

Review

Technology-Mediated Public Speaking Interventions for Educational Training and Anxiety Treatment: A Comprehensive Review

Dragoș-Ion Dogioiu ^{1,*} , Anca Andreea Morar ¹, Alin Dragoș Bogdan Moldoveanu ¹ , Ana Magdalena Anghel ¹  and Alexandru Ion Berceanu ² 

¹ Faculty of Automatic Control and Computer Science, National University of Science and Technology POLITEHNICA Bucharest, 313 Splaiul Independenței, Sector 6, 060042 Bucharest, Romania; anca.morar@cs.pub.ro (A.A.M.); alin.moldoveanu@cs.pub.ro (A.D.B.M.); ana.anghel@upb.ro (A.M.A.)

² Faculty of Film, National University of Theatre and Film “I.L. Caragiale”, 75–77 Matei Voievod Street, Sector 2, 021452 Bucharest, Romania; alexandru.berceanu@unatc.ro

* Correspondence: dragos_ion.dogioiu@stud.acs.upb.ro

Abstract

Public speaking training and public speaking anxiety (PSA) treatment increasingly rely on technology-mediated systems, including web, mobile, desktop, and immersive virtual reality (VR) tools. This comprehensive review identifies and maps information and communication technology (ICT)-based public speaking interventions for educational training and anxiety-oriented treatment, focusing on platform choice, audience simulation, feedback timing, physiological and behavioral sensing, therapeutic framing, practitioner involvement, and evaluation methods. A PRISMA-inspired process documented identification and screening across Scopus, Web of Science, PubMed, IEEE Xplore, and ERIC for English-language journal articles and conference papers published between 1 February 2015 and 1 February 2025. From 6602 records, 82 participant-validated studies were included. No formal risk-of-bias assessment was conducted. VR accounted for 82.93% of interventions, while research prototypes comprised 72.0% of the evidence base, demonstrating the field's strong emphasis on immersive and experimental development. At the same time, 64.6% of interventions provided no automated feedback and 36.6% used no physiological or behavioral sensing, highlighting opportunities for more responsive and data-informed systems. The findings can guide the development of future public-speaking training and treatment systems by highlighting recurring design components, promising implementation patterns, and priorities for stronger comparative validation.

Keywords: public speaking anxiety; virtual reality; review; communication skills training; public speaking anxiety treatment



Academic Editor: Andrea Sanna

Received: 2 June 2026

Revised: 24 July 2026

Accepted: 28 July 2026

Published: 30 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Public speaking competence is a highly relevant capability for graduates as well as a visible marker of professional readiness. Communication competence is important for professional and civic participation [1]. Structured practice and feedback can support the development of oral-presentation competence [2]. These gaps demand structured approaches to help learners acquire delivery skills in front of an audience.

Public-speaking apprehension exists on a continuum, from normative situational nervousness to elevated or impairing public speaking anxiety (PSA). PSA should not

be treated as synonymous with social anxiety disorder (SAD): public-speaking fear may occur as a circumscribed difficulty, within a performance-only presentation of SAD, or as part of a broader clinical pattern involving multiple social situations [3–5]. In this review, glossophobia is used only as a descriptive term for a pronounced fear of public speaking, not as a separate diagnosis.

Public-speaking fear has been reported across community and student samples, although estimates vary substantially across populations and measurement approaches [6–9]. Interventions therefore range from educational skills training for normative or subclinical apprehension to clinical approaches for persistent, impairing fear. Cognitive-behavioral accounts emphasize fear of negative evaluation and self-focused attention [3,10]. Intervention research includes exposure, cognitive-behavioral methods, and skills-based or self-management approaches [11,12].

This review examines information and communication technology (ICT)-enabled solutions for public-speaking training and PSA treatment across web, mobile, desktop, and immersive virtual reality (VR) platforms. These systems can standardize practice scenarios, reduce logistical barriers, and capture behavioral or physiological data. Early virtual audience research showed that simulated audience behavior can elicit PSA [13]. Mobile rehearsal applications subsequently provided immediate or delayed feedback on timing, body movement, and voice level [14]. Virtual reality exposure therapy (VRET) has demonstrated beneficial effects in SAD [15] and across anxiety-related disorders more broadly [16]. PSA-specific trials using 360° video and gamified exposure delivered through head-mounted displays (HMDs) further support the feasibility and potential efficacy of these approaches [17,18].

Previous reviews have addressed psychological and internet-based interventions for social anxiety or fear of public speaking but have paid less attention to the design of technology-mediated public-speaking systems. Platform choice, audience simulation, feedback timing, sensing, and teacher or therapist involvement therefore remain insufficiently mapped across educational and anxiety-treatment contexts.

Unlike prior syntheses that address only parts of this field, our review connects educational public-speaking training with anxiety-oriented treatment. Ebrahimi et al. examined psychological interventions and delivery mode but did not characterize system features such as audience design or feedback mechanics [19]. Esfandiari et al. compared internet-delivered and face-to-face cognitive behavioral therapy (CBT) across anxiety disorders, whereas our review is restricted to participant-validated public speaking interventions [20]. Guo et al. focused on the clinical efficacy of internet-based CBT for social anxiety disorder, while, in contrast, we examine the instructional and engineering choices that structure practice, including platform, audience configuration, feedback timing, biofeedback, and post-session analytics [21].

The objective of this comprehensive review is to identify, classify, and synthesize ICT-based solutions for public speaking training and PSA treatment according to their technological, pedagogical, and therapeutic characteristics. By integrating participant-validated studies from both educational training and anxiety-oriented intervention contexts, the review identifies recurring design patterns across educational training and anxiety-oriented intervention contexts, which may inform future public-speaking training and treatment systems.

2. Field Overview

Technology-mediated public speaking interventions range from simple mobile and desktop tools to immersive VR systems. Across educational, professional, and anxiety-oriented contexts, these systems are used to support speaking practice, improve perfor-

mance, and reduce public-speaking fear or anxiety. Although their immediate aims may differ, they commonly rely on repeatable speaking scenarios, adjustable difficulty, guided practice, and structured feedback.

Early work showed that technology-mediated feedback can augment or replace in-person coaching. The Presentation Trainer of Schneider et al. [22] tracked speakers' voice and posture to deliver real-time tips during practice. Similarly, the mobile app of Lui et al. [14] provided instant feedback on timing, volume, and body movement, allowing learners to practice independently. Web-based platforms and video recordings offered learners additional opportunities to rehearse in a less threatening environment. For instance, having students create video blogs for peer review offered a low-pressure practice environment that significantly reduced speaking anxiety in one study [23]. These innovations familiarized users with technology-mediated training and paved the way for immersive solutions.

Consumer VR made it possible to reproduce the social pressure of speaking before an audience in a controlled setting. When incorporated into structured training or exposure protocols, these simulations were also used to reduce PSA. An at-home 360° video exposure program yielded significant anxiety reductions [24], and a later randomized controlled trial reported reductions in PSA following 360° video VRET [17]. In educational settings, some studies reported improvements in selected presentation outcomes following VR practice. Students who rehearsed speeches in VR showed improved delivery performance without heightened anxiety, as compared to a control group [25].

With VR scenarios in place, developers began tuning virtual audience characteristics to adjust difficulty. Mostajeran et al. [26] examined whether virtual audience size influenced social anxiety and physiological arousal, but the observed effects did not follow a simple linear pattern. Factors like audience demeanor, degree of familiarity, and distance from the speaker have also been varied to simulate easier or more stressful speaking environments. Importantly, most systems keep real-time feedback minimal and defer detailed analytics to the post-speech debrief so as not to distract the speaker during performance.

Some systems incorporate physiological monitoring, although participant-facing biofeedback remains uncommon. Premkumar et al. [27] returned heart rate (HR) and electroencephalography (EEG)-derived information to users during VR exposure, producing steadier physiological change and lower self-reported arousal than exposure alone. Other systems use physiological signals only for retrospective assessment or clinician monitoring.

The reviewed literature also includes VR, web, and mobile approaches with different trade-offs between immersion, accessibility, and opportunities for repeated practice. Looking ahead, a few research prototypes are beginning to test large language model (LLM)-based audience interaction, such as Min et al. [28], which uses an LLM to generate live audience questions in a VR speech trainer. In this review, this feature is treated as an emerging direction rather than an established design pattern. Across the reviewed period, systems increasingly explored audience customization, adaptive responses, sensing, and more accessible delivery formats, extending public speaking training and treatment beyond what was possible with mirror practice and traditional classrooms.

3. Methodology

This comprehensive review was informed by selected reporting elements from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework to transparently document study identification, screening, and selection. However, screening and initial inclusion decisions were conducted by a single reviewer rather than independently duplicated by multiple reviewers. The selected papers include ICT solutions for public-speaking skills training in educational or professional settings, as well as

anxiety-oriented interventions delivered in clinical, educational, or self-guided contexts. A ten-year publication window, from 1 February 2015 to 1 February 2025, was selected to capture recent technological and intervention-design trends while maintaining a manageable review scope. The upper limit corresponds to the date on which the database searches were completed. The earliest included study within this window was published on 24 February 2015, as shown in Figure 1. Searches were run in five databases selected to balance coverage across education, health, psychology, and computing: Scopus, Web of Science, PubMed, IEEE Xplore, and ERIC. Searches were limited to records in English or with an English translation.

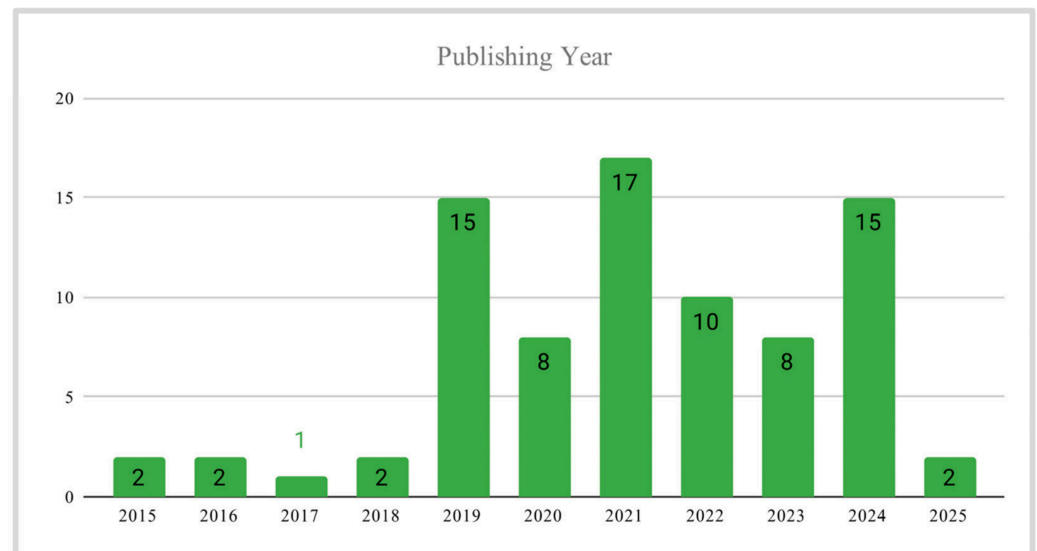


Figure 1. Distribution of included studies by publication year.

The search strategy combined four concept groups addressing: (1) public speaking and related communication activities; (2) anxiety, training, assessment, and improvement; (3) educational settings and populations; and (4) digital and immersive technologies. No geographic filters were applied. The following Boolean expression was developed and tested in Scopus and subsequently used as the baseline for database-specific adaptations:

((("public speaking" OR "speaking in public" OR "oral presentation" OR "presentation skills" OR "communication skills" OR "speech delivery" OR "rhetoric" OR "speech performance")AND ("fear" OR "anxiety" OR "nervousness" OR "glossophobia" OR "speech anxiety" OR "performance anxiety" OR "training" OR "coaching" OR "skills development" OR "evaluation" OR "assessment" OR "improvement" OR "enhancement")) AND ("education" OR "students" OR "academic" OR "school" OR "college" OR "university" OR "curriculum")) AND ("virtual reality" OR "VR" OR "augmented reality" OR "AR" OR "simulation" OR "simulation-based" OR "e-learning" OR "digital" OR "computer-based" OR "ICT-based" OR "technology-assisted" OR "serious games"))

The baseline search was adapted to the field structures and query interfaces of the other databases without changing its vocabulary, block organization, or Boolean relationships. In Scopus, the expression was applied to the title, abstract, and keyword fields. In Web of Science, the Scopus field specification was removed, and the complete expression was entered in the Topic field of Advanced Search, covering titles, abstracts, author keywords, and Keywords Plus. In PubMed, the same terms were retained, but each was assigned the '[Title/Abstract]' field tag, covering titles, abstracts, and author-supplied keywords. In IEEE Xplore, the four concept blocks were entered as separate All Metadata fields and

connected using AND. In ERIC, the complete Boolean expression was entered using the database's default unfielded Collection search. The transformations therefore concerned only database-specific field selection, field-tag syntax, and query-entry format. No terms, concept groups, or Boolean operators were added, removed, or substantively altered.

A paper was deemed eligible if all of the following were met:

The study evaluated an ICT-based solution centered on audience-facing speaking tasks, and targeting public-speaking performance, presentation competence, or speaking-related anxiety;

The solution was evaluated or validated with human participants;

The publication was a journal article or conference paper;

The article was in English (or provided an English translation);

The publication date fell between 1 February 2015 and 1 February 2025;

The full text was retrievable after targeted access attempts through institutional access, publisher or database links, or open-access copies.

For eligibility, public speaking was defined as a spoken performance before a real or simulated audience, with presentation competence or speaking-related anxiety as a central focus. PSA referred to anxiety associated with such tasks, while skills training targeted delivery and presentation competence. Debate, classroom presentations, interviews, stuttering tasks, and broader oral-language activities were included only when they met this definition. Unrelated conversation, language-learning, and communication studies were excluded.

Full texts were sought through institutional access routes available via the e-information platform, publisher and database links associated with the searched records, and supplementary Google Scholar checks for open-access copies. Records were excluded if any of the following applied: no participant validation, insufficient methodological detail to understand the solution, or studies without clear empirical evaluation. The paper selection methodology applied in this study yielded 6602 records. After removing duplicates, 4683 unique records remained. Meta-analyses and narrative reviews were then removed, leaving 3787 records for topical screening. The literature search process is summarized in the study identification and selection flow diagram (Figure 2). An extraction worksheet detailing the extracted data for all included studies is available as Supplementary Material (Supplementary Data Extraction and Coding Workbook).

Broad domain terms were used in Zotero to locate clusters of records likely to concern discipline-specific communication rather than public speaking, including oncology, emergency and critical care, pediatrics, surgery, pharmacy, neurology, infectious diseases, palliative and perinatal care, and veterinary medicine. This was not an automated exclusion procedure: the first author inspected the title of every flagged record and consulted the abstract whenever relevance was uncertain. Records were excluded according to the eligibility criteria rather than solely because they contained a domain term. Because the terms overlapped and served only as navigation aids, per-keyword exclusion counts were not generated. The records excluded at this stage were subsequently rechecked against the revised operational definition of public speaking, and potentially eligible records were returned to full-text assessment.

Titles and abstracts were then assessed against the inclusion criteria, with conservative retention when scope was ambiguous. This produced 97 articles for full-text eligibility assessment. Fifteen records were excluded at full-text review because they did not meet the core eligibility criteria, most commonly due to insufficient methodological detail, resulting in the final 82 studies. The data extraction template mirrored the variables in the paper dataset: bibliographic metadata (Year; Paper Author(s) and Reference Format; Title; DOI), sample size (Total Number of Participants), delivery format and maturity (Platform;

Solution Maturity), therapeutic framing (Therapy Mentioned), practitioner involvement (Therapist Participation), sensing modalities (Sensing Modalities Tracked), embodiment and scenario design (Simulated Virtual Body; Audience Simulation), guidance and feedback mechanisms beyond audience behavior (Non-Audience Related Solution Feedback and Timing—when is the feedback offered), artificial intelligence (AI) by means of LLM integration, and outcome measurement (Evaluation Instruments Used). Intervention duration and session number, follow-up, transfer to real-life speaking, dropout and adherence, adverse or negative effects, accessibility, tolerability, and assessor blinding were not predefined as separate extraction variables. These characteristics are therefore discussed descriptively where explicitly reported rather than quantified across the full corpus.

