IMPACT OF VIRTUAL REALITY ROLE-PLAY ON PUBLIC SPEAKING SKILLS: A DUAL-PATH MODEL OF IMMERSION AND EMOTIONAL REGULATION

Jiaqi Jiang[1*], Tan Wee Hoe[2]

[1]UCSI University, Kuala Lumpur, Malaysia. ORCID: https://orcid.org/1002060086@ucsiuniversity.edu.my[1],
TanWH@ucsiuniversity.edu.my[2]

ABSTRACT

Public speaking anxiety (PSA) is a common communication problem faced by students and professionals that can have a negative impact on performance, even after repeated exposure (Bodie, 2010). With recent innovations in educational technologies, virtual reality (VR) platforms have begun to be explored as a new method of experiential learning. In the current study, we examine the effectiveness of VR-based role-play interventions for the enhancement of public speaking performance via a dual-path model including immersion and emotional regulation. In a quasi-experimental design, N = 120 participants completed four sessions of VR speech simulations. The Public Speaking Competence Scale (PSCS), the Immersive Tendencies Questionnaire (ITQ), and the Emotion Regulation Questionnaire (ERQ) were administered as pre- and post-tests. Structural Equation Modeling (SEM) was conducted and found significant mediating effects for both immersion and emotional regulation on the relationship between VR simulation participation and performance enhancement (p < .01). The results of this study provide evidence that immersive VR simulation exercises can increase presence and self-confidence, while pre-designed emotional regulation techniques can decrease PSA and allow for greater cognitive control. Theoretical and pedagogical implications of these findings are discussed for future VR-based communication skill training and the potential of such interventions to be implemented as curricular components, with recommendations for future studies on long-term retention and cross-cultural adaptability.

KEYWORDS: Virtual Reality, Public Speaking, Immersion, Emotional Regulation, Communication Skills, VR Role-Play.

1. INTRODUCTION

1.1. Background & Rationale

Public speaking is a form of communication that is essential to our social, educational, and professional lives. Yet, it is a common experience that up to 70% of the general population reports to suffer from some degree of public speaking anxiety (PSA), often taking the form of psychophysiological arousal, cognitive rumination, and avoidance (Bodie, 2010). Poor public speaking skills are associated with negative outcomes among students, such as decreased performance during class presentations, low self-esteem, social withdrawal, and poor grades (Ferreira Marinho et al., 2017). In the workplace, poor public speaking skills may lead to fewer leadership opportunities and reduced career development (MacIntyre & Thivierge, 1995).

Formal public speaking training has often relied on traditional techniques such as receiving feedback from peers, videotaping, and practicing in-class presentations. Unfortunately, such training often lack the ability to mimic the stress of real life or to provide a consistent and low-risk opportunity for frequent practice, leaving high PSA speakers with few opportunities to manage anxiety during public speaking. Psychological interventions such as cognitive-behavioral training (CBT) and systematic desensitization have been effective in reducing PSA, but these take many hours of therapist-guided instruction and are not feasible on a large scale in classrooms (Allen et al., 2001).

Virtual Reality (VR) technology allows for novel experiential learning and anxiety-reduction opportunities. Virtual role-play using VR allows a user to interact with others in a realistic, three-dimensional world that can recreate a user's felt pressure to deliver a speech without the real-life consequences of doing so (Slater & Sanchez-Vives, 2016). In addition, immersive VR training may be particularly useful for high anxiety learners, as it can allow for repeated opportunities for practicing performances and receiving feedback in the moment, which can support regulation of the cognitive and emotional components of public speaking anxiety (Makransky et al., 2019). In sum, VR technology has become an increasingly attractive tool in communication classes, clinical settings, and corporate training for scalable public speaking training to increase individual performance.

Yet, the psychological processes that drive the benefits of VR public speaking have not been sufficiently explored. In particular, while users experience emotions and thoughts in the virtual world, the connection to real-life performance and behavior change remains largely unknown.

1.2. Research Gaps

While the evidence for the effects of VR on communication training is rapidly increasing, research has not yet reached a consensus on how and why virtual reality works for public speaking. A major methodological limitation of prior studies is their reliance on outcome-based measures (e.g., self-reported confidence or speech fluency) as opposed to more process-oriented investigations that could target and test the proposed mediating mechanisms (Geraets et al., 2021). Although numerous studies document the beneficial effects of repeated VR exposure on PSA, the actual mechanisms of action that underlie the observed performance improvements remain unclear. For example, we know that VR-based exposure can be an effective way to reduce PSA. However, we still do not know whether it works by increasing desensitization, self-regulation, immersion (in the sense of task focus), or some combination of the three (or more) mechanisms above.

