructs such as motivation and anxiety (Esfandiari & Mazharpour, 2025; Tegoan et al., 2025). Empirical studies illustrate this breadth: immersive VR flashcards and self-developed environments have been used for vocabulary acquisition (Santoso et al., 2025; Luan et al., 2025; Chen & Yuan, 2025; Zhang et al., 2025), spherical video VR for writing complexity (Shen et al., 2025), and 360° and metaverse environments for listening (Shokri & Rahimi, 2025; Çelik & Ersanlı, 2025). This diversity, however, complicates any attempt to draw firm conclusions about a single skill.

Methodological and technological advances have accelerated in the last five years. Meta-analyses now quantify the effects of technology and augmented reality on language gains (Yi et al., 2025; Yang & Zhang, 2025), while reviews chart the migration from text-based virtual worlds toward artificial-intelligence-enhanced metaverse ecosystems (Tegoan et al., 2025). Standalone consumer devices and off-the-shelf platforms such as ImmerseMe, Google Expeditions, and social metaverse spaces have lowered the barrier to classroom adoption, and reviews of higher education and even preschool contexts confirm that immersive tools are diffusing across the educational lifespan (Esen, 2025; Fu & Su, 2025).

Despite this momentum, an important gap concerns the speaking skill itself. Much of the synthesised evidence aggregates heterogeneous outcomes, so the specific question of whether-and under what conditions-immersive VR improves oral communication is obscured. Recent systematic reviews of VR and augmented reality in English education tend to survey all skills together (Christou et al., 2025; Ma et al., 2025) or to focus on K–12 randomised trials without isolating speaking outcomes (Sun & Song, 2026). A dedicated synthesis of speaking is therefore warranted, particularly given how central oral communication is to workplace preparedness (Younus et al., 2025).

A second gap is theoretical and methodological. Studies invoke situated learning, the cognitive-affective model of immersive learning, and social-constructivist accounts, yet they rarely test these frameworks against one another, and designs range from single-group pilots to controlled experiments. Meta-analytic work cautions that effect sizes for technology in language learning are moderated by design quality and outcome type (Yi et al., 2025; Xue, 2025). Without a critical appraisal of how speaking is operationalised and measured, optimistic perceptions risk being mistaken for demonstrated learning gains.

The urgency of clarifying this evidence is heightened by rapid classroom adoption. Institutions are investing in VR laboratories and metaverse platforms on the assumption that immersion translates into better speaking, even though some rigorous studies report non-significant proficiency effects. Practitioners need an evidence-based account of what immersive VR reliably delivers for oral communication, for whom, and through which pedagogical designs, so that scarce resources are directed toward configurations that work rather than toward novelty alone (Esfandiari & Mazharpour, 2025).

To address these gaps, this systematic review synthesises empirical studies published between 2021 and 2025 that examine immersive VR for English speaking and oral communication. The first research question is:

RQ1: How does immersive virtual reality affect English (L2) learners' speaking and oral communication outcomes reported in empirical studies?

RQ2: Which speaking-related constructs-oral proficiency, willingness to communicate, speaking anxiety, confidence, pronunciation, and communicative competence-are most frequently examined, and what patterns of effect are reported?

RQ3: Which learner populations, VR technologies, and pedagogical designs appear most effective, and what methodological and contextual gaps remain for future research?

By concentrating on a single, high-stakes skill and appraising the methodological confidence of each study, this review contributes a focused, critical evidence base that complements broader multi-skill syntheses and informs both theory and classroom practice.

Method

Research design and framework

A systematic literature review (SLR) was adopted because the aim was to identify, appraise, and synthesise the existing empirical evidence on a clearly bounded question in a transparent and reproducible manner (Tranfield et al., 2003). The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement for its design, conduct, and reporting (Page et al., 2021), drawing additionally on the explanatory guidance for reporting systematic reviews (Liberati et al., 2009). PRISMA was selected because it provides a standardised flow for documenting identification, screening, eligibility, and inclusion, which strengthens the auditability of the selection process.

Search strategy

Prior to constructing the search string, the review question was decomposed using the PICOS framework (Population, Intervention/Interest, Comparison, Outcome, Study design) to ensure the systematic identification of core search concepts. Each element was translated into a cluster of synonyms and combined with Boolean operators, as summarised in Table 1.

Table 1. PICOS framework guiding the search strategy

Element	Specification
Population (P)	English (EFL/ESL/ESP) learners at any educational level
Intervention (I)	Immersive virtual reality: head-mounted/high-immersion VR, 360° video VR, social VR/metaverse, and desktop VR
Comparison (C)	Conventional (non-VR) instruction, lower-immersion conditions, or no comparator (single-group designs)
Outcome (O)	Speaking and oral communication outcomes (oral proficiency, willingness to communicate, speaking anxiety, confidence, pronunciation/fluency, communicative competence)
Study design (S)	Empirical primary studies: experimental, quasi-experimental, qualitative, and mixed-methods

The clusters were combined into the following Boolean string, adapted to the syntax of each database. In Scopus the string was applied to the TITLE-ABS-KEY field; equivalent title/abstract fields were used in PubMed and ScienceDirect. Truncation (the asterisk) captured morphological variants (e.g., *speak** retrieves *speak*, *speaking*, *speaker*), and quotation marks enforced exact phrases.

("virtual reality" OR "immersive VR" OR "360-degree video" OR "metaverse" OR "virtual world" OR "head-mounted display" OR XR) AND ("English" OR "EFL" OR "ESL" OR "foreign language" OR "second language" OR "L2") AND (speak OR "oral" OR "pronunciation" OR "willingness to communicate" OR "communicative competence" OR "verbal communication")*

Limiters were applied for document type (journal article and review), language (English), and publication period (2021–2025) to capture the most recent generation of consumer-grade immersive devices. The string was piloted and refined iteratively: an initial version returned a large volume of medical and surgical VR records, prompting the addition of language-specific terms to improve precision.

Database and information sources

Three complementary databases were searched: Scopus (multidisciplinary and the primary source), PubMed (to capture health- and psychology-oriented VR research), and ScienceDirect (for full-text Elsevier coverage). The searches were executed in a single session and exported as structured metadata. The initial search yielded 948 records: 690 from Scopus, 158 from PubMed, and 100 from ScienceDirect.

Eligibility criteria

Table 2. Inclusion and exclusion criteria

Criterion	Inclusion	Exclusion
Language	English	Non-English
Document type	Journal article; review	Conference abstract, editorial, book chapter
Publication period	2021–2025	Before 2021
Population	EFL/ESL/ESP learners	Non-language or clinical populations
Intervention	Immersive VR (HMD, 360°, social VR/metaverse, desktop VR)	Augmented-reality-only or non-immersive tools
Outcome	Speaking / oral communication outcome measured or evidenced	Outcome unrelated to speaking (e.g., reading-only)
Target language	English	Other target languages (e.g., Chinese, Spanish)
Study type	Empirical primary study	Reviews, bibliometrics, conceptual papers
Accessibility	Retrievable record with usable abstract	Records without sufficient information

Study selection process

Selection proceeded in three stages. First, duplicates were removed by matching digital object identifiers and normalised titles, eliminating 11 records and leaving 937 for screening. Second, titles, keywords, and abstracts were screened against the eligibility criteria; records that were clearly outside the immersive-VR-for-English-speaking scope-including medical VR, augmented-reality-only tools, non-English target languages, and non-empirical papers-were excluded, yielding 162 reports for full-text assessment. Third, the 162 reports were assessed for eligibility, and 138 were excluded because their primary outcome was not oral/speaking communication ($n = 102$), the intervention was not immersive VR or speaking was not a measured outcome ($n = 30$), or the target language was not English ($n = 6$). Twenty-four studies met all criteria. Where borderline decisions arose, they were resolved by re-reading the abstract against the PICOS specification. The full flow is depicted in Figure 1.