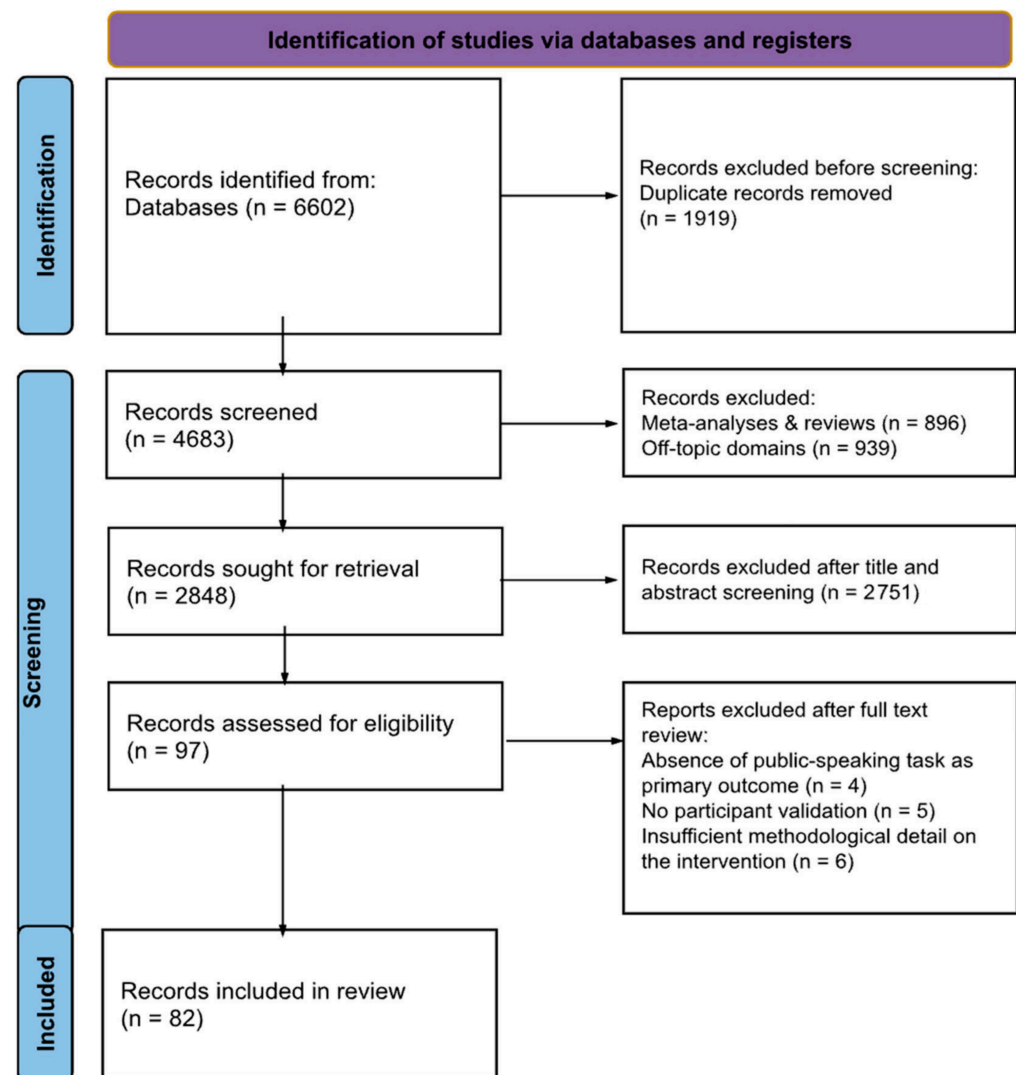


Figure 2. Study identification, screening, eligibility assessment, and inclusion flow diagram.

Screening, data extraction, and coding were conducted by the first author. Although the review did not employ formal independent duplicate screening or inter-rater agreement statistics, the database creation and classification process was reviewed regularly by the author team during bi-weekly meetings held between February and June 2025. During these meetings, uncertain inclusion decisions, borderline cases, and ambiguous coding categories were discussed collectively, and classifications were refined through author consensus. When records were difficult to classify at the title and abstract stage, a conservative approach was used by retaining them for further assessment rather than excluding them prematurely.

This workflow was intended to improve consistency and transparency within the scope of a comprehensive review. Records were organized in Zotero, which was also used to identify and remove duplicates during database construction.

4. Solutions Characteristics

4.1. Technology Aspects

4.1.1. Platform

Platform and maturity set the capacity for what training and treatment solutions can achieve, how quickly they can respond, and whether they fit the real context of classrooms and clinics. As can be observed in Figure 3, VR accounts for 82.93% of solutions, with smaller shares for personal computers (PC) at 7.32%, mobile and web at 4.88% each. Most entries are pilot research initiatives at 72%, with the rest of 28% being commercial, available on public marketplaces. These proportions shape today's platform choices. VR was commonly used for head-tracked audience simulation, PC-based systems for supervised or data-intensive setups, and web and mobile systems for more accessible forms of repeated practice. Research prototypes were often more experimental, while commercial or course-integrated tools emphasize reliability, onboarding and analytics that are simple to read.

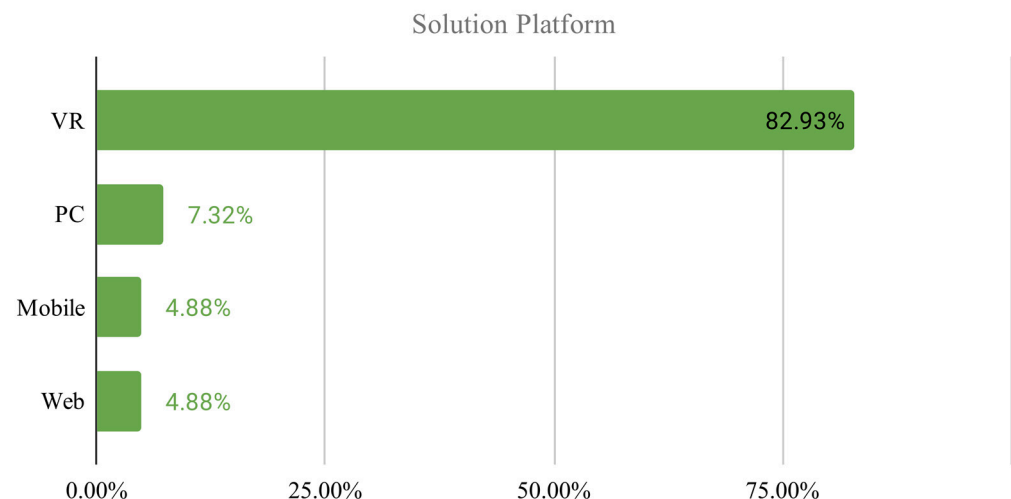


Figure 3. Solution platform.

VR is widespread because standalone and PC-tethered headsets can combine head-tracked immersion, spatial audio, and controllable audience simulation in a single training environment. At the same time, platform choice still involves practical trade-offs between portability, computational capacity, visual complexity, instructor oversight, and the stability of feedback or sensing features. Recent solutions display both the strengths and limits of the platform. Reeves et al. [17] deliver graded exposure with pre-recorded 360° video scenes arranged as a hierarchy by audience size, highlighting that video capture can support speech without interactive avatars. Bartyzel et al. [29] present a VR trainer developed in Unity for Meta Quest 2, in which voice features such as pitch and speech rate drive real-time animations of virtual characters and on-screen feedback, illustrating a system in which vocal features are linked directly to real-time character and environmental responses.

PC-tethered systems remain common in institutions, primarily favored for their enhanced computational capabilities that support complex environments, while also providing reliable mirrored displays for instructor monitoring. Recent studies deploy HTC Vive Pro or Oculus Rift on PCs in such settings, illustrating these practical benefits. Hosseini et al. [30] run a lab-based interview trainer where participants sit in front of a monitor, interact with an artificial interviewer, and later receive a recorded video or an

avatar replay as feedback, showing why a fixed PC workstation is appropriate when supervision and recording are prioritized over mobility. PC-tethered deployments remain common as laboratory setups for mirrored displays, live slide sharing, and desktop audio capture, enabling instructor oversight and standardized content delivery.

Web platforms offer broad device accessibility and may require little or no local installation, which can support early-stage practice and large-class deployment. The drawback is the limited monitoring capacity, often restricted to microphone use. Westwick et al. [31] present a scalable Learning Management System course, with narrated modules and webcam-recorded speeches submitted for instructor and peer review in discussion boards. The study spans 691 students across multiple sections. For many institutions, that is the right starting point: enough guidance to improve pacing and structure with little to no hardware overhead.

Mobile platforms can support at-home practice and brief speaking sessions using devices already available to many participants. Stupar-Rutenfrans et al. [24] deliver at-home, smartphone-based 360° video exposures arranged by audience size (empty, small, large), with 5 min scenes that can be replayed. The participants train in their own environment, as well as using their own presentation material. While the overall experience is less immersive, the benefit is increased frequency, which may be preferable in some cases.

In summary, platform selection involves trade-offs among immersion, accessibility, supervision, computational requirements, and opportunities for repeated practice. VR was the dominant platform for immersive audience simulation in the reviewed literature, whereas PC, web, and mobile systems supported different deployment contexts.

4.1.2. Solution Maturity

As illustrated in Figure 4, solution maturity leans toward research prototypes. These solutions implement more experimental features, such as advanced reactive crowds, sensor suites and rich analytics, yet many stop at the piloting stage because of performance, privacy, security and support. However, commercial entries tend to make conservative choices that keep sessions stable and interpretable, at the cost of lower innovation. Kryston et al. [25] integrate a PC-tethered Virtual Orator station into a large basic course, with automated metrics for gaze and vocal features, finding VR practice performs similarly to mirror or video lab practice. This example illustrates that adoption depends not only on platform features but also on how reliably a system fits existing course workflows.

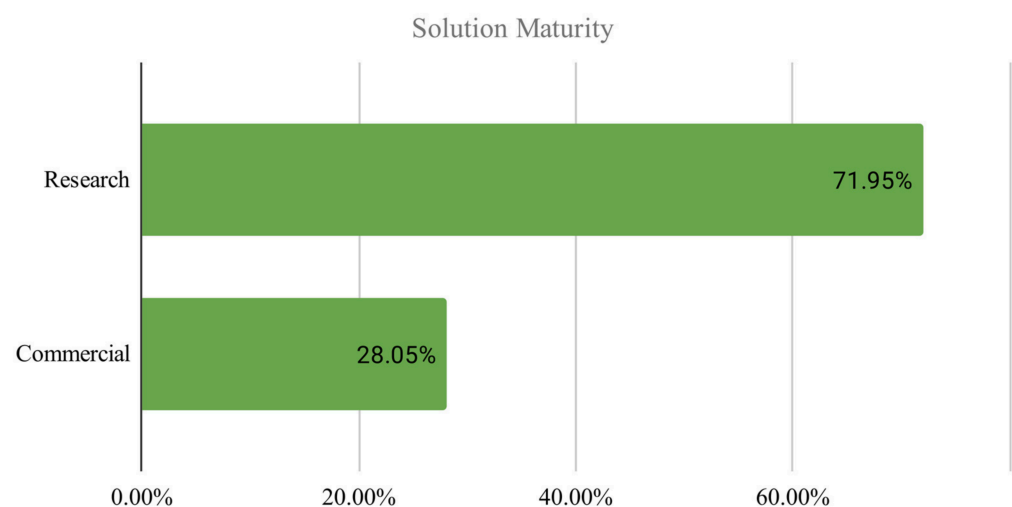


Figure 4. Solution maturity.

Solution maturity influences deployment context and feature priorities. In the reviewed corpus, research prototypes more often explored experimental sensing, audience, or AI features, whereas commercial systems generally emphasized stable workflows and ease of use.

4.2. Solutions Functionalities

4.2.1. Virtual Audience

The virtual audience is the set of simulated observers that a speaker faces inside a training system. In practice, it ranges from no spectators at all to static characters, pre-recorded 360° rooms, pre-animated avatars, reactive crowds that change with the user's delivery, sometimes with a large degree of control over their behavior. Educators and clinicians use these audiences to deliver social pressure and create structured rehearsal environments focused on skills such as pacing, prosody (rhythm, pitch, and intonation of speech), gesture, gaze, and stance. Virtual-audience designs vary in the relationship between the speaker and the simulated setting, the graphical representation of the audience, and the way difficulty changes are introduced when the audience adapts to the speaker. As a result of this comprehensive analysis, solutions were placed into one of the following categories, as can be observed in Figure 5: no audience, static, pre-recorded, animated, reactive, controlled by instructor/therapist, and customizable.

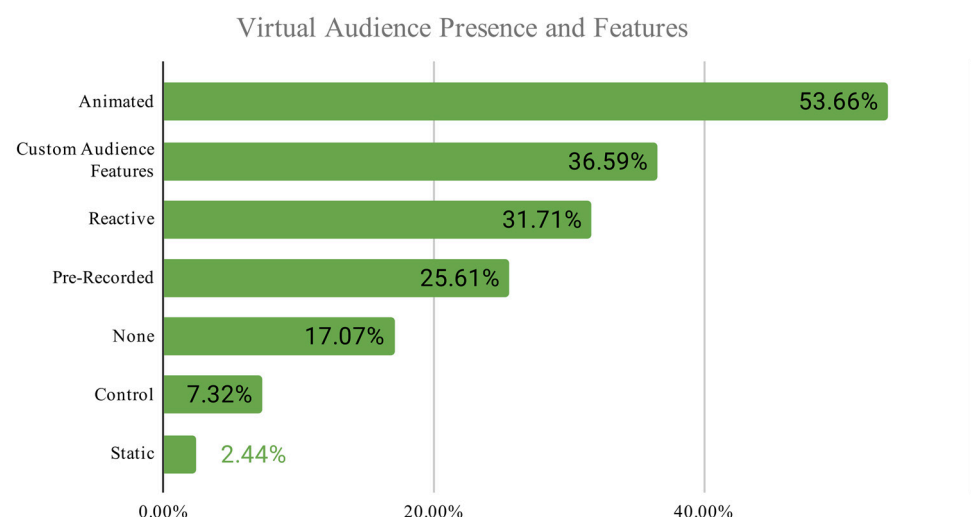


Figure 5. Virtual audience presence and features.

A small number of papers offer solutions with no virtual spectators. These are usually web or mobile applications that let learners rehearse delivery without any visible audience, often focusing on voice or structure. Kwok-Fai Lui et al. [14] present a smartphone coach that supports audience-free rehearsal by providing instant alerts on voice level and hand tremor during practice, and a post-session report on overall speech duration, body movement, and loudness. Schneider et al. [22] describe Presentation Trainer, a Kinect-based system that monitors vocal features such as loudness and pauses alongside posture and gestures, providing immediate visual or haptic cues. The solution focuses on delivery skills rather than speech content and cannot train gaze or audience scanning. Saukh and Maag [32] describe Quantle, a presentation coach application that computes pace, pitch, and pause duration in real time, provides subtle in-talk hints, and then offers post-session summaries for reflection. Across these audience-free systems, the shared emphasis is repeatable practice of articulation, pacing, and other measurable delivery features.

A static audience involves visible spectators that do not move, giving the speaker a sense of being looked at without the computational cost of animation. Schneider et al. [33]

extend Presentation Trainer with a VR module that places the speaker in a 3D classroom, utilizing a heads-up display (HUD) to overlay real-time visual coaching icons that follow the user's gaze. For performance reasons, the VR audience was kept very small and largely static. Static audiences provide visible observers with limited computational demand, but their lack of movement or responsiveness may reduce perceived social realism for some users.

The pre-recorded category wraps users in a 360° recording of a real space, trading interactivity for fidelity and realism. Reeves et al. [17] use pre-recorded 360° scenes as a graded exposure and show that this form of VRET reduces PSA relative to a control group, with effects broadly maintained for 10 weeks. Stupar-Rutenfrans et al. [24] use 360° recordings of the same room at three audience sizes for at-home exposure. Even with short sessions, the program led to lower PSA for students, particularly for those with higher baseline anxiety. Ferreira et al. [34] offer a low-cost trainer built on pre-recorded 360° videos. This approach wraps learners in real venues with high visual fidelity, but because the people in the video do not react to the speaker's performance, the dynamic interactivity of the audience is inherently limited. Pre-recorded audiences wrap learners in real venues, but the people in the video do not react, so feedback is limited during the talk and typically delivered after the attempt.

Animated audiences add idle motion and are the most common modern baseline, being light enough for standalone headsets. Barrett et al. [35] manipulate audience familiarity by replacing generic agents with photorealistic faces of people known to the speaker. Familiar faces in VR raise PSA relative to generic avatars, increasing difficulty, and pointing to identity cues in animated crowds as a design lever. Rodero and Larrea [36] stage practice in an auditorium where the instructor can drop in timed distractors, an audience cough, someone leaving or a challenging question, so speakers rehearse while resisting disruptions. Animated audiences are a pragmatic middle ground; they scale to standalone headsets while supporting graded exposure.