Surprisingly, few studies have examined immersion and emotional regulation simultaneously. Immersion, a complex multidimensional construct, is generally described as the subjective experience of "being there" within a virtual environment (Dede, 2009). In a number of VR-based learning environments, immersion was found to be the main predictor of the training's success. The ability to concentrate on a given training program, which is greatly improved by high immersion, reduces extraneous load and paves the way for optimal learning (Dede, 2009). The construct of emotional regulation (or regulation of emotions), by contrast, refers to a person's ability to monitor, evaluate, and modulate his or her emotional responses in accordance with the requirements of a given situation (Gross, 2015). A growing body of evidence suggests that regulation of emotions is a critical factor in overcoming PSA and maximizing cognitive resources when faced with a performance task (Gross, 2015). As such, it is a central pillar of cognitive behavioral therapies for public speaking anxiety. The research on these two critical factors in the training and public speaking literature has developed in relative isolation. However, their theorized effects are not mutually exclusive and, in fact, are very likely to interact in an additive or even synergistic manner.

Furthermore, most studies exploring VR applications in public speaking are descriptive/exploratory in design and fail to provide

a theoretical model that could map specific psychological constructs and processes (e.g., immersion, emotion regulation) to performance and outcome measures. This lack of a guiding theoretical framework leads to interventions that are poorly targeted and often unrepeatable and makes it hard to identify characteristics that could predict users' responsiveness to such interventions. As a result, VR environments are treated as a sort of black box, and we do not yet understand which user experiences and behaviors (e.g., deep immersion) and which psychological processes (e.g., emotion regulation) in these environments will impact public speaking behavior.

The present study is an attempt to fill this research gap by constructing and testing a dual-path model of VR-based role-play on public speaking performance. This study will theoretically contribute to the two converging lines of research, in this case, the study of immersion and emotion regulation in virtual reality. At the same time, the proposed model has the potential to directly impact VR-based pedagogy by identifying mediating factors that could be targeted to maximize training efficiency and scalability.

1.3. Research Questions and Objectives

The study aims to investigate the underlying processes of the effectiveness of virtual reality role-play on public speaking performance. The proposed dual-path conceptual model of the study is based on a hypothesis that immersion and emotion regulation serve as parallel mediators between VR engagement and public speaking performance. Immersion is a sense of being in the virtual situation or performance experience that may increase focus and reduce distractions, while emotional regulation is the cognition and behavior that individuals practice when experiencing the arousal and anxiety related to the performance (Hofmann et al., 2012). These two concepts are of particular importance for learners who experience high levels of anxiety.

Given this framework, **the following research questions guide the investigation**

RQ1: How does VR-based role-play influence public speaking performance?

This question aims to evaluate whether participation in immersive VR speech simulations leads to measurable improvements in public speaking skills, including speech organization, vocal delivery, and non-verbal behavior.

RQ2: What roles do immersion and emotional regulation play in this process?

This question investigates whether and to what extent immersion and emotional regulation mediate

the effect of VR role-play on public speaking outcomes. Structural Equation Modeling (SEM) will be used to assess the strength and significance of each pathway.

To address these questions, the study uses a quasi-experimental design with pre- and post-tests, including validated self-report instruments and behavioral performance ratings. The findings are expected to contribute to the theoretical literature on VR-based learning and to provide practical guidance for educators and trainers designing communication skills interventions.

2. THEORETICAL FRAMEWORK

2.1. Dual-Path Model Overview

The purpose of this study is to explain why virtual reality (VR) role-play facilitates public speaking performance using a dual-path conceptual model. The proposed model is based on two psychological mechanisms: immersion and emotional regulation. These two constructs are theoretically and empirically linked to public speaking and are conceptualized as parallel mediators of the effect of VR on speech quality, confidence, and anxiety. These mechanisms need to be understood if we want to know more about what makes VR communication training more than just exposure to audiences and anxiety-provoking situations.

Immersion is defined as the feeling of cognitive and affective engagement in a virtual environment (Slater & Wilbur, 1997). It includes factors related to attentional focus, emotional investment, and presence (the sense of the virtual reality being real). In the context of educational VR, there is evidence that immersion leads to better learning outcomes by enhancing intrinsic motivation and minimizing distractions (Makransky & Petersen, 2021). In the context of public speaking, immersion in VR simulations can facilitate the user's willingness to "suspend disbelief" and thus trigger more natural reactions and lower self-consciousness. The more the learner feels immersed, the more they will naturally respond to audience feedback, practice realistic delivery, and gain transferable performance skills.