Quality assessment

Each included study was appraised with the Mixed Methods Appraisal Tool (MMAT), version 2018 (Hong et al., 2018), selected because the corpus contained qualitative, quantitative, and mixed-methods designs. For every study, the design category was first identified (qualitative, quantitative randomised, quantitative non-randomised, quantitative descriptive, or mixed methods), after which the five corresponding MMAT criteria were rated as met (Yes = 1), partially met/unclear ($P = 0.5$), or not met (No = 0). No study was excluded on the basis of appraisal results. The resulting ratings are used descriptively to judge the methodological confidence of the evidence and to qualify interpretation; they are not treated as a validated numerical risk-of-bias scale and are not used to generate artificial statistical weights. Twenty studies met the manuscript's high-confidence threshold ($\geq 4.0/5$), while four were classified as medium confidence ($3.0–3.9/5$). Appraisal was based on the information available in the included records and abstracts, so ratings should be interpreted cautiously until full-text verification. A formal inter-rater agreement statistic was not recorded in the original review; this absence is reported transparently as a limitation rather than replaced with an assumed value.

Table 3. Methodological quality appraisal of included studies (MMAT 2018)

Author(s), Year	MMAT category	S1	S2	S3	S4	S5	Total	Confidence
Jehma (2021)	Quant. non-randomised	Y	Y	Y	P	P	4.0	High
Alsaffar (2021)	Qualitative	Y	Y	Y	Y	P	4.5	High
Wu & Hung (2022)	Quant. non-randomised	Y	Y	Y	Y	Y	5.0	High
Ebadi & Ebadijalal (2022)	Mixed methods	Y	Y	Y	Y	Y	5.0	High
Lee & Ko (2023)	Quant. non-randomised	Y	Y	Y	Y	P	4.5	High
Raman et al. (2023)	Mixed methods	Y	Y	P	Y	P	4.0	High
Hoang et al. (2023)	Mixed methods	Y	Y	Y	Y	P	4.5	High
Chen et al. (2023)	Quant. non-randomised	Y	Y	Y	Y	P	4.5	High
Rababah (2023)	Quant. randomised	Y	Y	P	P	P	3.5	Medium
Fan et al. (2024)	Quant. randomised	Y	Y	Y	Y	P	4.5	High
Yudintseva (2024)	Mixed methods	Y	Y	Y	Y	P	4.5	High
Lai (2024)	Quant. non-randomised	Y	Y	Y	Y	P	4.5	High
Ha (2024)	Mixed methods	Y	Y	P	Y	P	4.0	High
Raman et al. (2024)	Qualitative	Y	Y	P	P	P	3.5	Medium
Ding (2024)	Mixed methods	Y	Y	Y	Y	Y	5.0	High
Akay & Kessler (2024)	Mixed methods	Y	Y	Y	Y	P	4.5	High
Hung et al. (2024)	Mixed methods	Y	Y	Y	Y	P	4.5	High
Hoter et al. (2025)	Mixed methods	Y	Y	P	Y	P	4.0	High

Author(s), Year	MMAT category	S1	S2	S3	S4	S5	Total	Confidence
Nuemailhom et al. (2025)	Quant. descriptive	Y	Y	P	P	P	3.5	Medium
Ratnah et al. (2025)	Qualitative	Y	Y	Y	Y	P	4.5	High
Li et al. (2025)	Quant. randomised	Y	Y	Y	Y	Y	5.0	High
Fathi et al. (2025)	Mixed methods	Y	Y	Y	Y	Y	5.0	High
Wang et al. (2025)	Mixed methods	Y	Y	Y	Y	P	4.5	High
Hwang & Lee (2025)	Qualitative	Y	Y	P	P	P	3.5	Medium

Note. S1–S5 denote the five MMAT criteria for the study's design category (Y = 1, P/unclear = 0.5, N = 0). "High" $\geq 4.0/5$; "Medium" = 3.0–3.9. No study was excluded on quality grounds.

Data extraction procedure

A structured extraction form captured author(s) and year, learner and study context, design, sample, VR technology or platform, educational level, targeted speaking outcome(s), and principal findings. Extraction was based on the bibliographic metadata and abstracts available in the reviewed records, and findings were recorded in analytic memos before being condensed for synthesis. Because the original review did not retain a formal second-rater agreement dataset, Cohen's kappa cannot be calculated retrospectively from the current evidence file. Any future revision should re-run screening and extraction with at least two independent reviewers and report agreement statistics before final submission.

Bibliometric and descriptive analysis

Descriptive analyses characterised the corpus by publication year, reported geographic context, research design, and the frequency of measured outcome constructs. These distributions (Figures 2–5) provide a structural overview of the field and contextualise the thematic synthesis.

Data analysis and synthesis

Findings were synthesised using a narrative thematic approach (Thomas & Harden, 2008), combining qualitative, quantitative, and mixed-methods evidence without statistically pooling effects. Study findings were first coded inductively; codes were grouped into descriptive themes and then aligned with RQ1–RQ3. Reflexive thematic-analysis logic (Braun & Clarke, 2006) was used to support constant comparison across studies. Quantitative findings were retained in their original direction and significance status rather than converted into pooled effect sizes because the studies used heterogeneous designs, outcome constructs, instruments, comparison conditions, and reporting conventions. Accordingly, no pooled effect size, confidence interval, or heterogeneity statistic is reported. Where individual studies supplied numerical outcomes, the manuscript reports those study-level results narratively and avoids cross-study numerical comparisons that cannot be justified from the available data.

Reporting and documentation

The review was reported in accordance with the PRISMA 2020 checklist (Page et al., 2021). The identification, screening, eligibility, and inclusion counts reported throughout the manuscript are mutually consistent and are visualised in the flow diagram below.