In reactive audiences, audience members respond to the speaker's performance. This reactivity can be as simple as posture shifts mapped to speaking rate or as rich as blended rules that account for multiple user performance markers. Van Ginkel et al. [2] implement a VR practice room where the audience reacts to the speaker's gaze and voice level. Learning gains matched the human expert feedback rather than exceeding it. Truong et al. [37] present a VR simulator where speech and gaze drive character reactions: the system transcribes speech, pauses, filler words, and updates audience attention based on those features and the speaker's gaze. In a small user study, participants preferred the version with audience feedback and said it helped them know where to focus. Two linked studies from Palmas et al. [38,39] offer reactive crowds with training extras. The first implements a VR trainer whose audience attention, computed from pace, fillers, volume, gaze and posture, drives real-time crowd animations and yields a detailed post-session report. The second, from 2021, builds upon this baseline by introducing real-time coaching prompts and individual attention indicators for each avatar, achieving higher technology-acceptance ratings than the unaugmented reactive crowd. Bartyzel et al. [29] animate a VR audience directly from the speaker's voice: pitch variability, timbre, and words-per-minute update an engagement threshold that selects character reactions, and, in a second scene, the weather. Taken together, these systems demonstrate different ways of linking measurable speaking behaviors to visible audience responses and post-session summaries.

A distinct tier appears once audience control is introduced, under the guidance of a therapist or instructor. Jakubowski et al. [40] present Cicero VR, a classroom trainer with configurable avatars and distractions triggered by an instructor. The solution logs pace, loudness, gaze, and gesture and produces a post-session report. In the case of customizable

audiences, users or instructors tailor what the audience looks like and how it behaves. Monteiro et al. [41] customize the identity of animated listeners by swapping in familiar 3D-scanned faces. These subjects, audience control and customization will be developed further in Section 4.3.2, since solutions that implement said features also tend to employ specialists in directing them.

Across the spectrum, virtual audiences differ in how they represent and use social cues. Static or pre-recorded scenes can provide a realistic entry point, animated audiences can support repeated practice, reactive crowds can connect listener behavior to measurable speaking features, and configurable audiences can allow instructors or users to adjust scenario intensity. Promising audience designs can align audience signals with delivery goals, let instructors or learners shape intensity in small, predictable steps, and use logs for debriefs. Identity and context also matter; who the people in the audience seem to be can shift anxiety and effort as much as how they move.

4.2.2. Solution Feedback Methods

Feedback is the information the system returns to the speaker about performance, capable of turning rehearsal into structured learning. During speaking sessions, systems can show small cues that nudge pace, loudness, gaze, and posture while the speaker is still on task. After the speaking session, they can present a debrief that aggregates metrics, visualizes trends across sessions, and suggests areas of improvement. Some solutions combine both: instant nudges with a richer summary afterwards. Some systems additionally analyze intonation, pause structure, gesture amplitude, and gaze distribution, while others allow instructors or peers to comment on recorded sessions. As observed in Figure 6, solutions were placed into one of the following categories: no feedback, in-session feedback, after-session feedback, and both in-session and after-session feedback.

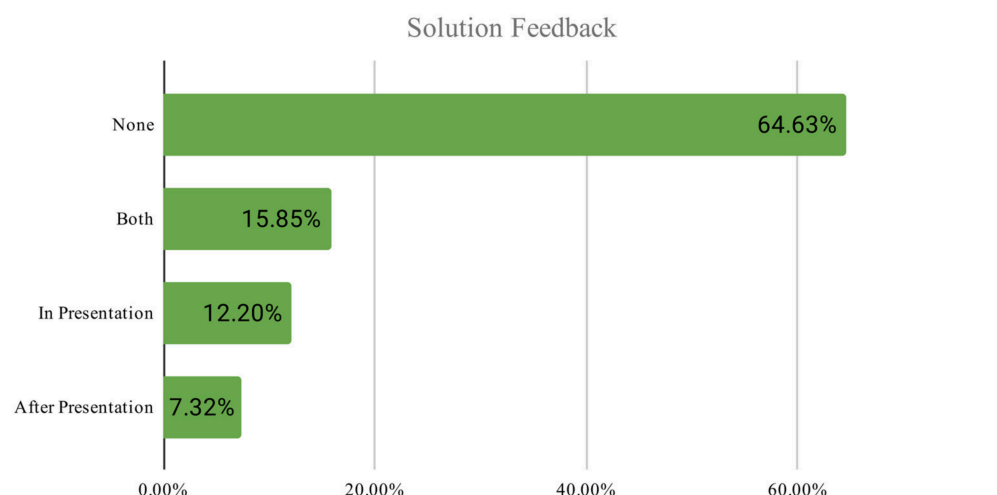


Figure 6. Solution Feedback.

Most systems offer no automated feedback, either because the study focuses on exposure effects or because authors prioritize scene control. In this group, the session itself is the intervention, and reflection happens outside the software. Education-facing pilots emphasize recorded practice with peer/mentor review over automated analytics. Leinonen et al. [42] add timeline comments, ratings, points, and a leaderboard to stimulate manual feedback and engagement in the EchoMe solution. In classrooms, the recordings of Colognesi et al. [43] let students rewatch and annotate performances, with added gains in delivery. Audience and embodiment studies rely on exposure during the session and assess outcomes only afterward. Kroczeck and Mühlberger [44] show that practicing before a

supportive virtual audience improves confidence in a subsequent in vivo speaking session, and Macey et al. [45] reduce PSA by increasing the height of the embodied user character in VR. Takac et al. [46] probe distress habituation across three consecutive speeches before an audience, emphasizing how repetition shapes habituation. Van Dis et al. [47] stage exposure on Day 1, then test spontaneous recovery and renewal a week later. The focus is on measuring the delayed return of fear, not on providing real-time performance feedback. Measurement-focused contributions exemplified by Kuai et al. [48] improve anxiety assessment by pairing VR speeches with functional near-infrared spectroscopy and expert ratings. Signals are analyzed retrospectively and not used as coaching cues.

In the in-presentation category, systems show live cues that learners can act on immediately. Nagao and Yokoyama [49] present Cyber Trainground, a virtual conference hall that computes voice metrics alongside eye contact, then turns them into feedback: lightweight audience reactions and pop-up messages that name the metric that triggered them. Premkumar et al. [27] augment self-guided VRET with real-time heart rate and frontal alpha asymmetry (FAA), an EEG marker linked to approach–avoidance motivation. Both signals are displayed as oscillating bars in the virtual lecture hall, which users are prompted to lower while presenting. Van Ginkel et al. [50] examine cues for eye contact and pace in VR, against delayed expert reports. Both conditions yield short-term gains with no statistical advantage.

Where systems delay information until speaking is over, after-presentation feedback usually takes the form of dashboards, textual summaries and scores. Kryston et al. [25] integrate a VR practice lab into a course and generate automated post-session feedback, but few students actually used those summaries. The authors recommend pairing analytics with instructor guidance. Ferreira et al. [34] offer a pre-recorded solution that delivers after-presentation reports on voice tone, eye contact, head movement, and speech duration, guiding gradual progression across difficulty levels. Hosseini et al. [30] compare two post-session feedback modes for interview practice: watching your own video versus watching a dissimilar avatar reenact your behavior. The avatar-based feedback increased identity-based interpretations of their behavior and shifted valence toward happiness during the feedback session, and it reduced galvanic skin response (GSR) peaks in the second half of that session (especially for highly anxious participants). In the subsequent interview, the avatar group showed higher average pitch and fewer weak words.

The final group delivers both in-presentation and after-presentation feedback. Belboukhaddaoui and Ginkel [51] compare in-presentation icons for eye contact and pace with after-presentation, computer-generated messages, and find no difference in delivery outcomes. They recommend brief, just-in-time nudges during the speech and richer analytics afterward, since the two moments support different parts of learning. Tangsripairoj et al. [52] introduce a solution with an in-session notes panel and an optional overlay for movement and eye contact. After each attempt, a results screen and browsable history summarize movement, time management, speech-content match, and eye contact. Ruiz-Capillas et al. [53] present a VR trainer with minigames that evaluate where and for how long speakers pause (to assess phrasing), alongside speaking volume adjusted for small, medium, or large rooms. After the attempt it provides a score summary so learners can plan the next exercise. Palmas et al. [39] evaluate a VR trainer that layers iconic, real-time voice and behavior cues plus audience attention markers over a reactive crowd, then closes with a post-talk star report.

In conclusion, feedback is most useful when it targets the behaviors that matter for delivery. Real-time audience reactions driven by vocal features, combined with structured post-session analytics, show how feedback can be explicitly connected to measurable aspects of delivery, and learners tend to rate such signals as more usable and relevant. Second,

pacing and transparency are crucial. Studies that explain what the live layer will do and limit in-presentation cues report higher acceptance and better use, while crowded screens or opaque metrics reduce uptake. Taken together, the reviewed systems illustrate mixed feedback phases, with brief nudges during the talk and richer reviews afterward, alongside clearer mapping to delivery criteria and, in a small number of systems, participant-facing biofeedback or scenario adaptation based on physiological arousal estimates.

4.2.3. Physiological and Behavioral Sensing

The reviewed systems capture physiological and neurophysiological signals, including HR, skin conductance as a measure of electrodermal activity (EDA), EEG, and functional near-infrared spectroscopy (fNIRS), alongside behavioral and performance signals such as voice, gaze, head movement, and body movement. These data may support retrospective assessment, clinician monitoring, participant-facing biofeedback, or bio-adaptive changes to the virtual environment. Recording a physiological signal alone was not considered biofeedback. Figure 7 summarizes the sensing modalities reported across the included studies.

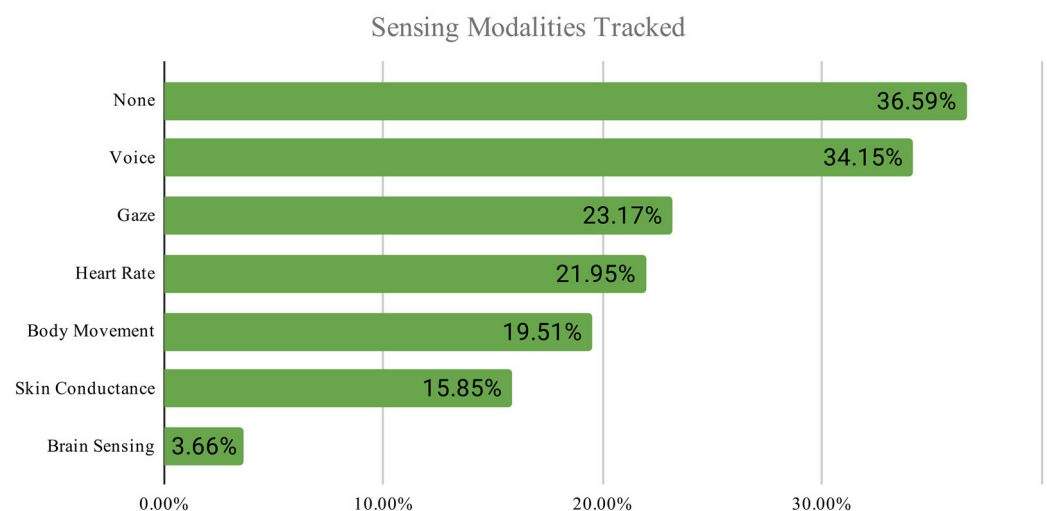


Figure 7. Sensing modalities tracked.

A minority of studies run exposure with no biofeedback and no sensor logging. The emphasis is placed on presence, audience design, or classroom workflows instead of physiology. Studies featuring large cohorts, or oriented on teaching, often employ analytics based on questionnaires and human evaluation. Huang [54] combines video-recorded oral presentations with mobile-assisted peer assessment in Tencent Docs, evaluating learning with assessment criteria, a standardized exam, and interviews. Glémarec et al. [55] present STAGE, a virtual audience control system for tutor-supervised speaking seminars, where instructors author “pedagogical narratives” and can adjust audience behavior via an interface. These designs are useful baselines for contexts that value simplicity and scale over continuous sensor measurement and its associated complexity.

Brain monitoring appeared in a small number of studies but served different functions. Premkumar et al. [27] returned EEG-derived frontal alpha asymmetry and HR information to participants during exposure, constituting participant-facing biofeedback. Salkevičius et al. [56] streamed EEG, HR, and skin-conductance data to clinicians for monitoring, whereas Kuai et al. [48] analyzed fNIRS data retrospectively in relation to anticipatory anxiety and speaking performance.

Skin conductance provides a continuous index of sympathetic arousal during speaking tasks, but it cannot identify anxiety on its own. Macey et al. [45] recorded EDA during a VR

speech task and observed lower self-reported PSA and a trend toward lower electrodermal activity in the taller-avatar condition. Mostajeran et al. [26] examined skin conductance alongside HR and cortisol under different audience-size conditions, while Rodero et al. [36] combined EDA with self-reported PSA before and after VR training. These findings show the value of EDA for tracking arousal-related change, but not as an isolated measure of anxiety.

Body and head movement are primarily behavioral or performance signals used to assess gesture, posture, stance, or movement during speaking. Niebuhr et al. [57] related gesture and stance patterns to vocal delivery, while Lui et al. [14] used phone accelerometry to detect hand tremor and movement during practice. Movement data may also support physiological estimation, as in Noori et al. [58], who predicted HR from headset motion.

Heart rate is frequently recorded as a continuous indicator of autonomic arousal, although its interpretation depends on individual baselines and may be affected by speech and movement. The VR trainer of Monteiro et al. [41] records HR while students present to familiar (3D scanned heads) or generic virtual audiences. This HR data is stored with a video recording of each session for review. The child-focused solution of Sülter et al. [59] avoids sensors by using brief exposures to pre-recorded audiences and self-report visual analog scales (nervousness, perceived heartbeat, sweaty palms). Yadav et al. [60] combine HR with speech and movement monitoring, profiling learners, showing that VR exposure reduces anxiety and suggesting potential for personalized, real-time feedback.

Gaze sensing captures where speakers direct their visual attention and is therefore a behavioral measure relevant to eye contact and audience engagement. LeFebvre et al. [61] estimated whether speakers looked toward the audience or their notes during VR presentations. Wechsler et al. [62] recorded gaze during attention training, alongside HR and skin conductance, and found that participants instructed to redirect attention toward social stimuli looked at audience faces more frequently.

Voice is a central behavioral and performance channel because many aspects of delivery can be adjusted during speaking. Acoustic sensing can quantify loudness, pitch variation, speech rate, pauses, and timbre, whereas automatic speech recognition (ASR) analyzes recognized words and content rather than vocal delivery itself. Ruiz-Capillas et al. [53] use minigames to train pause timing and speaking volume, providing live indicators and post-session scores. Bartyzel et al. [29] use pitch variability, timbre, and speech rate to drive audience and environmental reactions, representing behavior-adaptive rather than bio-adaptive modification. Tangsripairoj et al. [52] instead employ ASR to display live transcription and assess the similarity between spoken and expected content. Moldoveanu et al. [63] combine volume, words per minute, pauses, gaze, and hand movement to update audience attention and performance dashboards, while HR and EDA are displayed separately to the therapist for physiological monitoring. These examples show that voice sensing can support immediate coaching, retrospective analysis, and responsive environments, but vocal features remain performance-related signals rather than physiological or anxiety-specific measures.

In conclusion, most sensing studies used physiological or behavioral data for measurement, monitoring, or retrospective analysis, while participant-facing biofeedback and bio-adaptive modification remained less common. Physiological signals such as HR and EDA reflect nonspecific arousal rather than anxiety itself and require appropriate baselines, temporal alignment, preprocessing, and consideration of speech- and movement-related artifacts. Physiological, vocal, gaze, and movement data also raise privacy, data-governance, and automated-assessment concerns.

4.3. The Role of the Therapy in the Treatment Process

4.3.1. Types of Implemented Therapies in Treatment Solutions

Interventions can range from individual practice tools to structured exposure, which may include therapies such as CBT, acceptance and commitment therapy (ACT), and other types of psychotherapy. The majority of the solutions can be placed in the exposure therapy category (Figure 8), but they differ in pacing, the extent of clinician involvement, and how much the VR scenarios are tailored to the individual.