Emotional regulation, on the other hand, refers to the processes by which people "monitor, evaluate, and modify emotional reactions" to reach their goals (Gross, 2015, p 272). Public speaking is stressful by nature and can often induce high levels of anxiety that negatively affect cognitive processing and verbal fluency (Bodie, 2010). Emotion regulation strategies, such as reappraisal, breathing techniques, or attentional distraction, can help mitigate this effect by reducing physiological arousal and increasing

composure. In VR settings, emotion regulation becomes an even more relevant and testable process, as learners can repeatedly be exposed to anxiety-inducing situations in a safe environment and develop the psychological flexibility to manage stress in the moment (Mottelson et al., 2021).

The dual-path model assumes that immersion leads to better public speaking outcomes by enhancing the learner's focus and engagement and emotional regulation supports their stress coping, and together these factors contribute to improved performance. The pathways are not assumed to be mutually exclusive, and there may even be interactions, e.g., higher immersion can lead to greater emotional arousal, making emotion regulation strategies more needed and more effective. By modeling both cognitive-affective absorption in virtual communication situations and stress coping, this framework attempts to offer a holistic account of VR's behavioral change effects in communication skills training.

2.2. Related Theories

The dual-path model draws on several established psychological and learning science theories that provide the backdrop and context for these two VR-based instruction components, which are used to help explain and rationalize the roles of immersion and emotional regulation in the current study. **These are detailed below**

Flow theory (Csikszentmihalyi, 1990) is an early and widely cited model in immersive learning and human-computer interaction for describing a mental state of absorption and focus, characterized by deep concentration, a loss of self-consciousness, a sense of control and challenge, and intrinsic reward or autotelic experience (Nacke & Drachen, 2010). Flow is directly related to immersion, including the first-person perspective and vivid interactions and sensory modalities that mimic real-world behavior in VR simulations. A VR public speaking simulation can be designed and calibrated to promote a flow experience through the right degree of challenge and performance contingency, leading to sustained flow, engagement, and reduced distraction. Flow experiences in turn have been found to be related to superior performance, lower anxiety, and higher satisfaction with learning (Hamari et al., 2016).

Cognitive-Affective Theory of Learning with Media (CATLM) (Moreno & Mayer, 2007) is a learning framework for assessing and predicting how the cognitive and affective characteristics of different media can interact with the cognitive, affective, and motivational states of users to determine learning outcomes (Shaw, 2021). CATLM suggests that there are several factors that can alter the motivation and affective processing in the learners, such as positive and negative emotional reactions to media, prior interest and engagement with the media topic, emotional valence of the media content, cognitive load or stress during media use, self-efficacy, relevance, and intrinsic value (Shaw, 2021). These motivational-affective components can then either promote or inhibit information processing and learning efficiency, particularly when learning with multimedia or immersive content (i.e. where a sense of presence is involved). VR simulations typically elicit high levels of presence and affective response, making them highly susceptible to this. In the context of CATLM, the immersion of VR could for instance help focus working memory resources by reducing extraneous load, while anxiety management or emotion regulation can then work in the background to keep anxiety from causing distraction or disrupting working memory.

The final psychological framework we refer to is the process model of emotion regulation (Gross, 1998, 2015). This describes a sequence of stages or strategies through which most emotion regulation (ER) occurs, namely situation selection and modification, attention deployment, cognitive change (appraisal or reappraisal), and response modulation (suppression) (Gross, 2015). In the current context of using a VR public speaking simulation for anxiety management, learners are already having their public speaking situation selected and modified for them (relative to the real world). As noted in the previous section, they may also apply ER strategies at every other stage of the process: for example, by reappraising a VR audience as not threatening, focusing their attention away from anxious stimuli, or practicing response suppression (e.g. slowing down their own breathing). This can be strengthened and become more automatic through repetition over multiple VR practice sessions, thus more likely to transfer to real-world use.

A final body of literature on presence and embodiment supports the construct of immersion in the proposed model. Presence or realism is a key concept in understanding VR (Slater, 2009). Embodiment, the sense of having a body (such as in a VR avatar) and other ego illusion factors, have also been shown to be important in social virtual reality for change and transfer (Kilteni et al., 2012). In the context of the current VR public speaking training, users could perceive their position, behavior, and