PRISMA 2020 flow diagram

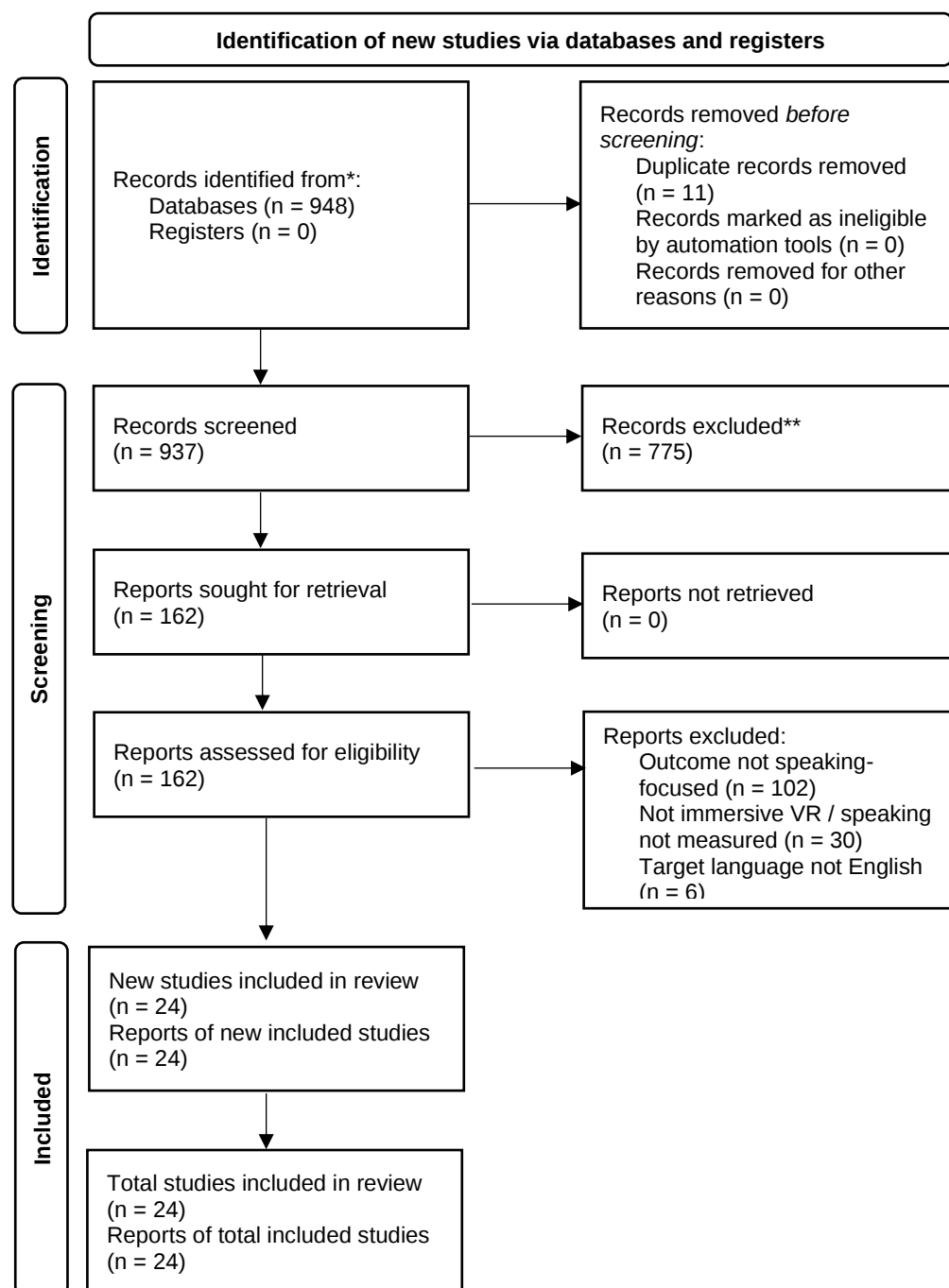


Figure 1. PRISMA 2020 flow diagram of the study selection process

Results and Discussions

Study selection results

The database searches returned 948 records (Scopus = 690, PubMed = 158, ScienceDirect = 100). After 11 duplicates were removed, 937 records were screened by title and abstract, of which 775 were excluded as clearly outside scope. The remaining 162 reports were assessed in full against the eligibility criteria, and 138 were excluded-102 because the primary outcome was not oral/speaking communication, 30 because the intervention was not immersive VR or speaking was not measured, and 6 because the target language was not English. The final corpus comprised 24 empirical studies (Figure 1). These counts are consistent across the abstract, methods, and results.

Descriptive characteristics

The 24 studies were published between 2021 and 2025, with a marked upward trend that peaked in 2024–2025, mirroring the wider diffusion of consumer VR and metaverse platforms (Figure 2). Reported study contexts were concentrated in Asia-South Korea, China, Taiwan, Indonesia, Malaysia, Thailand, and neighbouring systems-with a minority set in the Middle East, Europe, and North America; notably, several records did not specify a geographic context in the available metadata (Figure 3). Mixed-methods designs dominated ($n = 11$), followed by quasi-experimental ($n = 5$), qualitative ($n = 4$), experimental ($n = 3$), and quantitative descriptive ($n = 1$) studies (Figure 4). The full characteristics of each study are presented in Table 4, and a design-, platform-, and outcome-oriented classification with the principal findings is given in Table 5.

Table 4. Characteristics of the included studies

Author(s), Year	Title	Context	Design	Sample
Jehma (2021)	Second Life: A three-dimensional virtual world for developing Thai EFL learners' English communication skills	Thailand	Quasi-experimental	40 undergraduates
Alsaffar (2021)	Virtual reality software as preparation tools for oral presentations: Perceptions from the classroom	United States	Qualitative case study	5 students + 1 instructor
Wu & Hung (2022)	The effects of virtual reality infused instruction on elementary school students' English-speaking performance, willingness to communicate, and learning autonomy	Not reported	Quasi-experimental	56 sixth-graders (28/28)
Ebadi & Ebadijalal (2022)	The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency	Iran	Mixed-methods	20 upper-intermediate learners
Lee & Ko (2023)	Effects of self-regulation, goal orientation, and anxiety on EFL speaking in metaverse and face-to-face contexts	South Korea	Quasi-experimental	253 middle-school students
Raman et al. (2023)	Enhancing English verbal communication skills through virtual reality: A study on engagement, motivation, and autonomy among English as a second language learners	Not reported	Mixed-methods	166 senior high-school students
Hoang et al. (2023)	Extending a learning ecology with virtual reality mobile technology: Oral proficiency outcomes and students' perceptions	Not reported	Mixed-methods	23 undergraduates
Chen et al. (2023)	Facilitating EFL learners' active behaviors in speaking: A progressive question prompt-based peer-tutoring approach with VR contexts	Not reported	Quasi-experimental	Two intact classes
Rababah (2023)	Stepping into language mastery: Virtual reality simulations as catalysts for EFL pronunciation enhancement	Jordan	Experimental	Experimental/control groups
Fan et al. (2024)	A VR-assisted language learning approach to improve EFL learners' willingness to communicate	China	Experimental	70 college students
Yudintseva (2024)	An exploration of low- and high-immersive virtual reality modalities for willingness to communicate in English as a second language	Not reported	Mixed-methods	20 ESL graduate students
Lai (2024)	Enhancing multimodal output in CLIL education: The impact of VR games on fourth-grade students' English poster designs and presentations in Taiwan	Taiwan	Quasi-experimental	81 fourth-graders (40/41)
Ha (2024)	Exploring metaverse-blended learning in an English presentation class: Student perceptions	Not reported	Mixed-methods	24 EFL students
Raman et al. (2024)	Exploring the impact of VR integration on ESL learners' English verbal communication skills: A case study in a Malaysian high school	Malaysia	Qualitative case study	15 Form-6 students
Ding (2024)	The impact of high-immersion virtual reality on EFL learners' foreign language speaking anxiety: A mixed-method approach	China	Mixed-methods	140 learners (4 groups)

Author(s), Year	Title	Context	Design	Sample
Akay & Kessler (2024)	The impact of virtual reality (VR) pedagogy on L2 English learners' oral communication and pragmatic competence	Türkiye	Mixed-methods	60 university students (30/30)
Hung et al. (2024)	Virtual reality is not always a cure-all: Evidences from a quasi-experiment of EFL business speaking courses	Not reported	Mixed-methods	40 university students
Hoter et al. (2025)	Enhancing spoken English skills through an immersive virtual world: A pilot study focusing on motivation, confidence and teacher presence	Not reported	Mixed-methods	41 (+116 obs., 7 interviews)
Nuemaihom et al. (2025)	Examining EFL students' perceptions and experiences with AI-driven metaverse environments for developing communication skills	Indonesia	Quantitative (descriptive)	51 undergraduates
Ratnah et al. (2025)	Fostering ESP speaking and reducing anxiety through interactive, supportive, and VR-based methods in a public tourism polytechnic	Indonesia	Qualitative	20 polytechnic students
Li et al. (2025)	Pedagogical agent positioning in external videos improves English academic presentation proficiency in desktop virtual reality settings	Not reported	Experimental	64 students
Fathi et al. (2025)	Promoting intercultural communicative competence and willingness to communicate in EFL learners through virtual reality: A mixed-methods examination	Not reported	Mixed-methods	50 learners (25/25) + 9
Wang et al. (2025)	Prompts matter in virtual reality environments? A progressive prompt approach to enhance English-speaking performance in students from low socioeconomic status	Not reported	Mixed-methods	51 secondary learners (33/28)
Hwang & Lee (2025)	"Yes, I am more confident with my avatars": Integrating EFL students' speaking practice into metaverse	South Korea	Qualitative (practice report)	Undergraduate drama course

Table 5. Classification of included studies by VR technology, level, outcome focus, and key findings

Author(s), Year	VR platform	Level	Outcome focus	Key findings
Jehma (2021)	Second Life (virtual world)	Higher education	English communication skills	TOEIC pre-post scores improved after out-of-class practice in the virtual world; gains were independent of gender, major, or hours of computer use, indicating broadly distributed benefits.
Alsaffar (2021)	VirtualSpeech (VR mobile app)	Higher education	Oral presentation, confidence, public-speaking anxiety	All participants recommended VirtualSpeech as a rehearsal tool for oral presentations; most reported higher confidence when speaking English after use.
Wu & Hung (2022)	CoSpaces + VR headset	Primary	Speaking performance, WTC, autonomy	VR significantly improved grammar and lexical use in speaking, but produced no significant advantage for pronunciation, fluency and intonation, willingness to communicate, or learning autonomy.
Ebadi & Ebadijalal (2022)	Google Expeditions	Higher education	Oral proficiency, WTC	The VR group significantly outperformed controls on oral performance and willingness to communicate; interviews attributed gains to enhanced motivation, confidence, and cultural awareness.
Lee & Ko (2023)	Metaverse	Secondary	Speaking performance, anxiety, self-regulation	Self-regulation and mastery-goal orientation positively predicted speaking performance while speaking anxiety depressed it, with comparable effects in metaverse and face-to-face settings, suggesting equivalent learning environments.