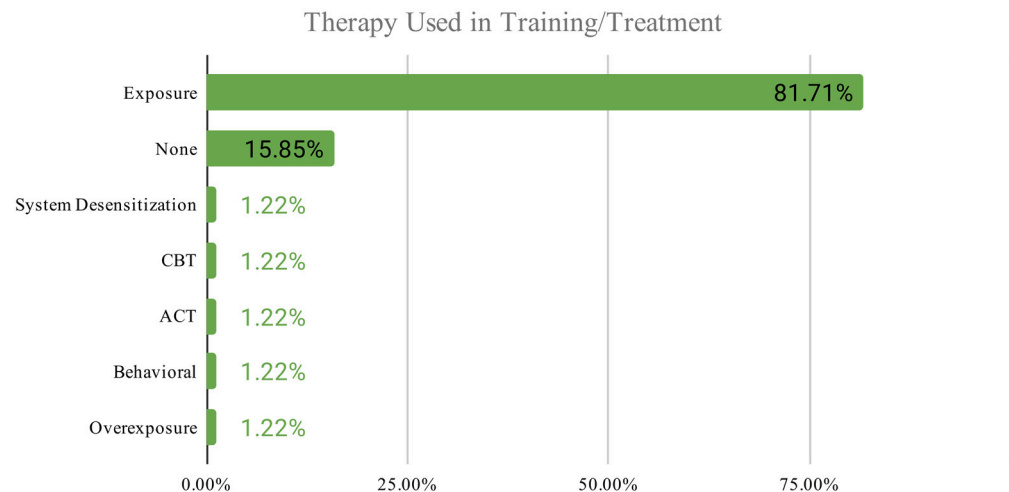


Figure 8. Therapy used in training/treatment.

A minority of systems implement no explicit therapy and frame the tool as skills training or classroom practice. For instance, Huang [54] used video-recorded presentations with peer-assisted assessment in large secondary classes, emphasizing scalability and teacher-guided reflection. Westwick et al. [31] tracked gains in self-perceived communication competence in an online public-speaking course, also without a formal therapeutic scaffold.

The core of the literature can be placed within the boundaries of exposure therapy, as in confronting individuals with stressors in controlled, repeatable settings. Valls-Ratés et al. [64] deliver three VR rehearsal sessions to high-school students and report reduced self-assessed anxiety and improved voice clarity. Lindner et al. [65] evaluate a clinician-led, one-session VRET in routine care and find a large immediate drop in PSA after the primary three-hour session. Reeves et al. [17] propose a randomized controlled trial of standalone 360-video VRET and show significant reductions in PSA, social anxiety, and fear of negative evaluation compared with control across both audience and empty-room variants. Sülter et al. [59] present a child-focused work that extends exposure to younger learners using pre-recorded scenes, with repeated brief sessions lowering in vivo presentation anxiety. In the paper set, Macdonald [66] is the only work that introduces a VR overexposure approach, with virtual audiences of up to 10,000 characters, and finds short-session gains in self-reported anxiety, confidence, and enjoyment.

Within behavioral therapeutic approaches, some studies combine exposure with behavior-change techniques. Zacarin et al. [67] ran six VR sessions layering diaphragmatic breathing, differential reinforcement, and functional analysis onto a hierarchy of progressively longer timed speeches, reporting SSPS and speech-quality gains. CBT often provides the framing for exposure, though many studies present exposure without employing the full CBT scaffold. Hanif et al. [68] present an Android application delivering CBT sessions plus graded VR exposure. The solution logs session times and progress, illustrating a functional CBT exposure package on mobile platforms. Third-wave approaches, a label used in the CBT literature for newer acceptance, mindfulness, and value-oriented

approaches as framed by Hayes [69], are uncommon in this set. Gorinelli et al. [70] implement an ACT intervention inside VR for university students: three lab sessions with graded social scenes while participants listen to audio ACT guidance that trains present-moment noticing, acceptance, defusion (stepping back from thoughts, seeing them as mental events, not facts), and choosing value-based actions. The trial reports reductions in social and communication anxiety and gains in psychological flexibility relative to a control group.

Overall, exposure-based designs predominated in the reviewed literature. Several individual studies reported reductions in PSA following brief or repeated VR exposure, but heterogeneity in study design and measurement prevents conclusions about optimal session number, clinician involvement, or personalization.

4.3.2. The Role of the Therapist in the Treatment Process

Therapists can shape PSA treatment at screening and goal setting before exposure, guidance during sessions, debriefing, or relapse planning if unwanted behaviors return. In VR-based solutions, they may sit beside the patient, connect through a dashboard, or remain outside the loop when the tool is designed for self-guided use. As a result of the analysis, solutions were placed into one of the following categories (Figure 9): solutions that include the therapist as part of the treatment loop, or fully unguided solutions, designed to work standalone.

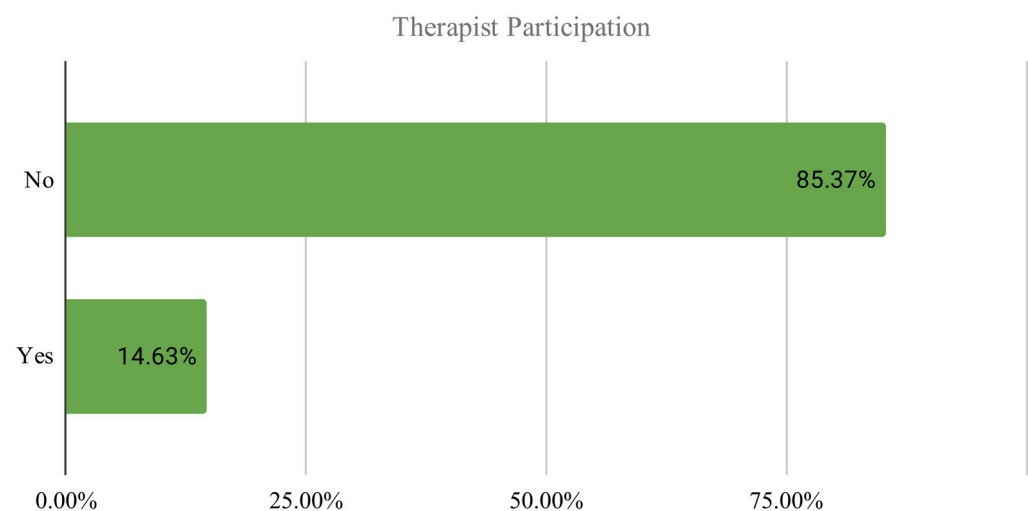


Figure 9. Therapist participation.

In solutions with no licensed therapist participation, systems are often framed as skills training, with peers, teachers, or instructors (rather than clinicians) providing human input. Reeves et al. [17] tested a standalone 360° video VRET, offering sessions weekly in a university research suite under supervision of a psychology doctoral student. The protocol scales well because clinical oversight is not mandatory for each session. Stupar-Rutenfrans et al. [24] tested mobile self-guided 360° VRET with an increasing number of virtual audience members on simple smartphone headsets, with pre/post anxiety checks. Also within the context of an educational institution, Kryston et al. [25] embedded a VR practice lab into an introductory public-speaking course, while comparing it to more traditional training methods such as mirror training, finding improvements across all training methods. These designs may fit institutional timetables and assessment cycles, although their implementation requirements and outcomes vary by context.

In the case of a licensed therapist playing an active role in the solution, clinicians can modulate real-time difficulty so that virtual situations can be adapted according to individual needs. Salkevičius et al. [56] present a VRET solution where a WebGL dashboard

allows clinicians to trigger panic-associated sound effects, simulate a computer crash, and set the virtual crowd's emotional reactions. Simultaneously, the system streams the patient's EEG, HR, and GSR data to the therapist in real time. Koller et al. 2019 and 2020 [71,72] developed a solution that keeps the therapist in the loop via a desktop graphical user interface (GUI) plus live avatar embodiment. The therapist selects an audience seat, takes over that avatar's body via Kinect, and speaks through it. In the 2019 pilot, adding this live channel did not reduce presence versus animations and was useful for guiding speakers. The 2020 paper formalizes "continuous interaction": hand-over gestures toggle between GUI control and in-VR embodiment, with spatialized voice and status indicators. Across the two papers, the emphasis is on therapist-mediated live interaction, including verbal guidance, nonverbal avatar control, and transitions from two-dimensional interaction to therapist embodiment within the virtual audience. Kahlon et al. [73] report a therapist-led one-session VRET for adolescents: difficulty advances through seven brief speech tasks with guidance between trials. The trial shows large, sustained PSA reductions, held at one- and three-month follow-ups. Zacarin et al. [67] present a clinic-based VR trainer that combines behavioral therapy with adjustable audience parameters. Therapists can set concrete speech goals like minimum speech duration and vocal pace, then review recorded behavior afterwards to give feedback and repeat targeted segments. Banakou et al. [74] test a single-session "imperceptible change" approach where a virtual therapist gradually becomes an audience. The therapist clones, with identical copies slipping out from behind the original, drifting to the sides, then morphing into different people. Most changes are confined to peripheral vision, as one interlocutor quietly turns into a full crowd without being extremely noticeable by the user. Using self-report outcomes, this single session performs at least as well as a five-session exposure.

Self-guided and teacher-supervised tools may support deployment within educational schedules, but they depend on clear instructions, appropriate safeguards, and interpretable feedback. Clinic-integrated systems instead allow therapists to contribute to screening, goal setting, exposure pacing, difficulty adjustment, and debriefing. Both brief and multi-session interventions were associated with favorable outcomes in some reviewed samples.

4.4. Evaluation Methods and Criteria

At a broad level, the evaluation instruments and measures reported across the included studies covered public-speaking and broader social anxiety, speaking performance and behavioral delivery, physiological or neurophysiological responses, and presence, usability, or user experience. Assessment timing included baseline or pre-intervention measurement, in-session assessment, immediate post-session or post-intervention evaluation, and, less consistently, follow-up. Physiological and neurophysiological outcomes are discussed in Section 4.2.3, while this section focuses primarily on questionnaires, behavioral measures, and rating instruments. Figure 10 summarizes the frequency of selected instruments, and Appendix A provides the full names of all abbreviated instruments used in this section and Appendix B.

In the case of core speech anxiety, four instruments are widely used: PRCS, PRPSA, PSAS and PRCA. Monteiro et al. [41] include the PRCS-SF to gauge public-speaking confidence while reporting session outcomes via HR and a posture-derived Body Score. Reeves et al. [17] assess a 360° VRET protocol with PSAS, LSAS-SR, BFNE, and presence (the sense of being physically located in the virtual space), documenting significant post-intervention reductions that persist at follow-up. Lindner et al. [75] assess PSA with PRCS-SF alongside LSAS-SR and BFNE. When general communication apprehension is of interest, Banakou et al. [74] include PRCA to complement public-speaking-specific measures.

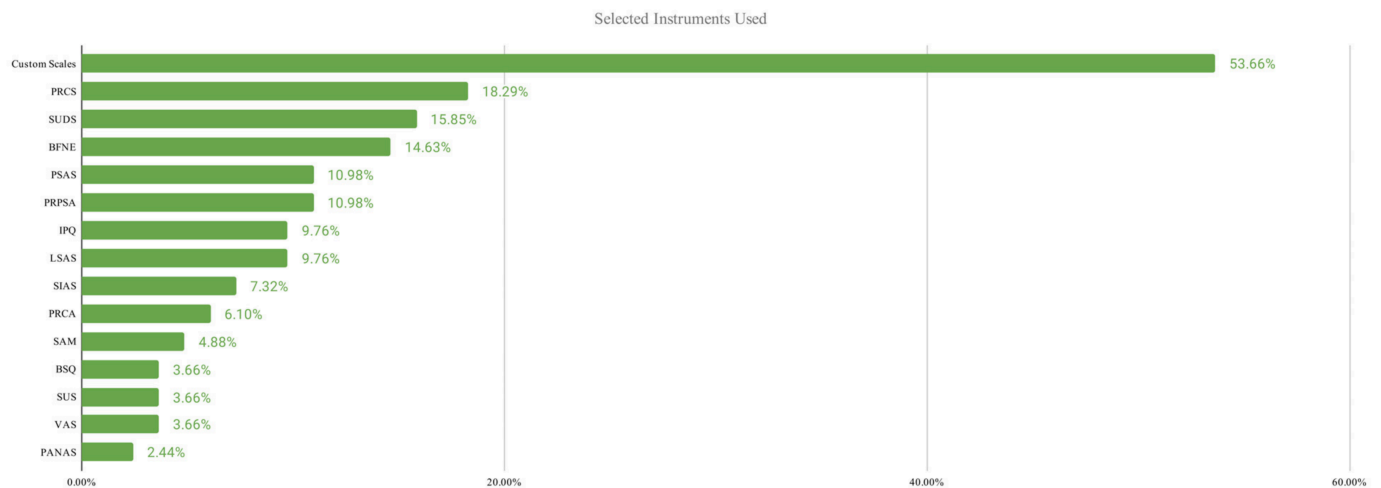


Figure 10. Selected Instruments Used.

Because public speaking sits inside the broader social anxiety family, LSAS, SIAS and BFNE frequently profile severity. Sarpourian et al. [76] assess outcomes with LSAS and PRPSA, adding an IPQ presence score, by comparing a VR exposure session with group counseling. Van Dis et al. [47] measure SUDS repeatedly during the speaking task and gather brief valence and arousal ratings before and after exposure blocks, while BFNE-II captures fear of negative evaluation. Premkumar et al. [77] assess outcomes in a self-guided VRET protocol using LSAS and BFNE, contextualizing scores with published norms for LSAS.

To document how the medium itself is experienced, studies often add measures for presence and usability. Kollöffel and Heuvel [78] evaluate VR presentation training with a mixed-methods design and quantify presence using the IPQ. Mostajeran et al. [26] examine the effect of virtual audience size using a set of questionnaires that includes IPQ, the SSQ, and affective self-reports such as SAM. Krocze and Mühlberger [44] measured presence with IPQ and MPS, finding that a supportive audience increased Social Presence and was followed by higher instructor-rated speaker confidence in a Zoom seminar. Bartyzel et al. [29] evaluate a voice-responsive VR speaking trainer with the SUS administered after use. Min and Jeong [28] introduce an LLM-assisted VR questions-and-answers (Q&A) practice system and pair SUS with the UEQ.

Momentary distress is routinely measured using SUDS and VAS, particularly in younger cohorts. Landkroon et al. [79] recorded SUDS at the start and at one-minute intervals during a single VR speech to track distress over time, with participants also rating the speech difficulty on a VAS. Wechsler et al. [62] assessed SUDS during speaking and also used the STAI state scale and PANAS to index state anxiety and positive/negative affect. Takac et al. [46] recorded SUDS within and between speeches to track habituation, with SIAS at baseline and IPQ post-exposure to check presence. Sülter et al. [59] present a child-focused proof of concept, using brief visual analog scales for nervousness, heart rate, and sweaty palms to capture state anxiety with minimal disruption.

Speaking performance outcomes were assessed through automated behavioral measures and human ratings. Automated approaches examined gaze distribution, speech rate, pauses, loudness, pitch and prosody, gesture, posture, and body movement, while instructor, expert, peer, or self-assessment rubrics addressed clarity, organization, persuasiveness, confidence, and overall delivery quality [25,48,57,64]. These outcomes should be distinguished from self-reported anxiety because reduced distress does not necessarily indicate improved speaking performance. Independent or blinded assessment was reported in

some studies, but assessor procedures were not described consistently enough for reliable quantitative comparison.

Finally, custom scales remain widely used in education as well as in clinical pilots. Colognesi et al. [43] grade elementary pupils' oral performances with a custom rubric, noting that when assigned to record videos, pupils can create successive iterations of their performances using self and peer feedback. Zacarin et al. [67] combine standard outcome scales, SSPS and SUDS, with brief bespoke tools such as a recording sheet, a semi-structured interview and a client-satisfaction questionnaire to track change, functional goals and acceptability within a behavioral therapy plus VR exposure protocol.

Across the paper set, outcome assessment combined validated anxiety scales, brief in-task distress ratings, automated behavioral measures, human performance ratings, and presence or usability instruments. Variation in the constructs assessed and the timing of measurement limits direct comparison across studies. Intervention dose, follow-up, transfer to real-life speaking, dropout and adherence, cybersickness or other adverse effects, symptom deterioration, accessibility, and tolerability were reported inconsistently and were not predefined extraction variables. They are therefore acknowledged where explicitly reported rather than presented as corpus-wide frequencies.

4.5. Performance Training Compared to Anxiety Treatment

The reviewed interventions can be distinguished by their primary orientation: educational or professional skills training, and anxiety-oriented treatment. This distinction is based on the principal goal and evaluation framework rather than the institutional setting alone. A university-based intervention may primarily target elevated PSA, while a VR system used in a laboratory may focus on presentation competence or technology validation. The two orientations therefore overlap and should not be treated as mutually exclusive categories.

Performance-training studies included learners in course-based and school-based settings who were not recruited primarily on the basis of a clinical diagnosis. Their main goals were to improve presentation competence, including vocal delivery, pacing, gaze, gesture, organization, confidence, and self-perceived competence. Reported outcomes in these examples included self-perceived communication competence, vocal characteristics, gesture, persuasiveness, charisma, anxiety, satisfaction, and perceived usefulness. The delivery models represented in these studies ranged from course-based instruction to unguided VR practice. Findings were mixed: some studies reported improvements in selected delivery or anxiety outcomes, while others found limited changes or no clear advantage over conventional practice [31,64].