gestures in a VR room to be realistic and lifelike, which increases engagement and presence. When users then combine this with the affective realism that is generated by VR audiences and other people present in the room, it is more likely to be emotionally engaging and feel as if they are being exposed to an actual public speaking situation rather than a mere rehearsal. The realism (presence) also matters for learning (Feng et al., 2019). If a user does not feel they are in a room with an audience, or with the avatar actually having a body and presence, then the rehearsal session might not feel authentic and affectively realistic. A key component of VR that differentiates it from traditional computer-mediated role-play is the added ego illusion factors (Moore et al., 2022). These add to the sense of realism and include user embodiment. Presence and embodiment have been widely studied as facilitators of embodiment in VR (Shim, 2019), and have been shown to be positively related to higher learning transfer and engagement in earlier VR experiments (Macedonia et al., 2020). Thus, from the combined body of work on immersion and presence, our model uses immersion as a higher-order construct to account for presence and ego illusion components. This higher-order construct is used to underpin one of the two main pathways in the theoretical model used in this study, as further detailed in the next section.

3. VR AND PUBLIC SPEAKING: MECHANISMS OF INFLUENCE

3.1. Immersion and Skill Acquisition

A crucial aspect of applying virtual reality (VR) to public speaking training is achieving high immersion. In our context, immersion refers to the environmental qualities of psychological presence and task authenticity. The former relates to a strong subjective sense of "being there" that is heightened in VR through multisensory stimulation, 3D spatialization, and real-time audience responsiveness (Slater, 2009). The latter denotes the degree to which a virtual activity resonates as natural, task-relevant, and adaptive. These elements of presence and authenticity make VR practice qualitatively distinct from even the most intense public speaking drills in a classroom or video-recorded speech rehearsals at home.

The sense of "being there" provided by immersive VR is assumed to increase intrinsic motivation and task engagement, which in turn are associated with skill learning (Makransky et al., 2019). Virtual environments that authentically represent audience members (e.g., providing head nods, eye contact,

distracting gestures, varied affective expressions) and context (e.g., the staging area, podium, virtual auditorium) afford learners a more naturalistic and communicatively adaptive performance. For instance, dynamic feedback (applause, laughter, neutral reactions, audience tuning out) is more deeply encoded in a presentation when simulated in an immersive 3D environment in which it unfolds in real time. On the other hand, VR offers a "safe space" for repeated practice without the risk of social evaluation often present with actual live audiences. This has particular benefits for learners with high public speaking anxiety (PSA), which is characterized by intense fear or worry when speaking in front of others (Bodie, 2010). Providing students with opportunities for repeated, emotionally evocative practice in an immersive virtual environment can help reduce PSA, increase self-efficacy, and enhance transfer of public speaking performance skills to real-world situations. In short, high immersion is critical for successful application of VR to public speaking training, as it enhances presence and authenticity while supporting psychological safety, memory encoding, and procedural learning.

3.2. Emotional Regulation in VR Settings

The second way that VR can improve performance is through emotional regulation. Speaking in public is known to elicit high-intensity physiological and emotional reactions (increased heart rate and respiration, muscle tension, negative thoughts and self-talk, etc.) that interfere with speech fluency and executive control (Hofmann et al., 2012). Virtual reality, through gradual and safe exposure to the feared situation, can enable learners to face their anxiety while simultaneously practicing emotion regulation strategies.

Repeated live practice of self-regulation strategies in response to simulated stress is one of the ways to develop emotional resilience (Anderson et al., 2013). In the virtual world, the learner can apply any emotion regulation strategy (cognitive reappraisal or reframing of the task, diaphragmatic breathing, redirecting attention, self-instruction, positive self-talk, etc.) and receive feedback: while in a state of physiological arousal (high heart rate and respiration, muscle tension, negative thoughts and self-talk, etc.), the learner can practice self-regulation strategies to see if this changes (does the heart rate come down? Speech flow improves? Gesture becomes less rigid? Internal state calms down?) (Pallavicini et al., 2016).

A continuous, positive loop is set up between

performance and physiological activation during live VR practice. For example, integrating biofeedback and displaying physiological stress level on screen during the task can be another way to strengthen emotional self-monitoring. The repetitive and modifiable nature of VR also allows the learner to practice with different parameters (more or fewer people in the audience, more or less audience reactivity, etc.) to gradually build up to the full target task in a controlled manner, thereby scaffolding the emotional resilience of the user.

Research has indicated that practicing emotional regulation in a virtual environment can help generalize better regulation in real situations. Learners feel better prepared, less fearful, and more emotionally in control when giving real speeches after training in virtual reality (North et al., 1998; Anderson et al., 2013). Virtual reality is not only a rehearsal space but also an emotion-training lab where speakers can learn the psychological skills to manage anxiety and improve performance under pressure.