Author(s), Year	VR platform	Level	Outcome focus	Key findings
Raman et al. (2023)	VR-integrated lessons	Secondary	Engagement, motivation, autonomy	Learners reported high engagement (M = 4.55), motivation (M = 4.36) and autonomy (M = 3.90), with strong inter-correlations; qualitative data highlighted authentic experience, retention, and improved communication.
Hoang et al. (2023)	VR mobile technology	Higher education	Oral proficiency	Significant gains emerged in overall oral proficiency-most clearly in fluency and coherence, lexical resource, and pronunciation-with interviews detailing both affordances and practical constraints of mobile VR.
Chen et al. (2023)	Progressive question-prompt VR peer-tutoring	Higher education	Speaking, self-efficacy	The progressive question-prompt VR peer-tutoring group achieved significantly greater speaking and self-efficacy gains than a conventional VR peer-tutoring group, and displayed more interaction and improvement attempts.
Rababah (2023)	VR pronunciation simulations	Higher education	Pronunciation	A VR simulation-based intervention was contrasted with conventional pronunciation teaching; the study advocates VR simulations for EFL pronunciation training and reports greater learner engagement and satisfaction in the VR condition.
Fan et al. (2024)	Self-developed VR platform	Higher education	Willingness to communicate	VR use exerted a significant positive effect on WTC; learners reported that interacting with virtual objects and characters eased anxiety and motivated them to communicate in the target language.
Yudintseva (2024)	Engage XR (low/high immersion)	Higher education	WTC, speaking anxiety, cognitive load	No significant WTC difference emerged between low- and high-immersion modalities; the high-immersion condition produced greater cognitive load and enjoyment, and speaking anxiety fell while self-confidence rose after both VR tasks.
Lai (2024)	VR games (CLIL)	Primary	Multimodal output (writing, oral presentation)	Grounded in self-regulated learning and a 4Cs assessment framework, the study compared VR-based and PowerPoint-based games for fourth-grade CLIL learners' English poster designs (writing) and oral presentations (speaking).
Ha (2024)	Gather Town (metaverse)	Higher education	Presentation skills, anxiety, engagement	Metaverse-blended instruction produced strong satisfaction and heightened engagement and enjoyment; speaking anxiety, metaverse interest, and English proficiency significantly predicted presentation anxiety in the metaverse context.
Raman et al. (2024)	ImmerseMe	Secondary	Verbal communication (fluency, accuracy)	VR-integrated lessons with ImmerseMe supported ESL learners' English verbal communication-particularly fluency and accuracy-and evidenced improvements in communicative competence during speaking assessments.
Ding (2024)	High-immersion VR	Higher education	Foreign-language speaking anxiety	Although most learners perceived high-immersion VR as useful for alleviating speaking anxiety, statistical analysis showed no significant effect on real-life anxiety, and learners found it hard to transfer reduced in-VR anxiety to authentic situations.
Akay & Kessler (2024)	ImmerseMe	Higher education	Oral communication, pragmatic competence	The VR group achieved larger average gains in oral pragmatic ability, but the between-group difference was not statistically significant; learners enjoyed the VR tasks

Author(s), Year	VR platform	Level	Outcome focus	Key findings
Hung et al. (2024)	VR (business speaking)	Higher education	Oral proficiency, speaking anxiety, emotions	while reporting technical and task-related challenges. ANCOVA revealed no significant group advantage for oral proficiency, speaking anxiety, or emotions, yet interviews indicated that the private VR environment lowered anxiety and functioned as a safe space to rehearse before real-life performance. Participants reported significant improvements in speaking (fluency, pronunciation, vocabulary, grammar) and listening; avatar-mediated immersion mitigated anxiety and boosted confidence, and teacher presence was highly valued for engagement.
Hoter et al. (2025)	OpenSim (Englehaven Island)	Higher education	Speaking, listening, confidence, motivation	Interaction with generative-AI virtual robots was perceived to improve English communication (vocabulary, intonation, gesture, fluency, and volume) and to strengthen learner autonomy, critical thinking, and confidence.
Nuemaihom et al. (2025)	AI-driven metaverse (GenAI robots)	Higher education	Oral communication skills	Confidence emerged as a proximal antidote to speaking anxiety when practice was frequent and supported; interactive formats and graded exposure lowered anxiety, with short, repeated off-the-shelf VR job simulations integrated as one reduction strategy.
Ratnah et al. (2025)	Off-the-shelf VR job simulation	Higher education	ESP speaking anxiety, confidence	Learners viewing an out-of-video pedagogical agent within desktop-VR lectures achieved better academic-presentation performance, allocated less attention to on-screen content, and showed distinct behavioural patterns relative to an in-video agent.
Li et al. (2025)	Desktop VR + pedagogical agent	Higher education	English academic presentation proficiency	VR-enhanced instruction produced significant gains in intercultural communicative competence and willingness to communicate; qualitative benefits included confidence, positive emotion, and cultural awareness, alongside technical and feedback-related challenges.
Fathi et al. (2025)	ImmerseMe	Higher education	Intercultural communicative competence, WTC	The progressive-prompt VR approach positively affected English-speaking performance and learning conceptions for low-SES learners, with qualitative evidence of increased motivation, enjoyment, and a sense of achievement, supporting educational equity.
Wang et al. (2025)	Progressive-prompt VR (PP-VR)	Secondary	English-speaking performance, learning conceptions	Avatar-mediated metaverse practice significantly improved students' confidence, motivation, and English-speaking skills by reducing the anxiety associated with physical presence, though connectivity and expression constraints were noted.
Hwang & Lee (2025)	Metaverse avatars	Higher education	Speaking skills, confidence, motivation, anxiety	