Anxiety-oriented studies more often recruited participants with elevated PSA, fear of negative evaluation, or broader social-anxiety symptoms. Their main goals were to reduce PSA and related anxiety or arousal through structured exposure. Reported outcomes included validated PSA and social-anxiety scales, fear of negative evaluation, in-session anxiety or arousal, physiological measures, and follow-up assessments. Supervision ranged from therapist-led protocols to fully self-guided exposure, while intervention dose ranged from a single structured session to repeated sessions with later follow-up. Several studies reported reductions in self-reported PSA across therapist-led, self-led, and standalone formats, although differences in design and measurement prevent conclusions about an optimal delivery model [17,75,77].

Educational studies therefore primarily define success through competence and performance, whereas anxiety-oriented studies prioritize symptom reduction and the ability to function in audience situations. Nevertheless, many interventions assess both performance and anxiety, reflecting the relationship between speaking skill, confidence, and

apprehension. Intervention duration, session number, adherence, and follow-up were not predefined extraction variables in this review; consequently, dose is described through representative studies rather than compared quantitatively across the full corpus. These differences support interpreting the two orientations as related but heterogeneous evidence streams rather than as a single body of effectiveness evidence.

5. Discussion

5.1. Design Implications and Preliminary Application Framework

The reviewed literature does not identify a single technological feature that reliably reduces PSA. A recurring cross-study pattern is the use of structured, repeated, or graded speaking exposure across pre-recorded, animated, guided, and self-guided formats. Reactive audiences can connect speaker behavior to adaptive social cues, but direct comparisons with static or pre-recorded audiences remain scarce. Likewise, participant-facing biofeedback appears in only a small subset of studies, so its added value over exposure alone remains uncertain. VR was frequently selected for controlled audience simulation and repeated practice. These limits inform the preliminary framework below, which treats audience reactivity, feedback, and sensing as configurable design components requiring further comparative validation.

This section builds on the recurring design patterns identified in the review to outline a preliminary framework for a future public-speaking training and treatment application. The framework is presented as a synthesis-driven design proposal that translates current evidence, implementation tendencies, and unresolved gaps into a coherent application model. Its purpose is to show how the main strands of the literature—educational skills training, anxiety-oriented exposure, audience simulation, feedback, and sensing—can be combined in a system that remains usable for learners, educators, and, when needed, clinical facilitators.

A central implication of the reviewed studies is that future systems should support graded speaking scenarios. These may vary by venue, audience size, audience familiarity, audience behavior, and task difficulty, allowing users to progress from low-pressure rehearsal to more demanding performance situations. Static or pre-recorded audiences can provide accessible entry points, while animated and reactive audiences can increase realism by responding to gaze, speech rate, pauses, vocal intensity, or other observable performance cues. Such reactivity should be meaningful rather than punitive: the goal is not to overwhelm the speaker, but to make the simulated audience a useful source of social pressure and feedback.

Feedback design is equally important. Several reviewed systems used limited, interpretable real-time cues tied to behaviors the speaker could adjust immediately, such as pace, loudness, eye contact, posture, or breathing, while more detailed metrics were commonly presented after the session. The preliminary framework therefore combines restrained in-session guidance with richer post-session analytics as a design hypothesis requiring comparative evaluation.

Physiological and behavioral sensing can further strengthen such a system, provided that signals are interpreted carefully. HR, EDA, gaze, voice, and movement can help describe arousal and performance, but they should not be treated as direct or isolated measures of anxiety. Their main value lies in supporting reflection, therapist or teacher monitoring, and adaptive training decisions when adequate baselines and contextual interpretation are available. Biofeedback is especially relevant when physiological information is returned to the user in a clear, actionable form, such as breathing support or arousal-awareness prompts. A small number of prototypes explored participant-facing biofeedback or bio-adaptive changes driven by physiological arousal estimates. These approaches

remain emerging and should be distinguished from the more common practice of recording physiological signals for monitoring or retrospective analysis.

The proposed framework should support both self-guided and facilitated use. In educational contexts, teachers may configure assignments, review reports, and guide repeated practice. In anxiety-oriented or clinical contexts, therapists may use the same system to adjust exposure difficulty, monitor distress, and structure debriefing. This dual-use model reflects the core contribution of the review: public speaking technologies are most promising when they do not treat training and treatment as isolated domains but as overlapping forms of structured practice requiring different levels of feedback, supervision, and personalization. Even so, the recurring convergence around adaptive audiences, structured feedback, and sensing suggests a direction for future public-speaking training and treatment systems.

5.2. Limitations

Several limitations should be considered when interpreting the findings and the preliminary framework proposed above. First, although uncertain cases were discussed within the author team, the review did not include formal independent duplicate screening, independent coding, or inter-rater agreement statistics. Second, no formal risk-of-bias or methodological-quality appraisal was performed, meaning that the synthesis maps technological, pedagogical, and therapeutic design patterns without weighting findings by study quality or ranking interventions by comparative effectiveness. Accordingly, findings from individual studies are used mainly to illustrate design patterns and implementation trends, rather than to suggest that all studies provide the same level of evidence.

Finally, the evidence base remains heterogeneous, with many small prototypes, feasibility studies, classroom interventions, and anxiety-oriented trials differing in population, duration, outcome measures, and follow-up. Positive technology-validation studies may also be more visible than null or unsuccessful implementations. Future work should therefore strengthen comparative evaluation and report transfer, adherence, cybersickness, adverse effects, accessibility, and tolerability more consistently.

6. Conclusions

Across 82 participant-validated studies, technology-mediated public speaking interventions showed substantial diversity in platform, audience design, feedback, sensing, and therapeutic framing. Nevertheless, the field remains strongly concentrated around VR research prototypes, while comparative evidence and longer-term validation remain limited. VR accounted for 82.93% of the included solutions and provided the dominant platform for audience simulation, repeated practice, exposure, and performance training. Several controlled, classroom, and exploratory studies reported reductions in PSA or improvements in selected speaking outcomes. However, variation in populations, intervention goals, study designs, and outcome measures, together with the absence of effect-size synthesis and formal risk-of-bias assessment, prevents firm conclusions about consistent effectiveness, platform superiority, or optimal intervention features.

Our analysis identifies several recurring design patterns across educational and anxiety-oriented systems. These include graded speaking scenarios, configurable or reactive virtual audiences, limited and actionable in-session cues, detailed post-session analytics, behavioral sensing, carefully interpreted physiological monitoring, and different levels of teacher or therapist involvement. Physiological biofeedback, bio-adaptive modification, and LLM-supported audience interaction appeared in smaller subsets of studies and remain emerging directions rather than established determinants of effectiveness.

Drawing on these recurring patterns and the gaps identified across the reviewed systems, we propose a preliminary application framework that connects educational skills training with anxiety-oriented intervention. The framework combines progressive scenario difficulty, audience adaptation, immediate coaching for modifiable speaking behaviors, longitudinal post-session analytics, optional physiological feedback, and configurable facilitator involvement. It translates the review findings into a coherent design direction for future public speaking applications.

The evidence base includes randomized and controlled trials alongside small prototypes, feasibility studies, technology validation experiments, and classroom evaluations. Future research should therefore test the proposed framework and its individual components through direct comparative studies, standardized anxiety and performance outcomes, longer follow-up, and assessment of transfer to real-life speaking. Future studies should report adherence and dropout, accessibility barriers, tolerability and adverse events, assessor independence or blinding. More consistent reporting, combined with direct comparative and longitudinal designs, is needed to determine which combinations of platform, audience design, feedback, sensing, and professional support produce sustained improvements in speaking performance or anxiety, for which populations, and under which educational or clinical conditions.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/info17080741/s1>, Workbook: Supplementary_Data_Extraction_and_Coding_Workbook.

Author Contributions: Conceptualization, D.-I.D., A.A.M. and A.D.B.M.; methodology, D.-I.D., A.A.M. and A.D.B.M.; software, D.-I.D.; validation, D.-I.D., A.A.M., A.D.B.M., A.M.A. and A.I.B.; formal analysis, D.-I.D., A.A.M. and A.D.B.M.; investigation, D.-I.D.; resources, D.-I.D.; data curation, D.-I.D.; writing—original draft preparation, D.-I.D.; writing—review and editing, D.-I.D., A.A.M., A.D.B.M., A.M.A. and A.I.B.; visualization, D.-I.D.; supervision, A.A.M., A.D.B.M., A.M.A. and A.I.B.; project administration, D.-I.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Regional Development Fund (ERDF) through the Smart Growth, Digitalization and Financial Instruments Programme 2021–2027 (PoCIDIF), within the project “Romanian Hub for Artificial Intelligence—HRIA”, SMIS Code 2021–351416.

Data Availability Statement: The data extraction and coding workbook supporting this review is provided in the Supplementary Materials.

Acknowledgments: During the preparation of this work, the authors utilized several versions of OpenAI’s GPT-5 models and Gemini 3.1 Pro to assist with proofreading, correcting grammatical errors, and rephrasing select sections to improve coherence and reduce the overall word count. After using this tool, the authors rigorously reviewed and edited the content.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. PSA Evaluation Metrics

1. Core Public Speaking and Communication Anxiety

These instruments specifically target the user’s fear, confidence, and cognitive patterns related to speaking in front of an audience or communicating in general.

PRCA: Personal Report of Communication Apprehension.

PRCS: Personal Report of Confidence as a Speaker.

PRCS-SF: Personal Report of Confidence as a Speaker—Short Form.

PRPSA: Personal Report of Public Speaking Anxiety.

PSAS: Public Speaking Anxiety Scale.

SSPS: Self-Statements during Public Speaking scale.

2. Broad Social Anxiety and Fear of Evaluation

These scales are used to profile the broader severity of the speaker's social fears.

BFNE: Brief Fear of Negative Evaluation.

BFNE-II: Brief Fear of Negative Evaluation—II.

LSAS: Liebowitz Social Anxiety Scale.

LSAS-SR: Liebowitz Social Anxiety Scale—Self Report.

SIAS: Social Interaction Anxiety Scale.

3. Momentary Distress and Affect (In-Session Measures)

These scales are typically administered right before, during, or immediately after a speech to capture real-time state anxiety, distress, and emotional valence.

PANAS: Positive and Negative Affect Schedule.

SAM: Self-Assessment Manikin.

STAI: State-Trait Anxiety Inventory.

SUDS: Subjective Units of Distress.

VAS: Visual Analog Scales.

4. Presence, Usability, and System Experience

These instruments document how the technology itself is experienced by the user, evaluating immersion and the software's ease of use.

IPQ: Igroup Presence Questionnaire.

MPS: Multimodal Presence Scale.

SSQ: Simulator Sickness Questionnaire.

SUS: System Usability Scale.

UEQ: User Experience Questionnaire.

5. Specialized and Custom Measures

This category covers bespoke evaluation criteria used for specific embodiment studies.

BSQ: Body Sensations Questionnaire.

Custom Scales: Non-standardized, bespoke rubrics or questionnaire.

Appendix B. Condensed Review Workbook

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART001	2025	Developing Immersive Virtual Reality Space for Public Speaking Training: English Debate Room in Spatial Platform [80]	4	VR	Research	–	Exposure	–	–	PRCA
ART002	2025	Exploring user reception of speech-controlled virtual reality environment for voice and public speaking training [81]	5	VR	Research	Voice	Exposure	Animated, Reactive, Custom Audience Features	In Presentation	PRCS-SF, SUS

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART003	2024	Augmenting self-guided virtual-reality exposure therapy for social anxiety with biofeedback: a randomized controlled trial. [27]	73	VR	Research	Heart Rate, Voice, Body Movement, Brain Sensing	Exposure	Animated, Reactive, Custom Audience Features	In Presentation	SPIN, PRCS, PSAS, LSAS, BFNE, VAS, Custom Scale (Presence), Behavioral Inhibition System Appraisal Subscale
ART004	2024	Avatar-Based Feedback in Job Interview Training Impacts Action Identities and Anxiety [30]	36	PC	Research	Skin Conductance	–	–	After Presentation	Measure of Anxiety in Selection Interview, SAM, Behavior Identification Form, Affective Rating System (Modified)
ART005	2024	Desensitizing Anxiety Through Imperceptible Change: Feasibility Study on a Paradigm for Single-Session Exposure Therapy for Fear of Public Speaking. [74]	45	VR	Research	–	Exposure	Animated, Reactive, Custom Audience Features	–	PRCA, State Perceived Index of Competence, Implicit Association Test, STAI, Custom Scale (Body Ownership and Audience Response)
ART006	2024	Effects of audience familiarity on anxiety in a virtual reality public speaking training tool [41]	10	VR	Research	Heart Rate	Exposure	Animated, Custom Audience Features	–	Custom Scale (Demographics), PRCS
ART007	2024	Empowering oral proficiency in a large-scale class: video-recorded oral presentations and mobile-assisted peer assessment in a Chinese Middle school [54]	94	Mobile	Commercial	–	–	–	–	Oral Test Rubric (Modified), Shaoguan City-Wide Standardized Exam, Custom Scale (Semi-Structured Interviews)
ART008	2024	Exploring the influence of audience familiarity on speaker anxiety and performance in virtual reality and real-life presentation contexts [35]	10	VR	Research	Voice	Exposure	Animated, Custom Audience Features	–	Custom Scale (Self-Confidence), PRCS-SF, Custom Scale (Reflection)
ART009	2024	Gestures and Feet Can Tell Us how You Speak! On the Relationships between Voice and Body Language in VR Public Speeches [57]	28	VR	Commercial	Voice, Body Movement	Exposure	Animated, Reactive, Custom Audience Features	–	–

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART010	2024	Improving virtual reality exposure therapy with open access and overexposure: a single 30 min session of overexposure therapy reduces public speaking anxiety [66]	29	VR	Research	–	Overexposure	Animated, Reactive, Custom Audience Features	–	Custom Scale (Anxiety, Confidence, and Enjoyment)
ART011	2024	Positive mood induction does not reduce return of fear: A virtual reality exposure study for public speaking anxiety. [82]	62	VR	Research	Skin Conductance	Exposure	Pre-Recorded	–	PRPSA, SAM, PANAS (Modified), Custom Scale (Speech Ratings), Mood and Need for Threat Questionnaire, Custom Scale (Post-Experimental)
ART012	2024	Public Speaking Q&A Practice with LLM-Generated Personas in Virtual Reality [28]	20	VR	Commercial	–	Exposure	Animated, Reactive, Unclear	–	Custom Scale (Post-Study), SUS, UEQ
ART013	2024	Speech Lab VR: A Virtual Reality System for Improving Presentation Skills [52]	10	VR	Research	Voice, Body Movement, Gaze	Exposure	Animated, Custom Audience Features	In Presentation, After Presentation	Custom Scale (Post-Test)
ART014	2024	The affordances of a mobile video-tagging tool for evaluating presentation skills in a second language [83]	35	Web	Commercial	–	–	–	–	Custom Scale (Self-Reflection), Custom Scale (Tool Evaluation)
ART015	2024	Virtual Reality on Public Speaking Phobia mitigation [34]	15	VR	Research	Voice, Body Movement, Gaze	Exposure	Pre-Recorded	After Presentation	–
ART016	2024	Voice-Responsive Virtual Reality Training Environment for Occupational and Non-Professional Voice Users [29]	5	VR	Research	Voice	Exposure	Animated, Reactive,	In Presentation	PRCS, SUS
ART017	2024	VR Public Speaking Simulations Can Make Voices Stronger and More Effortful [84]	47	VR	Commercial	Voice, Body Movement	Exposure	Pre-Recorded	–	SUDS