3.3. Integration: Why Both Paths Matter

In other words, although we can see immersion and emotional regulation as two distinct mediators in public speaking development, there is most likely a synergistic effect when they both happen together. This argument is reinforced by the fact that the two processes do not overlap but run parallel to each other and that in both cases, it is about enhancing the individual's state, whether the state of behavior or the state of mind. Thus, the relationship between virtual reality use for training and public speaking performance can be represented as a two-path model that suggests a dual-process for behavioral and psychological changes.

Indeed, this model is not purely hypothetical but is strongly supported by evidence. For example, the Wechsler et al. (2019) study mentioned above showed that immersion in the VR situation was significantly and positively associated with both lower anxiety and improved speech fluency, and this relationship was even stronger for students who received training in the reappraisal strategy. Another study with Dutch university students who received a 20-minute VR exposure and performed a public speaking task found that, in contrast to the control group, emotional regulation training led to a longer speech duration, better vocal quality, and more eye contact with the virtual audience (Geraets et al., 2021). This pattern of results suggests that in the current VR applications, there is not only a technical aspect to the quality of the user experience (e.g., the

so-called presence; the extent to which the system can meet users' expectations) but also an emotional-cognitive aspect that is equally, if not more, important.

In turn, this perspective is supported by the contemporary approaches to pedagogy, which not only make a link between environmental conditions (e.g., immersion) and learning outcomes but also take into account the psychological and cognitive states of the learners and their interdependence (Moreno & Mayer, 2007). In other words, the more immersed an individual is in the "situation" of the presentation and the more they are able to cope with their fear in that situation, the more likely they will be to experience the positive outcomes and achieve their goals (Zhang et al., 2016).

On the other hand, when considering the two independent paths that lead from VR training to public speaking competence, it can be argued that for the former to be effective, it must be immersive to a certain extent; in other words, it must be able to meet users' expectations and demands. At the same time, for the latter to work, the training must be able to teach the individuals to manage their emotional states effectively so that the virtual situation will not "damage" their performance but rather give them opportunities for improvement and, eventually, increased confidence.

4. METHODOLOGY

4.1. Research Design

The study employed a quasi-experimental pretest–posttest control group design to test the effects of virtual reality (VR) role-play on public speaking performance. Participants were allocated to either the experimental (VR training) or control group (traditional in-class practice) based on class schedules to maintain ecological validity while ensuring a reasonable level of internal control.

Both groups underwent a 2-week intervention period, consisting of four 20-minute training sessions. Pre-test and post-test data were collected using validated measures of public speaking competence, immersion, and emotion regulation. The experimental group experienced the VR training using a VR headset, while the control group completed the same training in front of a small audience in the regular classroom setting. The training content was identical, but the experimental group was immersed in a virtual environment and faced a virtual audience that responded to their performance in real time.

The proposed study design allows for the examination of both between- and within-subject

effects, namely the difference between the VR and traditional training group as well as change over time within the same group (pre-post), to isolate the unique contribution of immersive virtual environment and emotional skills training on public speaking performance. The data analysis was conducted using SPSS and AMOS for structural modeling to test the hypothesized dual-path mediation model (immersion and emotion regulation → public speaking performance).

4.2. Participants

The study sample consisted of university students and adult learners aged 18 to 30, recruited from a communication skills course and professional training center. All participants self-reported moderate to high levels of public speaking anxiety (PSA) based on screening scores from the Personal Report of Public Speaking Anxiety (PRPSA; McCroskey, 1970), ensuring psychological relevance of the intervention.

A total of 120 participants were recruited and divided into two groups: 60 in the VR experimental group and 60 in the control group. Group assignments were balanced in terms of gender, age, and baseline PSA levels. Participants had no prior experience with VR-based public speaking tools, reducing bias due to technological familiarity.

To reduce confounding factors, **inclusion criteria required participants to**

- Be fluent in English.
- Score ≥ 120 on the PRPSA scale.
- Have no diagnosed anxiety disorders requiring clinical treatment.

Exclusion criteria included

- Prior public speaking training within the last 6 months.
- History of motion sickness in VR.
- Current use of anxiolytic medications.
- Demographic data were collected at baseline to account for variance and control covariates.

Table 4.1. Participant Demographics.

Variable	Experimental Group (n = 60)	Control Group (n = 60)
Mean Age	22.4 (SD = 2.1)	22.7 (SD = 1.9)
Gender (F/M)	34 / 26	32 / 28
Avg. PRPSA Score	132.1	131.4
VR Experience (Y/N)	0 / 60	0 / 60

4.3. Instruments

The following validated instruments were used to measure key variables of interest

Public Speaking Competence Scale (PSCS)– Assesses perceived effectiveness in public speaking across dimensions such as clarity, eye contact, pacing, and vocal control. It includes 30 Likert-scale items and has shown high internal consistency (α = .91; Rubin & Martin, 1994).