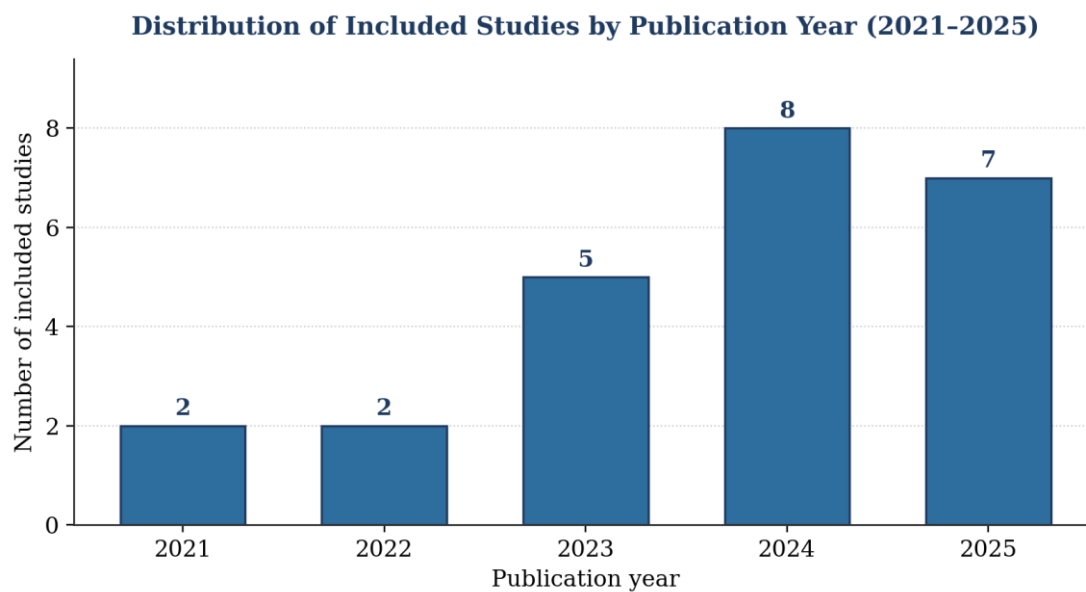


Figure 2. Distribution of included studies by publication year

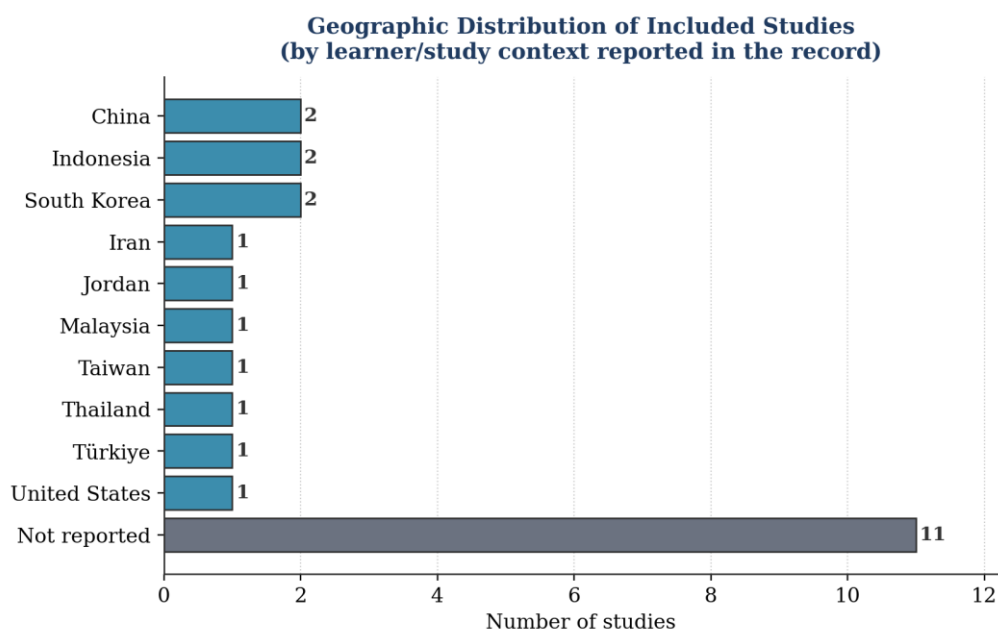


Figure 3. Geographic distribution of included studies by reported learner/study context

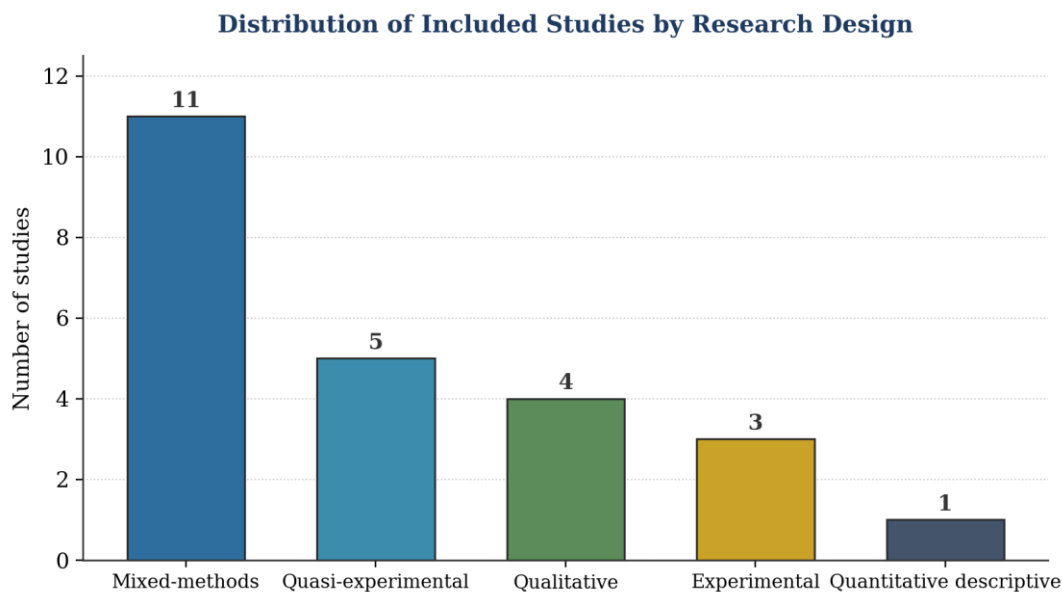


Figure 4. Distribution of included studies by research design

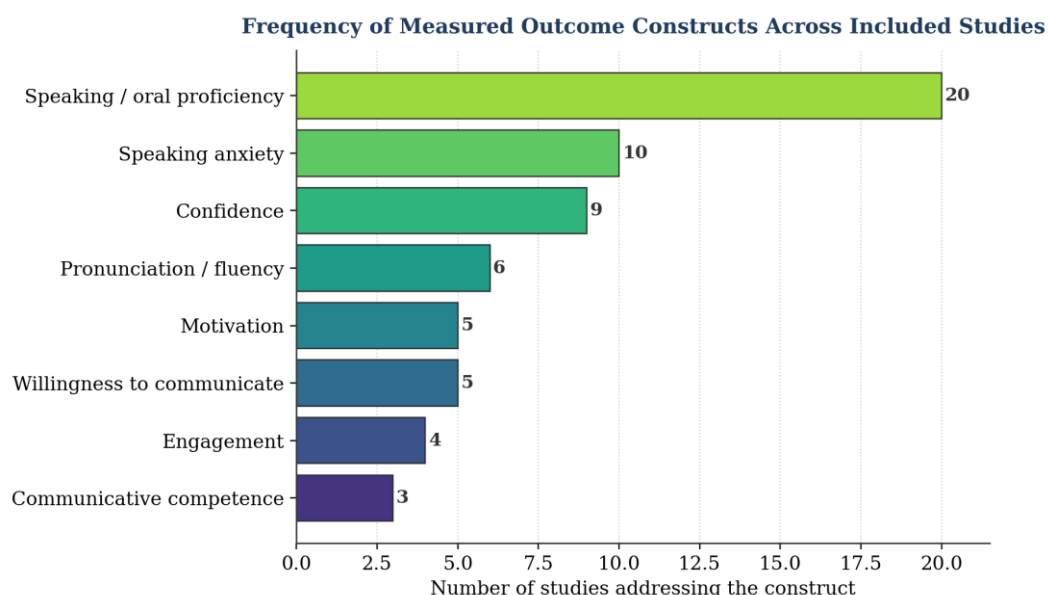


Figure 5. Frequency of measured outcome constructs across included studies

Thematic synthesis

RQ1: Effects of immersive VR on speaking and oral communication

Across the corpus, immersive VR was associated with gains in oral communication, but the strength and nature of those gains varied sharply by outcome type. Several controlled studies reported clear benefits for speaking performance and oral proficiency. Ebadi & Ebadijalal (2022) found that an experimental group using Google Expeditions significantly outperformed a control group on both oral performance and willingness to communicate, while Chen et al. (2023) showed that a progressive question-prompt VR peer-tutoring design produced significantly greater speaking and self-efficacy gains than conventional VR peer-tutoring. Similarly, Hoang et al. (2023) recorded significant improvements in overall oral proficiency—most visibly in fluency, coherence, and lexical resource—and Wang et al. (2025) demonstrated positive effects on English-speaking performance for low-socioeconomic-status learners. Foundational virtual-world studies such as Jehma (2021) likewise reported improved communication skills after sustained practice.