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART018	2023	Effect of Metacognitive-Based Digital Graphic Organizer on Learners' Oral Presentation Skill and Self-Regulation of Learning Awareness [85]	27	PC	Commercial	–	–	–	–	Self-Regulation of Learning Scale
ART019	2023	Encouraging participant embodiment during VR-assisted public speaking training improves persuasiveness and charisma and reduces anxiety in secondary school students [86]	70	VR	Commercial	Voice, Body Movement	Exposure	Animated, Custom Audience Features	In Presentation, After Presentation	SUDS, Custom Scale (Persuasiveness and Charisma)
ART020	2023	Feeling Small or Standing Tall? Height Manipulation Affects Speech Anxiety and Arousal in Virtual Reality. [45]	61	VR	Research	Heart Rate, Skin Conductance	Exposure	Animated, Unclear	–	SAM, PSAS
ART021	2023	Immersive Phobia Therapy through Adaptive Virtual Reality and Biofeedback [63]	4	VR	Research	Heart Rate, Skin Conductance, Voice, Gaze	Exposure	Animated, Reactive, Custom Audience Features	In Presentation	–
ART022	2023	Improving the oral language skills of elementary school students through video-recorded performances [43]	256	PC	Research	–	–	–	–	Custom Rubric (Oral Communication Evaluation Matrix)
ART023	2023	Public speaking training in front of a supportive audience in Virtual Reality improves performance in real-life. [44]	73	VR	Research	–	Exposure	Animated, Custom Audience Features	–	SPIN, BFNE, PRCS, Generalized Self-Efficacy Scale, IPQ, MPS
ART024	2023	Using 360 VR Videos in Presentation Rehearsals: Comparing Anxiety Levels [87]	52	VR	Research	–	Exposure	Pre-Recorded	–	Foreign Language Classroom Anxiety Scale (Modified)
ART025	2023	Virtual reality acceptance and commitment therapy intervention for social and public speaking anxiety: A randomized controlled trial [70]	76	VR	Research	Heart Rate, Skin Conductance	ACT	Pre-Recorded	–	SIAS, PRCA, Mental Health Continuum-SF, Perceived Stress Scale, VAS, Comprehensive Assessment of ACT Processes, Self Compassion Scale-SF, BFNE

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART026	2022	Controlling the Stage: A High-Level Control System for Virtual Audiences in Virtual Reality [88]	16	VR	Research	–	Exposure	Animated, Reactive, Custom Audience Features	–	Custom Scale (Semi-Structured Interviews), PSAS
ART027	2022	Exploring Individual Differences in Public Speaking Anxiety in Real-Life and Virtual Presentations [89]	55	VR	Commercial	Heart Rate, Skin Conductance, Skin Temperature	Exposure	Animated, Reactive, Custom Audience Features	–	Custom Scale (Demographics), STAI, Communication Anxiety Inventory, PRPSA, Big Five Inventory, BFNE, Reticence Willingness to Communicate, Custom Scale (Prior Daily Experiences), Communication Anxiety Inventory (Modified), STAI (Modified), State-Anxiety Enthusiasm, BSQ, Custom Scale (Presentation Preparation Performance), Custom Scale (VR Presence), Sense Questionnaire, Slater-Usch-Steed Questionnaire
ART028	2022	Future-Oriented Positive Mental Imagery Reduces Anxiety for Exposure to Public Speaking. [79]	43	VR	Research	–	Exposure	Pre-Recorded	–	PRCS, VR Experience Scale, Custom Scale (Visual Analog Scales), SUDS, Behaviors Checklist
ART029	2022	Public Speaking Simulator with Speech and Audience Feedback [37]	10	VR	Research	Voice, Gaze	Exposure	Animated, Reactive	–	–
ART030	2022	SpeakApp-Kids! Virtual reality training to reduce fear of public speaking in children—A proof of concept [59]	89	VR	Research	Heart Rate, Skin Conductance	Exposure	Pre-Recorded	–	VAS, PRPSA (Modified), Social Anxiety Scale for Children (Modified), Custom Scale (Speech Preparation)

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART031	2022	Talk and Play: a VR Application to Improve Volume and Silences Use in Verbal Communication [53]	14	VR	Research	Voice	Exposure	–	In Presentation, After Presentation	Custom Scale (Gamer Typology), Custom Scales (Multiple Constructs)
ART032	2022	The effect of virtual reality therapy and counseling on students' public speaking anxiety. [76]	30	VR	Research	–	Exposure	Pre-Recorded	–	LSAS, PRPSA, IPQ, IGP
ART033	2022	Unguided virtual-reality training can enhance the oral presentation skills of high-school students [64]	50	VR	Commercial	Voice, Body Movement	Exposure	Pre-Recorded	In Presentation	SUDS, Custom Scale (Satisfaction)
ART034	2022	Using Virtual Reality Simulation to Reduce Stage Fright during Public Appearances [90]	6	VR	Research	Heart Rate	Exposure	Animated	–	Custom Scale (System Quality), Custom Scale (Perceived Stress Level)
ART035	2022	Virtual Reality with Distractors to Overcome Public Speaking Anxiety in University Students [36]	100	VR	Commercial	Skin Conductance	Exposure	Animated, Reactive, Custom Audience Features (Instructor Control)	–	PSAS
ART036	2021	360° Video virtual reality exposure therapy for public speaking anxiety: A randomized controlled trial. [17]	51	VR	Research	–	Exposure	Pre-Recorded, Multiple Videos Available	–	PSAS, LSAS-SR, BFNE, IPQ
ART037	2021	Exploring Eye Contact in Virtual Environments: The Composer Mirror Tool, Areas of Interest, and Public Speaking Competency [61]	12	VR	Research	Gaze	Exposure	Pre-Recorded	–	Custom Scale (Open and Closed Questions)
ART038	2021	Higher anxiety rating does not mean poor speech performance: dissociation of the neural mechanisms of anticipation and delivery of public speaking. [48]	24	VR	Research	Brain Sensing	Exposure	Animated	–	PRCS, Social Performance Rating Scale
ART039	2021	"Imagine All the People": Imagined Interactions in Virtual Reality When Public Speaking [91]	17	VR	Research	–	Exposure	Pre-Recorded	–	PRPSA, Survey of Imagined Interactions

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART040	2021	Incorporating Virtual Reality Training in an Introductory Public Speaking Course [25]	140	VR	Commercial	Voice, Gaze	Exposure	Animated, Reactive, Custom Audience Features	After Presentation	BFNE, Intrinsic Motivation Inventory, Spatial Presence Experience Scale, PRPSA, STAI, Rosenberg Self-Esteem Scale, Self-Consciousness Scale, PRCS, NASA Task Load Index, Cognitive Demand Subscale
ART041	2021	Indifferent or Enthusiastic? Virtual Audiences Animation and Perception in Virtual Reality [55]	20	VR	Research	Gaze	Exposure	Animated, Reactive, Custom Audience Features	–	Custom Scale (Satisfaction)
ART042	2021	Look at the Audience? A Randomized Controlled Study of Shifting Attention From Self-Focus to Nonsocial vs. Social External Stimuli During Virtual Reality Exposure to Public Speaking in Social Anxiety. [62]	41	VR	Research	Heart Rate, Skin Conductance, Gaze	Exposure	Animated	–	SUDS (Modified), STAI, PRCS, BSQ, IPQ, SPIN, Custom Scale (Participant Impression), BFNE, PRCS (Modified), IPQ, Custom Scale (Rating Items), SUDS, STAI, PANAS, BSQ
ART043	2021	Old Fears Die Hard: Return of Public Speaking Fear in a Virtual Reality Procedure. [47]	32	VR	Commercial	Heart Rate	Exposure	Pre-Recorded	–	Structured Clinical Interview for DSM-5 Disorders, PRPSA, BFNE-II, Behaviors Checklist, Eysenck Personality Questionnaire, Anxiety Sensitivity Index, Custom Scale (VR Experiences), Custom Scale (VR Valence and Arousal), Custom Scale (Speech Topic Difficulty), SUDS

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART044	2021	Pilot randomized trial of self-guided virtual reality exposure therapy for social anxiety disorder [92]	44	VR	Commercial	Voice, Gaze	Exposure	Animated, Custom Audience Features	In Presentation	SAD Symptom Severity, SIAS, Measure of Anxiety in Selection Interviews, Penn State Worry Questionnaire, Patient Health Questionnaire, Custom Scale (Post-Treatment Items), IPQ, SSQ, SUDS
ART045	2021	Real and virtual classrooms can trigger the same levels of stuttering severity ratings and anxiety in school-age children and adolescents who stutter. [93]	10	VR	Research	–	Exposure	Pre-Recorded	–	PRCS-SF, LSAS (Modified), SUDS
ART046	2021	Reflections on 21st Century Skill Development Using Interactive Posters and Virtual Reality Presentations [94]	90	PC	Commercial	–	–	Pre-Recorded	–	Custom Scale (Skill Development and Tool Perception)
ART047	2021	Technology Acceptance Model: Investigating Students' Intentions toward Adoption of Immersive 360° Videos for Public Speaking Rehearsals [95]	86	VR	Research	–	Exposure	Pre-Recorded	–	Computer-Mediated Communication Anxiety Scale (Modified), Computer Self-Efficacy Scale (Modified), Electronic Propinquity Scale, Perceived Ease of Use Scale (Modified), Perceived Usefulness Scale (Modified), Attitude Toward Technology Scale (Modified), Custom Scale (Behavioral Intentions), Custom Scale (Open-Ended Responses)

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART048	2021	The Added Benefit of an Extra Practice Session in Virtual Reality on the Development of Presentation Skills: A Randomized Control Trial [96]	35	VR	Research	–	Exposure	Animated, Reactive, Custom Audience Features	In Presentation, After Presentation	PRCA, Custom Scale (VR Utility), Custom Rubric (Oral Presentation Skills)
ART049	2021	The Effectiveness of Self-Guided Virtual-Reality Exposure Therapy for Public-Speaking Anxiety. [77]	32	VR	Research	Heart Rate	Exposure	Animated, Custom Audience Features	–	STAI, PSAS, PRCS-SF, LSAS, BFNE, SUDS
ART050	2021	Virtual Reality exposure therapy for public speaking anxiety in routine care: a single-subject effectiveness trial. [65]	23	VR	Research	–	Exposure	Animated, Custom Audience Features (Instructor Control)	–	SUDS, PSAS, LSAS-SR, BFNE, Patient Health Questionnaire, Generalized Anxiety Disorder 7-item, Brunnsviden Brief Quality of Life Scale, Negative Effects Questionnaire
ART051	2021	Virtual Reality Public Speaking Training: Experimental Evaluation of Direct Feedback Technology Acceptance [39]	200	VR	Research	Voice, Body Movement	Exposure	Animated, Reactive, Unclear	In Presentation, After Presentation	Technology Acceptance Model, PRCS
ART052	2021	Virtual reality software as preparation tools for oral presentations: Perceptions from the classroom [97]	5	VR	Commercial	–	Exposure	Pre-Recorded	In Presentation, After Presentation	Custom Scale (Instructor Surveys)
ART053	2020	Avoidance of social threat: Evidence from eye movements during a public speaking challenge using 360°-video. [98]	84	VR	Research	Gaze	Exposure	Pre-Recorded	–	LSAS-SR, Custom Scale (Demographics)
ART054	2020	Awecure VR: A solution for phobia disorder using virtual reality therapy [68]	Unclear	VR	Research	Gaze	CBT, Exposure, System Desensitization	Animated	In Presentation	Post-Study Usability Questionnaire (Modified)

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART055	2020	Continuous Interaction for a Virtual Reality Exposure Therapy System [72]	8	VR	Research	–	Exposure	Animated, Reactive, Custom Audience Features (Therapist Control)	–	ISONORM 9241-110 S, Custom Scale (Semi-Structured Interview)
ART056	2020	Cyber Trainingground: Building-Scale Virtual Reality for Immersive Presentation Training [49]	3	VR	Research	Voice, Body Movement, Gaze	Exposure	Animated, Reactive, Custom Audience Features	In Presentation	–
ART057	2020	The Effects of Virtual Audience Size on Social Anxiety during Public Speaking [26]	24	VR	Research	Heart Rate, Skin Conductance, Gaze	Exposure	Animated, Custom Audience Features	–	SAM, STAI, Custom Scale (Self-Estimated Fear), Custom Scale (State Social Anxiety), Custom Scale (Co-Presence and Social Presence), SSQ, IPQ
ART058	2020	The Impact of Computer-Mediated Immediate Feedback on Developing Oral Presentation Skills: An Exploratory Study in Virtual Reality [50]	22	VR	Research	Voice	Exposure	Animated, Unclear	In Presentation	Custom Rubric (Oral Presentation Skills), Custom Scale (Experiential Evaluation)
ART059	2020	Use of Video Blogs in Alleviating Public Speaking Anxiety among ESL Learners [23]	54	Web	Research	–	–	–	–	PRPSA
ART060	2020	Virtual reality training of presentation skills: How real does it feel? A mixed-method study [78]	46	VR	Commercial	–	Exposure	Animated, Reactive, Custom Audience Features (Instructor Control)	–	IPQ, NEO Personality Inventory
ART061	2019	Acceptance and Effectiveness of a Virtual Reality Public Speaking Training [38]	73	VR	Research	Voice, Body Movement, Gaze	Exposure	Animated, Reactive, Custom Audience Features	After Presentation	Custom Scales (Multiple Constructs)
ART062	2019	Behavioral therapy and virtual reality exposure for public speaking anxiety [67]	6	VR	Commercial	Skin Conductance	Exposure, Behavioral	Animated, Custom Audience Features	In Presentation, After Presentation	SSPS, SUDS, Custom Scale (Recording Sheet), Custom Scale (Semi-Structured Presentation), Custom Scale (Interview Script), Client Satisfaction Questionnaire

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART063	2019	Beyond Reality-Extending a Presentation Trainer with an Immersive VR Module. [33]	24	VR	Research	Voice, Body Movement, Gaze	Exposure	Static	In Presentation, After Presentation	Custom Scale (Performance Review)
ART064	2019	Cicero VR—Public Speaking Training Tool and an Attempt to Create Positive Social VR Experience [40]	36	VR	Research	Voice, Body Movement, Gaze	Exposure	Animated, Reactive, Custom Audience Features	In Presentation, After Presentation	Game Experience Questionnaire (Modified)
ART065	2019	Emotions-Responsive Audiences for VR Public Speaking Simulators Based on the Speakers' Voice [99]	19	VR	Research	Voice, Gaze	Exposure	Static, Reactive	–	–
ART066	2019	Fostering oral presentation competence through a virtual reality-based task for delivering feedback [2]	35	VR	Research	Voice, Body Movement, Gaze	Exposure	Animated, Reactive	After Presentation	Custom Scales (Multiple Choice Tests), Custom Rubric (Oral Presentation Skills), Custom Scale (Evaluation)
ART067	2019	Fostering Oral Presentation Skills by the Timing of Feedback: An Exploratory Study in Virtual Reality [51]	30	VR	Research	Voice, Gaze	Exposure	Animated, Unclear	In Presentation, After Presentation	Custom Rubric (Oral Presentation Skills)
ART068	2019	Heart Rate Prediction from Head Movement during Virtual Reality Treatment for Social Anxiety [58]	4	VR	Research	Heart Rate, Body Movement	Exposure	Animated	–	–
ART069	2019	Public speaking anxiety decreases within repeated virtual reality training sessions. [46]	21	VR	Commercial	Heart Rate	Exposure	Animated, Reactive, Custom Audience Features	–	SUDS, SIAS, PRCS, IPQ
ART070	2019	Quantle: Fair and Honest Presentation Coach in Your pocket [32]	Unclear	Mobile	Commercial	Voice	–	–	In Presentation, After Presentation	–
ART071	2019	Rich Interactions in Virtual Reality Exposure Therapy: A Pilot-Study evaluating a System for Presentation Training [71]	24	VR	Research	–	Exposure	Animated	–	SPIN (Modified), Custom Scale (Co-Presence and Social Presence), Custom Scale (Motion Sickness)

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART072	2019	Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial. [75]	50	VR	Commercial	–	Exposure	Pre-Recorded	–	PSAS, LSAS-SR, BFNE, Patient Health Questionnaire, Generalized Anxiety Disorder 7-item, Brunnsvikén Quality of Life Scale
ART073	2019	Virtual reality exposure therapy for adolescents with fear of public speaking: a non-randomized feasibility and pilot study. [73]	27	VR	Research	Heart Rate	Exposure	Animated, Reactive, Custom Audience Features (Therapist Control)	–	PSAS, SIAS, Gatinéau Presence Questionnaire
ART074	2019	Virtual reality interfaces and population-specific models to mitigate public speaking anxiety [60]	122	VR	Commercial	Heart Rate, Skin Conductance, Voice	Exposure	Unclear	–	STAI, Communication Anxiety Inventory, PRPSA, Big Five Inventory, BFNE, Reticence Willingness to Communicate, Custom Scales (Prior Daily Experiences and Demographics), Communication Anxiety Inventory (Modified), State-Anxiety Enthusiasm, BSQ, Custom Scale (Presentation Preparation Performance)
ART075	2019	Virtually Unexpected: No Role for Expectancy Violation in Virtual Reality Exposure for Public Speaking Anxiety. [100]	43	VR	Commercial	Heart Rate	Exposure	Pre-Recorded	–	PRCS, SSPS, Custom Scale (List of Expectancies), BAT, SUDS
ART076	2018	Battling the Fear of Public Speaking: Designing Software as a Service Solution for a Virtual Reality Therapy [56]	6	VR	Research	Heart Rate, Skin Conductance, Brain Sensing	Exposure	Animated, Reactive, Custom Audience Features (Therapist Control)	In Presentation, After Presentation	–