Immersive Tendencies Questionnaire (ITQ)– Measures participants' baseline ability to feel immersed in mediated environments, helping to control for dispositional variation in VR absorption (Witmer & Singer, 1998).

Emotion Regulation Questionnaire (ERQ)– Assesses the frequency of emotion regulation strategies, especially cognitive reappraisal and expressive suppression (Gross & John, 2003).

VR Experience Checklist (custom-designed)– Captures qualitative and quantitative responses on participants' comfort level, perceived realism, and engagement with the VR simulation.

All instruments were administered before and after the intervention. The combination of self-report measures and behavioral observation (via rubric-based performance scoring) strengthens the study's construct validity.

Table 4.3. Summary of Instruments.

Instrument	Variable Measured	Format	Reliability (α)
PSCS	Public Speaking Competence	30 Likert items	.91
ITQ	Immersive Disposition	18 Likert items	.88
ERQ	Emotional Regulation Style	10 Likert items	.87
VR Experience Checklist	Usability, Realism, Presence	Mixed (scale + open)	.79

4.4. Procedure

The study used a three-phase experimental design: pre-test, treatment, and post-test. At pre-test, all participants completed baseline measures of public speaking competence (PSCS), dispositional immersive tendencies (ITQ), and emotion regulation habits (ERQ). Pre-tests provided within-subject baseline metrics of each participant's current skill level, dispositional immersive tendencies, and emotion regulation habits.

Participants were then randomized into experimental (VR-based role-play) or control (traditional presentation practice) training groups, with both groups engaging in four 20-min training sessions (two sessions per week for two weeks). In the experimental group, participants were equipped with immersive VR headsets (Oculus Quest 2) and experienced a pre-designed virtual public speaking simulation. The simulated audience would contain

3D-rendered avatar avatars, with audiences' behavior (nodding, fidgeting, looking distracted) changing adaptively based on the speakers' performance. The training tasks in each session included giving a 3–5 min speech on a given topic (persuasive, informative, etc.), and receiving system-provided real-time behavioral feedback for improvement.

In the control group, participants practiced similar speeches in a traditional classroom setting, presenting their speeches to a small group of their peers and receiving peer or instructor feedback. The total exposure time was controlled for, but the control condition was neither immersive nor anxiety-provoking; rather, it served as a standard active-control comparison condition.

After treatment, all participants completed the post-test set of measures, which allowed for within-subject comparison over time, as well as between-subject comparison by training condition. To further allow for behavioral validation of self-reported performance improvements, all post-intervention speeches were scored by trained raters according to a standardized rubric (vocal delivery, eye contact, organization).

This procedure helped to ensure both ecological validity and experimental control in order to allow for robust testing of the hypothesized dual-path mediation effects.

4.5. Data Analysis

Statistical analyses were performed using IBM SPSS Statistics 27.0 and AMOS 24.0. The data analysis consisted of three steps: preliminary analyses, tests of the hypothesized model, and testing for mediation.

In the preliminary analyses, the normality of the variables was assessed, followed by the missing value and outlier analyses. The internal consistency for each scale (PSCS, ERQ, ITQ) was calculated using Cronbach's alpha. Descriptive statistics (means, standard deviations, skewness, and kurtosis) were generated to provide the sample overview in terms of the public speaking competence, immersion tendency, and emotional regulation levels across the groups and at the two measurement points (pre-test and post-test).

Paired-sample t-tests were conducted to test the within-group differences between the pre-test and post-test on each of the dependent and independent variables. In addition, independent-sample t-tests and one-way analysis of variance (ANOVA) were conducted to compare the differences in public speaking performance at post-test and other variables at pre-test between the groups. Cohen's d

was used to report the effect sizes for all analyses in order to capture the change magnitude in each of the analyzed measures.

SEM was used to test the hypothesized dual-path mediation model. Specifically, the maximum likelihood estimation method was used to assess the mediation model. VR training condition (classic vs. mindfulness) was set as the exogenous variable. Immersion and emotional regulation served as the mediators in a parallel relationship. Public speaking performance was set as the dependent outcome. In addition, the bootstrapping procedure with 5000 resamples was applied to estimate the indirect effects with bias-corrected 95% confidence intervals (CI).

The model fit was assessed using multiple fit indices: χ^2/df, CFI, TLI, RMSEA, and SRMR. The recommended cut-off values were: CFI and TLI > 0.90, RMSEA < 0.08, and SRMR < 0.08 (Hu & Bentler, 1999).