A contrasting cluster of methodologically robust studies reported non-significant effects on measured proficiency and anxiety, tempering the optimistic narrative. Ding (2024) found that high-immersion VR did

not significantly reduce learners' real-life foreign-language speaking anxiety, even though most participants perceived it as helpful, and observed that reduced in-VR anxiety was difficult to transfer to authentic situations. Akay & Kessler (2024) reported larger average gains in oral pragmatic competence for the VR group that were nonetheless statistically non-significant, and Hung et al. (2024)- aptly titled to signal that VR is "not always a cure-all"- found no significant advantage for oral proficiency, anxiety, or emotions despite qualitative evidence of a calmer rehearsal space. Wu & Hung (2022) similarly found that VR improved grammatical and lexical use in speaking but not pronunciation, fluency, willingness to communicate, or autonomy. The answer to RQ1 is therefore conditional: immersive VR frequently improves speaking-related behaviour and affect, yet its impact on rigorously measured oral proficiency is positive but inconsistent.

RQ2: Constructs examined and patterns of effect

The measured constructs were dominated by speaking/oral proficiency (20 studies), followed by speaking anxiety (10), confidence (9), pronunciation and fluency (6), motivation and willingness to communicate (5 each), engagement (4), and communicative competence (3), as visualised in Figure 5. A consistent pattern emerged whereby affective and dispositional constructs responded more reliably than cognitive-linguistic ones. Confidence rose across diverse platforms: Alsaffar (2021) reported higher speaking confidence after rehearsing with a VR presentation app, Hwang & Lee (2025) found that avatar-mediated metaverse practice significantly improved confidence and reduced the anxiety of speaking before peers, and Hoter et al. (2025) documented boosted confidence and motivation in an immersive virtual world.

Willingness to communicate showed generally positive but not universal effects. Fan et al. (2024) reported a significant increase in willingness to communicate attributed to eased anxiety and enjoyable interaction, and Fathi et al. (2025) found significant gains in both intercultural communicative competence and willingness to communicate. For pronunciation specifically, Rababah (2023) advocated VR simulations as a means of improving articulation over conventional instruction. By contrast, Yuditseva (2024) detected no significant difference in willingness to communicate between low- and high-immersion modalities, although speaking anxiety fell and self-confidence rose after both VR tasks. For anxiety specifically, the evidence was split between perceived reduction and non-significant measured change, echoing the RQ1 pattern. In sum, immersive VR appears to act most dependably on the affective gateway to speaking-confidence, enjoyment, and willingness-while its measurable effect on proficiency and anxiety is more variable.

RQ3: Populations, technologies, and effective designs

The included learners spanned primary (Wu & Hung (2022); Lai (2024)), secondary (Lee & Ko (2023); Raman et al. (2023); Raman et al. (2024); Wang et al. (2025)), and, predominantly, higher-education settings. Technologically, the corpus documents a clear evolution from early desktop virtual worlds such as Second Life (Jehma (2021)) toward high-immersion head-mounted VR (Ding (2024)), 360° and mobile VR (Hoang et al. (2023)), off-the-shelf platforms such as ImmerseMe and Google Expeditions (Ebadi & Ebadijalal (2022); Raman et al. (2024); Akay & Kessler (2024); Fathi et al. (2025)), and, increasingly, social VR and metaverse spaces (Lee & Ko (2023); Ha (2024); Hoter et al. (2025); Hwang & Lee (2025)), including AI-driven metaverse environments (Nuemaihom et al. (2025)).

Pedagogical design, rather than immersion per se, appeared decisive. Studies that embedded structured scaffolding produced the most convincing gains: progressive question prompts and peer tutoring (Chen et al. (2023); Wang et al. (2025)), goal-oriented and self-regulated tasks (Lee & Ko (2023)), and deliberate teacher presence within the virtual world (Hoter et al. (2025)). Attention to instructional micro-design also mattered, as Li et al. (2025) showed that the positioning of a pedagogical agent in desktop-VR lectures significantly affected academic-presentation performance. Conversely, studies that relied on immersion alone, without task structure, tended to report weaker or non-significant proficiency effects (Wu & Hung (2022); Hung et al. (2024)). The most effective configurations therefore combine immersive presence with task-based, scaffolded, and socially supported speaking practice.

Comparative and critical analysis

Methodologically, the field leans heavily on mixed-methods ($n = 11$) and quasi-experimental studies ($n = 5$), followed by qualitative ($n = 4$), experimental ($n = 3$), and quantitative descriptive ($n = 1$) designs. Four studies received medium methodological-confidence classifications and 20 received high classifications under the manuscript's MMAT interpretation. The evidence base is therefore heterogeneous and cannot be treated as a single effect estimate. Perceptual and self-report measures are pervasive, and several positive conclusions rest on questionnaire and interview data rather than independently rated speaking performance. This creates a recurring tension between perceived benefits and measured effects, visible when enthusiastic perceptions in Raman et al. (2023) and Nuemaihom et al. (2025) are considered alongside non-significant statistical

outcomes in Ding (2024), Akay and Kessler (2024), and Hung et al. (2024). The dominance of higher-education samples and incomplete contextual reporting further constrains generalisability. The descriptive pattern therefore supports cautious claims about likely pedagogical utility rather than strong causal claims about the superiority of immersive VR across speaking outcomes.

Discussion

Interpreted as a whole, the evidence suggests that immersive VR reshapes the affective and motivational conditions of speaking more dependably than it accelerates linguistic proficiency. The recurrent findings of heightened confidence, willingness to communicate, and engagement, alongside reduced in-task anxiety, are coherent with situated-learning and cognitive-affective accounts of immersive learning: presence and authenticity lower the perceived social risk of speaking, freeing learners to produce more language. This interpretation aligns with broader reviews describing immersive technologies as powerful engagement and affect regulators (Christou et al., 2025; Tegoan et al., 2025).

The apparent advantage for affective outcomes should not be interpreted as an effect-size claim. The present review did not pool outcomes, and comparable confidence intervals were not consistently available across the 24 studies. Therefore, terms such as 'more consistent' refer to the direction and recurrence of reported findings, not to a statistically demonstrated difference between domains. Measurement validity also constrains interpretation: speaking proficiency was operationalised through different tests and rubrics, while anxiety and willingness to communicate were measured with non-identical instruments and task contexts. Differences in sample size, intervention duration, comparison conditions, and statistical power can therefore generate apparent inconsistencies that cannot be resolved through narrative comparison alone.

Theoretically, the results extend affective models of second-language speaking by locating the primary mechanism of immersive VR in anxiety reduction and willingness to communicate rather than in direct linguistic instruction. They also qualify strong claims about immersion: consistent with meta-analytic cautions that technology effects are moderated by design and outcome type (Yi et al., 2025; Xue, 2025), immersion appears to be a facilitating condition whose linguistic payoff depends on pedagogy. The finding that reduced in-VR anxiety may not transfer to real-world speaking (as reported by Ding (2024)) is theoretically important, pointing to a presence–transfer gap that current frameworks under-specify.

Practically, the synthesis favours task-based, scaffolded, and socially mediated VR designs over immersion for its own sake. Educators are advised to embed progressive prompts, peer interaction, and teacher presence; to manage cognitive load, especially in high-immersion conditions; and to build explicit bridges from virtual rehearsal to authentic performance. For under-resourced settings, the evidence that VR can support low-socioeconomic-status learners (Wang et al. (2025)) and workplace-oriented speaking (Ratnah et al. (2025)) is encouraging, though it must be weighed against persistent technical and access barriers documented in workplace-communication reviews (Younus et al., 2025).