Nr.	Year	Title	Total Number of Participants	Platform	Solution Maturity	Sensing Modalities Tracked	Therapy Present	Audience Simulation	Non-Audience Related Solution Feedback and Timing	Instruments Used
ART077	2018	Improving Presentation Skill Through Gamified Application—Gamification in Practice [42]	47	Web	Research	—	—	—	—	Custom Scales (Usability Tests)
ART078	2017	Beat the Fear of Public Speaking: Mobile 360° Video Virtual Reality Exposure Training in Home Environment Reduces Public Speaking Anxiety. [24]	35	Mobile	Research	—	Exposure	Pre-Recorded, Multiple Videos Available	—	PRCA, Emotion Regulation Questionnaire, STAI
ART079	2016	A Digital Divide? Assessing Self-Perceived Communication Competency in an Online and Face-to-Face Basic Public Speaking Course [31]	2500	Web	Research	—	—	—	—	Self-Perceived Communication Competence Scale
ART080	2016	Can You Help Me with My Pitch? Studying a Tool for Real-Time Automated Feedback [22]	40	PC	Research	Voice, Body Movement	—	—	In Presentation	Custom Scale (User Experience), Custom Scale (Self-Assessment and Self-Awareness)
ART081	2015	A novel mobile application for training oral presentation delivery skills [14]	20	Mobile	Research	Voice, Body Movement	—	—	In Presentation, After Presentation	Custom Scale (Feedback)
ART082	2015	Tools and evaluation methods for discussion and presentation skills training [101]	Unclear	PC	Commercial		—	—	After Presentation	Custom Rubric (Content, Organization, Impact), Custom Rubric (Impact, Content, Organization, Presentation)

References

1. Morreale, S.P.; Osborn, M.M.; Pearson, J.C. Why Communication Is Important: A Rationale for the Centrality of the Study of Communication. *J. Assoc. Commun. Adm.* **2000**, *29*, 1–25.
2. van Ginkel, S.; Gulikers, J.; Biemans, H.; Noroozi, O.; Roozen, M.; Bos, T.; van Tilborg, R.; van Halteren, M.; Mulder, M. Fostering Oral Presentation Competence through a Virtual Reality-Based Task for Delivering Feedback. *Comput. Educ.* **2019**, *134*, 78–97. [CrossRef]
3. Bodie, G.D. A Racing Heart, Rattling Knees, and Ruminative Thoughts: Defining, Explaining, and Treating Public Speaking Anxiety. *Commun. Educ.* **2010**, *59*, 70–105. [CrossRef]
4. Spence, S.H.; Rapee, R.M. The Etiology of Social Anxiety Disorder: An Evidence-Based Model. *Behav. Res. Ther.* **2016**, *86*, 50–67. [CrossRef] [PubMed]
5. Leichsenring, F.; Leweke, F. Social Anxiety Disorder. *N. Engl. J. Med.* **2017**, *376*, 2255–2264. [CrossRef] [PubMed]
6. Grieve, R.; Woodley, J.; Hunt, S.E.; McKay, A. Student Fears of Oral Presentations and Public Speaking in Higher Education: A Qualitative Survey. *J. Furth. High. Educ.* **2021**, *45*, 1281–1293. [CrossRef]
7. Dwyer, K.K.; Davidson, M.M. Is Public Speaking Really More Feared Than Death? *Commun. Res. Rep.* **2012**, *29*, 99–107. [CrossRef]

8. Ferreira Marinho, A.C.; Mesquita De Medeiros, A.; Côrtes Gama, A.C.; Caldas Teixeira, L. Fear of Public Speaking: Perception of College Students and Correlates. *J. Voice* **2017**, *31*, 127.e7–127.e11. [[CrossRef](#)] [[PubMed](#)]
9. Kunttu, K.; Pesonen, T.; Saari, J. *Student Health Survey 2016: A National Survey among Finnish University Students*; Finnish Student Health Service: Helsinki, Finland, 2016.
10. Hofmann, S.G. Cognitive Factors That Maintain Social Anxiety Disorder: A Comprehensive Model and Its Treatment Implications. *Cogn. Behav. Ther.* **2007**, *36*, 193–209. [[CrossRef](#)] [[PubMed](#)]
11. Allen, M.; Hunter, J.E.; Donohue, W.A. Meta-analysis of Self-report Data on the Effectiveness of Public Speaking Anxiety Treatment Techniques. *Commun. Educ.* **1989**, *38*, 54–76. [[CrossRef](#)]
12. Dwyer, K.K. The Multidimensional Model: Teaching Students to Self-manage High Communication Apprehension by Self-selecting Treatments. *Commun. Educ.* **2000**, *49*, 72–81. [[CrossRef](#)]
13. Pertaub, D.-P.; Slater, M.; Barker, C. An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience. *Presence Teleoperators Virtual Environ.* **2002**, *11*, 68–78. [[CrossRef](#)]
14. Lui, A.K.-F.; Ng, S.-C.; Wong, W.-W. A Novel Mobile Application for Training Oral Presentation Delivery Skills. *Commun. Comput. Inf. Sci.* **2015**, *559*, 79–89. [[CrossRef](#)]
15. Anderson, P.L.; Price, M.; Edwards, S.M.; Obasaju, M.A.; Schmertz, S.K.; Zimand, E.; Calamaras, M.R. Virtual Reality Exposure Therapy for Social Anxiety Disorder: A Randomized Controlled Trial. *J. Consult. Clin. Psychol.* **2013**, *81*, 751–760. [[CrossRef](#)] [[PubMed](#)]
16. Carl, E.; Stein, A.T.; Leivhn-Coon, A.; Pogue, J.R.; Rothbaum, B.; Emmelkamp, P.; Asmundson, G.J.G.; Carlbring, P.; Powers, M.B. Virtual Reality Exposure Therapy for Anxiety and Related Disorders: A Meta-Analysis of Randomized Controlled Trials. *J. Anxiety Disord.* **2019**, *61*, 27–36. [[CrossRef](#)] [[PubMed](#)]
17. Reeves, R.; Elliott, A.; Curran, D.; Dyer, K.; Hanna, D. 360° Video Virtual Reality Exposure Therapy for Public Speaking Anxiety: A Randomized Controlled Trial. *J. Anxiety Disord.* **2021**, *83*, 102451. [[CrossRef](#)] [[PubMed](#)]
18. Kahlon, S.; Lindner, P.; Nordgreen, T. Gamified Virtual Reality Exposure Therapy for Adolescents with Public Speaking Anxiety: A Four-Armed Randomized Controlled Trial. *Front. Virtual Real.* **2023**, *4*, 1240778. [[CrossRef](#)]
19. Ebrahimi, O.V.; Pallesen, S.; Kenter, R.M.F.; Nordgreen, T. Psychological Interventions for the Fear of Public Speaking: A Meta-Analysis. *Front. Psychol.* **2019**, *10*, 488. [[CrossRef](#)] [[PubMed](#)]
20. Esfandiari, N.; Mazaheri, M.A.; Akbari-Zardkhaneh, S.; Sadeghi-Firoozabadi, V.; Cheraghi, M. Internet-Delivered versus Face-to-Face Cognitive Behavior Therapy for Anxiety Disorders: Systematic Review and Meta-Analysis. *Int. J. Prev. Med.* **2021**, *12*, 153. [[CrossRef](#)] [[PubMed](#)]
21. Guo, S.; Deng, W.; Wang, H.; Liu, J.; Liu, X.; Yang, X.; He, C.; Zhang, Q.; Liu, B.; Dong, X.; et al. The Efficacy of Internet-based Cognitive Behavioural Therapy for Social Anxiety Disorder: A Systematic Review and Meta-analysis. *Clin. Psychol. Psychother.* **2021**, *28*, 656–668. [[CrossRef](#)] [[PubMed](#)]
22. Schneider, J.; Börner, D.; van Rosmalen, P.; Specht, M. Can You Help Me with My Pitch? Studying a Tool for Real-Time Automated Feedback. *IEEE Trans. Learn. Technol.* **2016**, *9*, 318–327. [[CrossRef](#)]
23. Madzlan, N.A.; Seng, G.H.; Kesevan, H.V. Use of Video Blogs in Alleviating Public Speaking Anxiety among ESL Learners. *J. Educ. E-Learn. Res.* **2020**, *7*, 93–99. [[CrossRef](#)]
24. Stupar-Rutenfrans, S.; Ketelaars, L.E.H.; van Gisbergen, M.S. Beat the Fear of Public Speaking: Mobile 360° Video Virtual Reality Exposure Training in Home Environment Reduces Public Speaking Anxiety. *Cyberpsychology Behav. Soc. Netw.* **2017**, *20*, 624–633. [[CrossRef](#)] [[PubMed](#)]
25. Kryston, K.; Goble, H.; Eden, A. Incorporating Virtual Reality Training in an Introductory Public Speaking Course. *J. Commun. Pedagog.* **2021**, *4*, 133–151. [[CrossRef](#)]
26. Mostajeran, F.; Balci, M.B.; Steinicke, F.; Kühn, S.; Gallinat, J. The Effects of Virtual Audience Size on Social Anxiety during Public Speaking. In *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Atlanta, GA, USA, 22–26 March 2020; IEEE: New York, NY, USA, 2020; pp. 303–312.
27. Premkumar, P.; Heym, N.; Myers, J.A.C.; Formby, P.; Battersby, S.; Sumich, A.L.; Brown, D.J. Augmenting Self-Guided Virtual-Reality Exposure Therapy for Social Anxiety with Biofeedback: A Randomised Controlled Trial. *Front. Psychiatry* **2024**, *15*, 1467141. [[CrossRef](#)] [[PubMed](#)]
28. Min, Y.; Jeong, J.-W. Public Speaking Q&A Practice with LLM-Generated Personas in Virtual Reality. In *Proceedings of the 2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Bellevue, WA, USA, 21–25 October 2024; IEEE: New York, NY, USA, 2024; pp. 493–496.
29. Bartyzel, P.; Lgras-Cybulska, M.; Tadeja, S.K.; Majdak, M.; Łukawski, G.; Hekiert, D. Voice-Responsive Virtual Reality Training Environment for Occupational and Non-Professional Voice Users. In *Proceedings of the 2024 IEEE Conference Virtual Real 3D User Interfaces Abstr. Workshop VRW*, Orlando, FL, USA, 16–21 March 2024; IEEE: New York, NY, USA, 2024; pp. 208–213.
30. Hosseini, S.; Quan, J.; Deng, X.; Miyake, Y.; Nozawa, T. Avatar-Based Feedback in Job Interview Training Impacts Action Identities and Anxiety. *IEEE Trans. Affect. Comput.* **2024**, *15*, 1608–1620. [[CrossRef](#)]

31. Westwick, J.N.; Hunter, K.M.; Haleta, L.L. A Digital Divide? Assessing Self-Perceived Communication Competency in an Online and Face-to-Face Basic Public Speaking Course. *Basic Commun. Course Annu.* **2016**, *28*, 48–86.
32. Saukh, O.; Maag, B. Quantle: Fair and Honest Presentation Coach in Your Pocket. In Proceedings of the 2019 18th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), Montreal, QC, Canada, 16–18 April 2019; pp. 253–264.
33. Schneider, J.; Romano, G.; Drachsler, H. Beyond Reality-Extending a Presentation Trainer with an Immersive VR Module. *Sensors* **2019**, *19*, 3457. [[CrossRef](#)]
34. Ferreira, L.; Cerqueira, J.; Jonassi, J.; Jesus, A.; Amaral, C.; Mateus-Coelho, N. Virtual Reality on Public Speaking Phobia Mitigation. *Procedia Comput. Sci.* **2024**, *239*, 2251–2259. [[CrossRef](#)]
35. Barrett, A.; Pack, A.; Monteiro, D.; Liang, H.-N. Exploring the Influence of Audience Familiarity on Speaker Anxiety and Performance in Virtual Reality and Real-Life Presentation Contexts. *Behav. Inf. Technol.* **2024**, *43*, 787–799.
36. Rodero, E.; Larrea, O. Virtual Reality with Distractors to Overcome Public Speaking Anxiety in University Students. *Commun. Media Educ. Res. J.* **2022**, *30*, 85–96. [[CrossRef](#)]
37. Truong, B.; Le, T.-N.; Le, K.-D.; Tran, M.-T.; Nguyen, T.V. Public Speaking Simulator with Speech and Audience Feedback. In Proceedings of the 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Singapore, 17–21 October 2022; IEEE Comp Soc; IEEE VGTC; ACM SIGGRAPH; Zoom; Qualcomm; Nvidia; Oppo; Advent2 Labs Consultat; Hiverlab; IEEE: New York, NY, USA, 2022; pp. 855–858.
38. Palmas, F.; Cichor, J.; Plecher, D.A.; Klinker, G. Acceptance and Effectiveness of a Virtual Reality Public Speaking Training. In Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Beijing, China, 14–18 October 2019; IEEE: New York, NY, USA, 2019; pp. 363–371.
39. Palmas, F.; Reinelt, R.; Cichor, J.E.; Plecher, D.A.; Klinker, G. Virtual Reality Public Speaking Training: Experimental Evaluation of Direct Feedback Technology Acceptance. In Proceedings of the 2021 IEEE Virtual Reality and 3D User Interfaces (VR), Lisbon, Portugal, 27 March–1 April 2021; IEEE: New York, NY, USA, 2021; pp. 463–472.
40. Jakubowski, M.; Wardaszko, M.; Winniczuk, A.; Podgorski, B.; Cwil, M. Cicero VR—Public Speaking Training Tool and an Attempt to Create Positive Social VR Experience. In Proceedings of the Virtual, Augmented and Mixed Reality: Applications and Case Studies; Chen, J., Fragomeni, G., Eds.; Springer: Cham, Switzerland, 2019; Volume 11575, pp. 297–311.
41. Monteiro, D.; Wang, A.; Wang, L.; Li, H.; Barrett, A.; Pack, A.; Liang, H.-N. Effects of Audience Familiarity on Anxiety in a Virtual Reality Public Speaking Training Tool. *Univers. Access Inf. Soc.* **2024**, *23*, 23–34. [[CrossRef](#)]
42. Leinonen, E.; Haapaniemi, M.; Mattila, J.; Firouzian, A.; Pulli, P. Improving Presentation Skill Through Gamified Application—Gamification in Practice. In Proceedings of the 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA), Košice, Slovakia, 23–25 August 2018; pp. 95–100.
43. Colognesi, S.; Coppe, T.; Lucchini, S. Improving the Oral Language Skills of Elementary School Students through Video-Recorded Performances. *Teach. Teach. Educ.* **2023**, *128*, 104141. [[CrossRef](#)]
44. Kroczeck, L.O.H.; Mühlberger, A. Public Speaking Training in Front of a Supportive Audience in Virtual Reality Improves Performance in Real-Life. *Sci. Rep.* **2023**, *13*, 13968. [[CrossRef](#)] [[PubMed](#)]
45. Macey, A.-L.; Järvelä, S.; Fernández Galeote, D.; Hamari, J. Feeling Small or Standing Tall? Height Manipulation Affects Speech Anxiety and Arousal in Virtual Reality. *Cyberpsychology Behav. Soc. Netw.* **2023**, *26*, 246–254. [[CrossRef](#)] [[PubMed](#)]
46. Takac, M.; Collett, J.; Blom, K.J.; Conduit, R.; Rehm, I.; De Foe, A. Public Speaking Anxiety Decreases within Repeated Virtual Reality Training Sessions. *PLoS ONE* **2019**, *14*, e0216288. [[CrossRef](#)] [[PubMed](#)]
47. van Dis, E.A.M.; Landkroon, E.; Hagenaars, M.A.; van der Does, F.H.S.; Engelhard, I.M. Old Fears Die Hard: Return of Public Speaking Fear in a Virtual Reality Procedure. *Behav. Ther.* **2021**, *52*, 1188–1197. [[CrossRef](#)] [[PubMed](#)]
48. Kuai, S.-G.; Liang, Q.; He, Y.-Y.; Wu, H.-N. Higher Anxiety Rating Does Not Mean Poor Speech Performance: Dissociation of the Neural Mechanisms of Anticipation and Delivery of Public Speaking. *Brain Imaging Behav.* **2021**, *15*, 1934–1943. [[CrossRef](#)] [[PubMed](#)]
49. Nagao, K.; Yokoyama, Y. Cyber Trainground: Building-Scale Virtual Reality for Immersive Presentation Training. In Proceedings of the 2020 IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), Calgary, AB, Canada, 17–24 August 2020; pp. 192–200.
50. Van Ginkel, S.; Ruiz, D.; Mononen, A.; Karaman, C.; de Keijzer, A.; Sitthiworachart, J. The Impact of Computer-Mediated Immediate Feedback on Developing Oral Presentation Skills: An Exploratory Study in Virtual Reality. *J. Comput. Assist. Learn.* **2020**, *36*, 412–422. [[CrossRef](#)]
51. Belboukhaddaoui, I.; van Ginkel, S. Fostering Oral Presentation Skills by the Timing of Feedback: An Exploratory Study in Virtual Reality. *Res. Educ. Media* **2019**, *11*, 25–31. [[CrossRef](#)]