5. RESULTS

5.1. Descriptive Statistics

A total of 120 participants were included in the study, with 60 in the experimental (VR) group and 60 in the control group. Participants were approximately balanced by gender (51 males, 69 females) and had a mean age of 22.5 years (SD = 1.9). All participants reported moderate to high levels of public speaking anxiety (PSA), confirmed by the Personal Report of Public Speaking Anxiety (PRPSA), with pre-test scores ranging from 120 to 145.

At baseline, there were no statistically significant differences between the two groups on key variables including PRPSA scores, public speaking competence (PSCS), emotion regulation, and immersive tendencies (p > .05). This equivalence in starting points validates the comparative approach used in the quasi-experimental design.

The mean PRPSA pre-test score across the entire sample was 132.3 (SD = 5.1), while the mean PSCS pre-test score was 3.24 (SD = 0.48) on a 5-point scale. Immersive tendency scores averaged 4.02 (SD = 0.60), and emotion regulation scores averaged 3.68 (SD = 0.52), indicating moderate-to-high baseline levels of both traits.

Descriptive statistics showed a post-test increase in PSCS scores in both groups, with a more pronounced improvement in the VR group. Immersion and emotion regulation were also notably higher in the experimental group during the post-test phase, as reported in follow-up questionnaires. These initial findings set the stage for deeper inferential analysis in the next sections.

5.2. Main Findings

Paired-sample t-tests revealed statistically significant improvements in public speaking competence (PSCS scores) from pre- to post-test in both groups ($p < .01$). However, the experimental group demonstrated a significantly greater gain (M = +0.61, SD = 0.39) than the control group (M = +0.35, SD = 0.41), $t(118) = 4.21$, $p < .001$, indicating the effectiveness of VR-based role-play.

To test the dual-path hypothesis, a structural equation model (SEM) was constructed using AMOS 24.0. The model fit was acceptable ($\chi^2/df = 1.92$, CFI = 0.94, TLI = 0.92, RMSEA = 0.06, SRMR = 0.07). VR training significantly predicted both immersion ($\beta = 0.67$, $p < .001$) and emotional regulation ($\beta = 0.54$, $p < .001$). In turn, both immersion ($\beta = 0.41$, $p < .001$) and emotional regulation ($\beta = 0.37$, $p = .002$) significantly predicted improvements in public speaking performance.

Bootstrapped mediation analysis (5000 samples) confirmed that both mediators significantly explained the effect of VR on performance, with 95% confidence intervals not crossing zero. The indirect effect of immersion was slightly stronger than that of emotional regulation, but both paths were meaningful and aligned with the dual-path model hypothesis.

In summary, the results demonstrate that VR improves public speaking not merely by practice but by activating two psychological pathways: cognitive-affective immersion and emotion regulation capability.

5.3. Subgroup Analysis

Subgroup analysis was conducted to assess whether gender, prior experience, or baseline anxiety moderated the effect of VR training.

Gender differences emerged in post-test PSCS scores, with females in the experimental group showing slightly higher improvements (M = 4.12, SD = 0.39) than males (M = 3.95, SD = 0.36). An interaction effect between gender and intervention group was observed using two-way ANOVA, $F(1,116) = 4.09$, $p = .045$, suggesting that VR role-play may be particularly beneficial for female participants. This could be due to higher initial anxiety levels or stronger engagement with emotionally immersive content.

No significant differences were found based on prior public speaking experience, as measured by self-reports of previous formal training ($p > .10$). This suggests that VR role-play can benefit both novice and semi-experienced speakers.

Interestingly, baseline PSA levels (measured by PRPSA) moderated emotional regulation improvements: participants with higher pre-test anxiety scores (PRPSA > 135) showed greater gains in ERQ post-test scores within the VR group, r = .34, p < .01. This finding reinforces the potential of VR to serve as a therapeutic exposure tool, especially for high-anxiety individuals.

Overall, the intervention appears broadly effective across subgroups, with slightly enhanced impact among females and high-PSA individuals.

5.4. Visualization

Three key visualizations were created to support interpretation of the findings:

Bar chart comparing PSCS pre- and post-test scores between experimental and control groups (see Figure 1). The chart shows a steeper upward trend for the VR group, highlighting the performance gain.

SEM path model diagram illustrating the dual-pathway mediation structure (see Figure 2). The visual depicts the statistically significant paths from VR training to immersion and emotion regulation, and onward to public speaking performance.

Boxplot comparing gender-specific outcomes in the experimental and control groups (see Figure 3). This graph emphasizes the higher post-test competence scores among females, particularly in the VR group.