Situated against prior reviews, this speaking-focused synthesis both confirms and refines the literature. Multi-skill reviews of VR and augmented reality in English education report broadly positive but heterogeneous effects (Christou et al., 2025; Esfandiari & Mazharpour, 2025; Ma et al., 2025), while reviews of metaverse language education similarly emphasise pedagogical design, methodological diversity, and implementation constraints (Fan & Chiang, 2023; Jiao et al., 2024). A 2024 review of XR in English classrooms identified speaking among the main target areas and reported benefits for motivation, engagement, satisfaction, and language learning, while also noting implementation challenges (Computers & Education: X Reality, 2024). Reviews focused on design recommendations for immersive VR English-learning applications likewise highlight the need to align immersion with instructional goals rather than treating immersion as an outcome in itself. More recent reviews continue to report heterogeneous evidence, including a 2025 systematic review of higher-education VR, a 2025 review of 360° video-based VR in English education, and a 2025 systematic review of K–12 randomised trials. Collectively, these syntheses reinforce the present review's interpretation: VR is promising, but outcomes depend on task design, scaffolding, context, and measurement rather than immersion alone. The present review adds a narrower speaking-specific lens and identifies the recurring divergence between affective gains and transfer to authentic speaking proficiency.

Limitations

Future research should (a) use adequately powered controlled and preferably longitudinal designs with independently rated speaking performance; (b) pair virtual rehearsal with authentic-context assessment to test transfer explicitly; (c) compare immersion levels, platform types, and speaking sub-skills using common outcome measures; (d) include under-represented learners and settings while reporting educational,

geographical, and technological context transparently; and (e) use at least two independent reviewers for screening and extraction, with inter-rater agreement statistics reported before final synthesis.

In direct response to the research questions, RQ1 indicates that immersive VR is generally associated with improved oral-communication-related outcomes, but measured proficiency effects are inconsistent. RQ2 shows that speaking/oral proficiency was the most frequently examined construct (20 studies), followed by speaking anxiety (10), confidence (9), pronunciation/fluency (6), motivation (5), willingness to communicate (5), engagement (4), and communicative competence (3). The descriptive pattern is most consistently positive for confidence, willingness to communicate, and engagement, while proficiency and anxiety remain mixed. RQ3 indicates that the corpus is concentrated in higher education and Asian contexts and increasingly includes social VR, metaverse, and AI-enhanced environments; the strongest recurring pedagogical pattern is the combination of immersive practice with task-based prompts, peer interaction, self-regulation, and teacher support.

Accordingly, four future directions are proposed. Researchers should (a) design adequately powered, controlled, and preferably longitudinal studies that measure independently rated speaking performance; (b) test transfer explicitly by pairing virtual rehearsal with authentic-context assessment; (c) run systematic comparisons across immersion levels, platforms, and speaking sub-skills to identify optimal configurations; and (d) extend the evidence base to under-represented learners and settings while reporting context transparently.

In direct response to the research questions: for RQ1, immersive VR generally supports oral communication but yields inconsistent effects on rigorously measured proficiency; for RQ2, affective constructs-confidence, willingness to communicate, and engagement-respond most dependably, whereas proficiency and anxiety show mixed results; and for RQ3, higher-education learners in Asian contexts predominate, social VR and metaverse platforms are ascendant, and task-based, scaffolded, socially supported designs are the most effective.

Conclusions

This systematic review indicates that immersive VR has meaningful potential for supporting L2 English oral communication, but the evidence does not justify treating immersion itself as a sufficient condition for improved speaking proficiency. Across 24 empirical studies, the clearest and most recurring benefits concerned confidence, willingness to communicate, engagement, and perceived safety for speaking. Effects on objectively or structurally measured oral proficiency and speaking anxiety were more variable, and evidence from high-immersion conditions shows that reduced anxiety within VR does not necessarily transfer to authentic speaking situations. The practical implication is that VR should be implemented as a pedagogically structured environment-supported by purposeful speaking tasks, progressive prompts, peer interaction, and teacher presence-rather than as a technology-driven intervention without instructional scaffolding.

Acknowledgments

The Authors would like to thank Faculty of Sport Sciences, Universitas Negeri Padang for providing excellent facilities, as well as for being very supportive of this research from the beginning. Authors also thank the team members who had generously shared their brilliant ideas, engaging discussion, and academic supports during the study.

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://doi.org/10.64152/10125/73568>
- Alsaffar, M. J. (2021). Virtual reality software as preparation tools for oral presentations: Perceptions from the classroom. **Theory and Practice in Language Studies*, 11*(10), 1146–1160. <https://doi.org/10.17507/tpls.1110.02>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>

- Celik, F., & Ersanlı, C.Y. (2025). Developing and Integrating Metaverse Into EFL Listening Skills to Increase Learners' Self-Determination Based on Social Constructivist Theory. *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.70151>
- 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(9). <https://doi.org/10.1080/10494820.2021.1878232>
- Chen, C., & Yuan, Y. (2025). Effectiveness of Virtual Reality on Chinese as a second language vocabulary learning: perceptions from international students. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2023.2192770>
- Christou, E., Parmaxi, A., & Christoforou, M. (2025). Implementation and application of extended reality in foreign language education for specific purposes: A systematic literature review. **Universal Access in the Information Society*, 24*, 2061–2076. <https://doi.org/10.1007/s10209-025-01191-w>
- Christou, E., Vassiliou, P., & Parmaxi, A. (2025). Augmented reality in language learning: a systematic literature review of the state-of-the-art and task design considerations. *Innovation in Language Learning and Teaching*. <https://doi.org/10.1080/17501229.2025.2504706>
- 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). <https://doi.org/10.1080/09588221.2020.1854311>
- Esen, S. (2025). Digital Tools in Second Language Learning in Higher Education: A Systematic Review of Recent Research. *Janus.net*. <https://doi.org/10.26619/1647-7251.dt0325.4>
- Esfandiari, S., & Mazharpour, D. (2025). A Systematic Review of Augmented Reality and Virtual Reality Integration in English as a Foreign Language Education. *Applied Research on English Language*. <https://doi.org/10.22108/are.2025.144750.2479>
- Fan, H., Jing, B., Yi, Y., & Luan, L. (2024). A VR-assisted language learning approach to improve EFL learners' willingness to communicate. *Global Chinese Conference on Computers in Education Proceedings*.
- Fathi, J., Zou, D., & Zhaleh, K. (2025). Promoting intercultural communicative competence and willingness to communicate in EFL learners through virtual reality: A mixed-methods examination. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2025.2496827>
- Fu, G., & Su, F. (2025). A systematic review of technology-assisted preschoolers' language learning: technologies, theoretical foundations, and implications. *Early Years*. <https://doi.org/10.1080/09575146.2025.2587844>
- Ha, M.-J. (2024). Exploring metaverse-blended learning in an English presentation class: Student perceptions. *CALL-EJ*, 25(2).
- 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(2). <https://doi.org/10.1080/17501229.2022.2070626>
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). Mixed Methods Appraisal Tool (MMAT), version 2018. Registration of Copyright (#1148552), Canadian Intellectual Property Office.
- Hoter, E., Yazbak Abu Ahmad, M., & Azulay, H. (2025). Enhancing spoken English skills through an immersive virtual world: A pilot study focusing on motivation, confidence and teacher presence. *Language Learning Journal*. <https://doi.org/10.1080/09571736.2025.2552375>
- Hung, S.-T.A., Chen, W.-J., & Chien, S.-Y. (2024). Virtual reality is not always a cure-all: Evidences from a quasi-experiment of EFL business speaking courses. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2023.2200811>
- Hwang, Y., & Lee, J.H. (2025). "Yes, I am more confident with my avatars": Integrating EFL students' speaking practice into metaverse. *RELC Journal*. <https://doi.org/10.1177/00336882241251952>
- Jehma, H. (2021). Second Life: A three-dimensional virtual world for developing Thai EFL learners' English communication skills. *Turkish Online Journal of Distance Education*, 22(4). <https://doi.org/10.17718/tojde.770909>
- Jiao, Y., DeWitt, D., & Abdul Razak, R. B. (2024). Exploring the potential of metaverse-based platforms for enhancing English as a foreign language (EFL) learning: A comprehensive systematic review.