52. Tangsripiroj, S.; Nunthapatpokin, N.; Saelim, P.; Savittrakul, C. Speech Lab VR: A Virtual Reality System for Improving Presentation Skills. In *Proceedings of the 2024 8th International Conference on Information Technology (InCIT), Chonburi, Thailand, 14–15 November 2024*; IEEE: New York, NY, USA, 2024; pp. 295–300.
53. Ruiz-Capillas, M.E.; Romero-Hernandez, A.; Manero, B. Talk & Play: A VR Application to Improve Volume and Silences Use in Verbal Communication. In *Proceedings of the 30th International Conference on Computers in Education (ICCE 2022), Kuala Lumpur, Malaysia, 28 November–2 December 2022*; Iyer, S., Shih, J., Chen, W., Md Khambari, M., Eds.; Asia-Pacific Society for Computers in Education (APSCE): Kuala Lumpur, Malaysia, 2022; pp. 505–510.
54. Huang, M. Empowering Oral Proficiency in a Large-Scale Class: Video-Recorded Oral Presentations and Mobile-Assisted Peer Assessment in a Chinese Middle School. *Comput. Assist. Lang. Learn.* **2024**, *39*, 395–415. [\[CrossRef\]](#)
55. Glemarec, Y.; Lugin, J.-L.; Bosser, A.-G.; Jackson, A.C.; Buche, C.; Latoschik, M.E. Indifferent or Enthusiastic? Virtual Audiences Animation and Perception in Virtual Reality. *Front. Virtual Real.* **2021**, *2*, 666232. [\[CrossRef\]](#)
56. Salkevicius, J.; Navickas, L. Battling the Fear of Public Speaking: Designing Software as a Service Solution for a Virtual Reality Therapy. In *Proceedings of the 2018 6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), Barcelona, Spain, 6–8 August 2018*; IEEE: New York, NY, USA, 2018; pp. 209–213.
57. Niebuhr, O.; Valls-Ratés, I. Gestures and Feet Can Tell Us How You Speak! On the Relationships between Voice and Body Language in VR Public Speeches. In *Proceedings of the 2024 IEEE International Professional Communication Conference (ProComm), Pittsburgh, PA, USA, 14–17 July 2024*; IEEE: New York, NY, USA, 2024; pp. 96–104.
58. Noori, F.M.; Kahlon, S.; Lindner, P.; Nordgreen, T.; Torresen, J.; Riegler, M. Heart Rate Prediction from Head Movement during Virtual Reality Treatment for Social Anxiety. In *Proceedings of the 2019 International Conference on Content-Based Multimedia Indexing (CBMI), Dublin, Ireland, 4–6 September 2019*; IEEE: New York, NY, USA, 2019; pp. 1–5.
59. Sülter, R.E.; Ketelaar, P.E.; Lange, W.-G. SpeakApp-Kids! Virtual Reality Training to Reduce Fear of Public Speaking in Children—A Proof of Concept. *Comput. Educ.* **2022**, *178*, 104384. [\[CrossRef\]](#)
60. Yadav, M.; Sakib, M.N.; Feng, K.; Chaspari, T.; Behzadan, A. Virtual Reality Interfaces and Population-Specific Models to Mitigate Public Speaking Anxiety. In *Proceedings of the 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), Cambridge, UK, 3–6 September 2019*; IEEE: New York, NY, USA, 2019; pp. 1–7.
61. LeFebvre, L.; LeFebvre, L.E.; Allen, M. Exploring Eye Contact in Virtual Environments: The Compositor Mirror Tool, Areas of Interest, and Public Speaking Competency. *Commun. Stud.* **2021**, *72*, 1053–1072. [\[CrossRef\]](#)
62. Wechsler, T.F.; Pfaller, M.; van Eickels, R.E.; Schulz, L.H.; Mühlberger, A. Look at the Audience? A Randomized Controlled Study of Shifting Attention from Self-Focus to Nonsocial vs. Social External Stimuli During Virtual Reality Exposure to Public Speaking in Social Anxiety. *Front. Psychiatry* **2021**, *12*, 751272. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Moldoveanu, A.; Mitrut, O.; Jinga, N.; Petrescu, C.; Moldoveanu, F.; Asavei, V.; Anghel, A.M.; Petrescu, L. Immersive Phobia Therapy through Adaptive Virtual Reality and Biofeedback. *Appl. Sci.* **2023**, *13*, 10365. [\[CrossRef\]](#)
64. Valls-Ratés, I.; Niebuhr, O.; Prieto, P. Unguided Virtual-Reality Training Can Enhance the Oral Presentation Skills of High-School Students. *Front. Commun.* **2022**, *7*, 910952. [\[CrossRef\]](#)
65. Lindner, P.; Dagö, J.; Hamilton, W.; Miloff, A.; Andersson, G.; Schill, A.; Carlbring, P. Virtual Reality Exposure Therapy for Public Speaking Anxiety in Routine Care: A Single-Subject Effectiveness Trial. *Cogn. Behav. Ther.* **2021**, *50*, 67–87. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Macdonald, C. Improving Virtual Reality Exposure Therapy with Open Access and Overexposure: A Single 30-Minute Session of Overexposure Therapy Reduces Public Speaking Anxiety. *Front. Virtual Real.* **2024**, *5*, 1506938. [\[CrossRef\]](#)
67. Zacarin, M.R.J.; Borloti, E.; Haydu, V.B. Behavioral Therapy and Virtual Reality Exposure for Public Speaking Anxiety. *Trends Psychol.* **2019**, *27*, 491–507. [\[CrossRef\]](#)
68. Hanif, K.; Bawany, N.; Siddiq, K.; Shareef, H.; Rizwan, M.; Amir, R. Awecure VR: A Solution for Phobia Disorder Using Virtual Reality Therapy. *Suranaree J. Sci. Technol.* **2020**, *27*, 30032–30033.
69. Hayes, S.C. Acceptance and Commitment Therapy, Relational Frame Theory, and the Third Wave of Behavioral and Cognitive Therapies. *Behav. Ther.* **2004**, *35*, 639–665. [\[CrossRef\]](#)
70. Gorinelli, S.; Gallego, A.; Lappalainen, P.; Lappalainen, R. Virtual Reality Acceptance and Commitment Therapy Intervention for Social and Public Speaking Anxiety: A Randomized Controlled Trial. *J. Context. Behav. Sci.* **2023**, *28*, 289–299. [\[CrossRef\]](#)
71. Koller, M.; Schäfer, P.; Lochner, D.; Meixner, G. Rich Interactions in Virtual Reality Exposure Therapy: A Pilot-Study Evaluating a System for Presentation Training. In *Proceedings of the 2019 IEEE International Conference on Healthcare Informatics (ICHI), Xi'an, China, 10–13 June 2019*; pp. 1–11.
72. Koller, M.; Rauh, S.F.; Lundström, A.; Bogdan, C.; Meixner, G. Continuous Interaction for a Virtual Reality Exposure Therapy System. In *Proceedings of the 2020 IEEE International Conference on Healthcare Informatics (ICHI), Oldenburg, Germany, 30 November–3 December 2020*; IEEE: New York, NY, USA, 2021; pp. 1–11.
73. Kahlon, S.; Lindner, P.; Nordgreen, T. Virtual Reality Exposure Therapy for Adolescents with Fear of Public Speaking: A Non-Randomized Feasibility and Pilot Study. *Child Adolesc. Psychiatry Ment. Health* **2019**, *13*, 47. [\[CrossRef\]](#) [\[PubMed\]](#)

74. Banakou, D.; Johnston, T.; Beacco, A.; Senel, G.; Slater, M. Desensitizing Anxiety Through Imperceptible Change: Feasibility Study on a Paradigm for Single-Session Exposure Therapy for Fear of Public Speaking. *JMIR Form. Res.* **2024**, *8*, e52212. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Lindner, P.; Miloff, A.; Fagnäs, S.; Andersen, J.; Sigeman, M.; Andersson, G.; Furmark, T.; Carlbring, P. Therapist-Led and Self-Led One-Session Virtual Reality Exposure Therapy for Public Speaking Anxiety with Consumer Hardware and Software: A Randomized Controlled Trial. *J. Anxiety Disord.* **2019**, *61*, 45–54. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Sarpourian, F.; Samad-Soltani, T.; Moulaei, K.; Bahaadinbeigy, K. The Effect of Virtual Reality Therapy and Counseling on Students' Public Speaking Anxiety. *Health Sci. Rep.* **2022**, *5*, e816. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Premkumar, P.; Heym, N.; Brown, D.J.; Battersby, S.; Sumich, A.; Huntington, B.; Daly, R.; Zysk, E. The Effectiveness of Self-Guided Virtual-Reality Exposure Therapy for Public-Speaking Anxiety. *Front. Psychiatry* **2021**, *12*, 694610. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Kollöffel, B.; Heuvel, K.O. Virtual Reality Training of Presentation Skills: How Real Does It Feel? A Mixed-Method Study. In *Proceedings of the 48th SEFI Annual Conference on Engineering Education, SEFI 2020, Enschede, The Netherlands, 20–24 September 2020*; Societe Europeenne pour la Formation des Ingenieurs: Brussels, Belgium, 2020; pp. 241–250.
79. Landkroon, E.; van Dis, E.A.M.; Meyerbröker, K.; Saleminck, E.; Hagenaars, M.A.; Engelhard, I.M. Future-Oriented Positive Mental Imagery Reduces Anxiety for Exposure to Public Speaking. *Behav. Ther.* **2022**, *53*, 80–91. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Lestiono, R.; Setyaningrum, R.W. Developing Immersive Virtual Reality Space for Public Speaking Training: English Debate Room in Spatial Platform. In *Proceedings of the 2025 19th International Conference on Ubiquitous Information Management and Communication (IMCOM), Bangkok, Thailand, 3–5 January 2025*; IEEE: New York, NY, USA, 2025; pp. 1–5.
81. Bartyzel, P.; Igras-Cybulska, M.; Hekiart, D.; Majdak, M.; Lukawski, G.; Bohne, T.; Tadeja, S. Exploring User Reception of Speech-Controlled Virtual Reality Environment for Voice and Public Speaking Training. *Comput. Graph.* **2025**, *126*, 104160. [\[CrossRef\]](#)
82. van Veen, S.C.; Zbozinek, T.D.; van Dis, E.A.M.; Engelhard, I.M.; Craske, M.G. Positive Mood Induction Does Not Reduce Return of Fear: A Virtual Reality Exposure Study for Public Speaking Anxiety. *Behav. Res. Ther.* **2024**, *174*, 104490. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Asik, A.; Sert, O.; Miller, P. The Affordances of a Mobile Video-Tagging Tool for Evaluating Presentation Skills in a Second Language. *Reflective Pract.* **2024**, *25*, 145–163. [\[CrossRef\]](#)
84. Valls-Ratés, I.; Niebuhr, O.; Prieto, P. VR Public Speaking Simulations Can Make Voices Stronger and More Effortful. In *Proceedings of the Proceedings of the 16th International Conference on Computer Supported Education, CSEDU 2024, Angers, France, 2–4 May 2024*; SCITEPRESS Digital Library: Setúbal, Portugal, 2024; Volume 1, pp. 685–693.
85. Robillos, R.J. Effect of Metacognitive-Based Digital Graphic Organizer on Learners' Oral Presentation Skill and Self-Regulation of Learning Awareness. *Int. J. Instr.* **2023**, *16*, 639–654. [\[CrossRef\]](#)
86. Valls-Rates, I.; Niebuhr, O.; Prieto, P. Encouraging Participant Embodiment during VR-Assisted Public Speaking Training Improves Persuasiveness and Charisma and Reduces Anxiety in Secondary School Students. *Front. Virtual Real.* **2023**, *4*, 1074062. [\[CrossRef\]](#)
87. Huang, H.-W.; Cai, S.; Li, Y.; Dusza, D.G. Using 360 VR Videos in Presentation Rehearsals: Comparing Anxiety Levels. In *Proceedings of the 2023 5th International Conference on Computer Science and Technologies in Education (CSTE), Xi'an, China, 21–23 April 2023*; IEEE: New York, NY, USA, 2023; pp. 237–241.
88. Glémarec, Y.; Lugrin, J.-L.; Bosser, A.-G.; Buche, C.; Latoschik, M.E. Controlling the Stage: A High-Level Control System for Virtual Audiences in Virtual Reality. *Front. Virtual Real.* **2022**, *3*, 876433. [\[CrossRef\]](#)
89. Yadav, M.; Sakib, M.N.; Nirjhar, E.H.; Feng, K.; Behzadan, A.H.; Chaspari, T. Exploring Individual Differences of Public Speaking Anxiety in Real-Life and Virtual Presentations. *IEEE Trans. Affect. Comput.* **2022**, *13*, 1168–1182. [\[CrossRef\]](#)
90. Seiler, R.; Coviello, R. Using Virtual Reality Simulation to Reduce Stage Fright During Public Appearances. In *Proceedings of the International Conference on e-Learning (EL 2022), Lisbon, Portugal, 12–15 July 2022*; pp. 45–51.
91. LeFebvre, L.E.; LeFebvre, L.; Allen, M. "Imagine All the People": Imagined Interactions in Virtual Reality When Public Speaking. *Imagin. Cogn. Personal.* **2021**, *40*, 189–222.
92. Zainal, N.H.; Chan, W.W.; Saxena, A.P.; Taylor, C.B.; Newman, M.G. Pilot Randomized Trial of Self-Guided Virtual Reality Exposure Therapy for Social Anxiety Disorder. *Behav. Res. Ther.* **2021**, *147*, 103984. [\[CrossRef\]](#)
93. Moïse-Richard, A.; Ménard, L.; Bouchard, S.; Leclercq, A.-L. Real and Virtual Classrooms Can Trigger the Same Levels of Stuttering Severity Ratings and Anxiety in School-Age Children and Adolescents Who Stutter. *J. Fluency Disord.* **2021**, *68*, 105830. [\[CrossRef\]](#) [\[PubMed\]](#)
94. Warrick, A.; Woodward, H. Reflections on 21st Century Skill Development Using Interactive Posters and Virtual Reality Presentations. *EuroCall* **2021**, *2021*, 290. [\[CrossRef\]](#)
95. Vallade, J.I.; Kaufmann, R.; Frisby, B.N.; Martin, J.C. Technology Acceptance Model: Investigating Students' Intentions toward Adoption of Immersive 360° Videos for Public Speaking Rehearsals. *Commun. Educ.* **2021**, *70*, 127–145.
96. Boetje, J.; van Ginkel, S. The Added Benefit of an Extra Practice Session in Virtual Reality on the Development of Presentation Skills: A Randomized Control Trial. *J. Comput. Assist. Learn.* **2021**, *37*, 253–264.

97. Alsaffar, M.J. Virtual Reality Software as Preparation Tools for Oral Presentations: Perceptions from the Classroom. *Theory Pract. Lang. Stud.* **2021**, *11*, 1146–1160. [[CrossRef](#)]
98. Rubin, M.; Minns, S.; Muller, K.; Tong, M.H.; Hayhoe, M.M.; Telch, M.J. Avoidance of Social Threat: Evidence from Eye Movements during a Public Speaking Challenge Using 360°-Video. *Behav. Res. Ther.* **2020**, *134*, 103706. [[CrossRef](#)] [[PubMed](#)]
99. El-Yamri, M.; Romero-Hernandez, A.; Gonzalez-Riojo, M.; Manero, B. Emotions-Responsive Audiences for VR Public Speaking Simulators Based on the Speakers' Voice. In *Proceedings of the 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, Maceio, Brazil, 15–18 July 2019; IEEE: New York, NY, USA, 2019; Volume 2161–377X, pp. 349–353.
100. Scheveneels, S.; Boddez, Y.; Van Daele, T.; Hermans, D. Virtually Unexpected: No Role for Expectancy Violation in Virtual Reality Exposure for Public Speaking Anxiety. *Front. Psychol.* **2019**, *10*, 2849. [[CrossRef](#)] [[PubMed](#)]
101. Nagao, K.; Tehrani, M.P.; Fajardo, J.T.B. Tools and Evaluation Methods for Discussion and Presentation Skills Training. *Smart Learn. Environ.* **2015**, *2*, 5. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.