- *International Journal of Information and Education Technology, 14*(2), 231–239. <https://doi.org/10.18178/ijiet.2024.14.2.2044>
- Lai, C.-J. (2024). Enhancing multimodal output in CLIL education: The impact of VR games on fourth-grade students' English poster designs and presentations in Taiwan. *Humanities and Social Sciences Communications*, 11. <https://doi.org/10.1057/s41599-024-03999-y>
- Lee, J., & Ko, Y. (2023). Effects of self-regulation, goal orientation, and anxiety on EFL speaking in metaverse and face-to-face contexts. *English Teaching (South Korea)*, 78(4). <https://doi.org/10.15858/engtea.78.4.202312.219>
- Li, X., Lan, Y.-J., Pi, Z., Qi, G.Y., Grant, S., & Sun, J. (2025). Pedagogical agent positioning in external videos improves English academic presentation proficiency in desktop virtual reality settings. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13531>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Medicine*, 6(7), e1000100. <https://doi.org/10.1371/journal.pmed.1000100>
- Luan, L., Hwang, G.-J., Yi, Y., Lu, Z., & Jing, B. (2025). The effects of a self-developed virtual reality environment on college EFL learners' vocabulary learning. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2024.2344056>
- Ma, C., Haire, G., Smith, S., & Newbutt, N. (2025). Enhancing English as Foreign Language Education with Virtual Reality and Augmented Reality in K–12 Classrooms: A Systematic Review from 2019 to 2023. *Technology in Language Teaching and Learning*. <https://doi.org/10.29140/tltl.v7n3.102535>
- Nuemaihom, A., Muthmainnah, & Al Yakin, A. (2025). Examining EFL students' perceptions and experiences with AI-driven metaverse environments for developing communication skills. *Forum for Linguistic Studies*, 7(5). <https://doi.org/10.30564/fls.v7i5.9453>
- Nuryana, Z., Lu, S., Ni, X., Xu, W., & Wibawa, A. P. (2025). The impact of augmented and virtual reality on sustainable education: A bibliometric analysis. *Next Research*. <https://doi.org/10.1016/j.nexres.2025.100999>
- 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>
- Rababah, L.M. (2023). Stepping into language mastery: Virtual reality simulations as catalysts for EFL pronunciation enhancement. *Journal of Applied Learning and Teaching*, 6(S1). <https://doi.org/10.37074/jalt.2023.6.s1.10>
- Raman, K., Hashim, H., & Ismail, H. H. (2023). Enhancing English verbal communication skills through virtual reality: A study on engagement, motivation, and autonomy among English as a second language learners. **International Journal of Learning, Teaching and Educational Research*, 22*(12), 237–261. <https://doi.org/10.26803/ijlter.22.12.12>
- Raman, K., Hashim, H., & Ismail, H.H. (2024). Exploring the impact of VR integration on ESL learners' English verbal communication skills: A case study in a Malaysian high school. *Arab World English Journal*. <https://doi.org/10.24093/awej/call10.10>
- Ratnah, R., Saputri, F., & Rosmaladewi, R. (2025). Fostering ESP speaking and reducing anxiety through interactive, supportive, and VR-based methods in a public tourism polytechnic. **International Journal of Language Education*, 9*(4), 985. <https://doi.org/10.26858/ijole.v9i4.84021>
- Santoso, S., Surjawan, D.J., & Handoyo, E.D. (2025). The Impact of Immersive Virtual Reality Flashcard on L2 Vocabulary Acquisition on Logographic Language. *International Journal of Information and Education Technology*. <https://doi.org/10.18178/ijiet.2025.15.8.2377>
- Shadiev, R., Yerkenaz, D., Yesmagambetova, A., Sapiyeva, A., Yermekbayeva, K., & Beisembayeva, Z. (2025). A systematic literature review on the use of 360-degree video-based virtual reality in English language education. **Digital Applied Linguistics*, 3*, 102704. <https://doi.org/10.29140/dal.v3.102704>
- Shen, B., Xing, W., Lin, Z., Jiang, M.Y.-C., Zou, D., & Jong, M.S.-Y. (2025). The Effects of Spherical Video-Based Virtual Reality (SVVR) on Chinese University EFL Learners' Writing Complexity. *Asia-Pacific Education Researcher*. <https://doi.org/10.1007/s40299-024-00952-0>
- Shokri, R., & Rahimi, M. (2025). Integrating 360°VR Videos into L2 Listening Instruction: Examination of the Effects of Two Types of VR Learning Experiences on Learning Gains. *International Journal of Information and Education Technology*. <https://doi.org/10.18178/ijiet.2025.15.11.2437>

- 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, L., & Song, X. (2026). The effectiveness of virtual reality for K-12 foreign language learning: a systematic review of recent randomized controlled trials. *Front Psychol.* <https://doi.org/10.3389/fpsyg.2025.1714481>
- Tegoan, N., Grandhi, S., Wibowo, S., & Yang, R. (2025). The present and future directions of metaverse, artificial intelligence, and second life environments in new language education within higher education: A systematic review. *Journal of Applied Learning and Teaching.* <https://doi.org/10.37074/jalt.2025.8.2.19>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, 45. <https://doi.org/10.1186/1471-2288-8-45>
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 14(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>
- Wang, H.-J., Zou, D., Liu, C.-C., Wang, Y., & Li, W. (2025). Prompts matter in virtual reality environments? A progressive prompt approach to enhance English-speaking performance in students from low socioeconomic status. *Innovation in Language Learning and Teaching.* <https://doi.org/10.1080/17501229.2025.2569488>
- Wu, Y.-H., & Hung, S.-T. (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). <https://doi.org/10.1177/07356331211068207>
- Xue, Y. (2025). What makes spherical video virtual reality popular in education? Evidence from a meta-analysis employing Stata. *Education and Information Technologies.* <https://doi.org/10.1007/s10639-024-13106-y>
- Yang, P., & Zhang, W. (2025). Effectiveness of augmented reality on EFL learner's language gains: a meta-analysis. *Innovation in Language Learning and Teaching.* <https://doi.org/10.1080/17501229.2024.2323529>
- Yi, S., Li, W., Zhang, Y., & Shadiev, R. (2025). Exploring the impact of technology on foreign language learning: a multivariate meta-meta-analysis study. *Educational Technology Research and Development.* <https://doi.org/10.1007/s11423-024-10412-7>
- Younus, M., bin Mohamed, A.A., binti Ali, F., & Bin Ab Hamid, M.R. (2025). Technical Oral Presentations (TOP) in EFL Engineering Education: A Systematic Literature Review on Workplace Communication Preparedness. *World Journal of English Language.* <https://doi.org/10.5430/wjel.v16n2p229>
- Yudintseva, A. (2024). An exploration of low- and high-immersive virtual reality modalities for willingness to communicate in English as a second language. *Computers and Education: X Reality*, 4. <https://doi.org/10.1016/j.cexr.2024.100076>
- 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*. <https://doi.org/10.3389/frvir.2025.1522380>
- Zhang, M.M., Hashim, H., & Yunus, M.M. (2025). Evaluating metaverse-based L2 vocabulary learning effectiveness using a proposed metric of vocabulary forgetting percentage. *Frontiers in Education.* <https://doi.org/10.3389/educ.2025.1641638>