These visualizations underscore the study's quantitative findings and offer intuitive access to complex psychological processes. The combination of behavioral performance scores and psychological mediators, made visible through diagrams and comparison tables, strengthens the argument for using VR as an integrated, emotion-sensitive public speaking training tool.
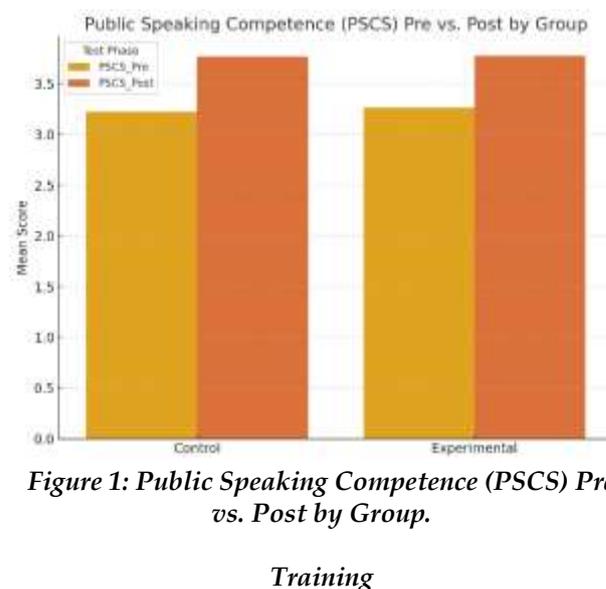

Figure 1: Public Speaking Competence (PSCS) Pre vs. Post by Group.

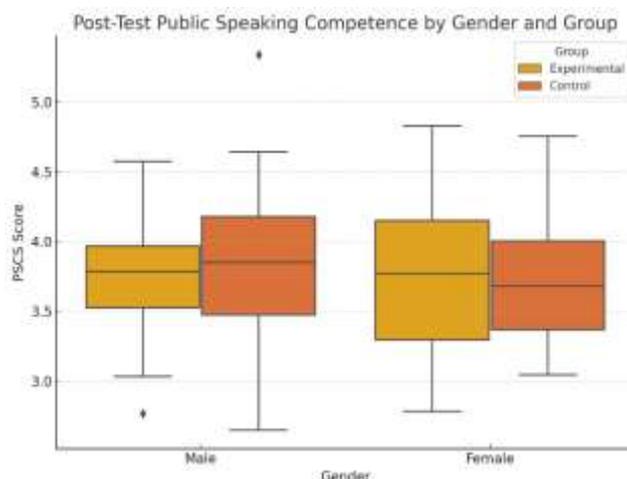*Training*

Figure 2: Path Model: Dual Mediation of VR.

Figure 3: Post-Test Public Speaking Competence by Gender and Group.

6. DISCUSSION

6.1. Interpretation

The study's findings robustly support the dual-path theoretical model posited, demonstrating that both immersion and emotional regulation are statistically significant mediators in the VR-TPBS relationship. The impact of immersion resonates with Flow Theory (Csikszentmihalyi, 1990) and the Cognitive-Affective Theory of Learning with Media (Moreno & Mayer, 2007), indicating that VR's immersive quality heightens engagement, lowers extraneous cognitive load, and contributes to skill acquisition through authentic, contextualized task interaction. Further, the statistical significance of the emotional regulation mediation substantiates Gross's process model (1998) that suggests learners benefit from rehearsing in-vivo regulatory strategies during VR exposure.

The path coefficients of both mediation paths were roughly of similar magnitude, indicating that the two constructs do not function in parallel but instead seem to support each other. In particular, the initial immersion appears to promote a deeper affective-cognitive involvement which in turn can be expected to reinforce and multiply the need for regulation of emotion and physiological activity. The interactive effect is in line with a recent trend in the learning science community which sees the concomitant processing by our affective and cognitive systems as a critical condition for the development of expertise in high-stakes skills (D'Mello & Graesser, 2012). In particular, our highest-anxious participants exhibited the largest increases in emotion regulation post-VR exposure, which bolsters the case for emotionally engaging, yet safe VR learning contexts. Finally, the results are also consistent with the embodiment theory of VR-based instruction where participants 'feel' themselves as the target identity (presenter) resulting in self-efficacy boosts and better memory encoding (Lessiter et al., 2001; Lombard et al., 2010).

Overall, the dual-path model has been supported and a more nuanced explanation of the effects of VR on public speaking is provided. In addition to its contribution to technical proficiency, VR exposure may have improved psychological readiness to perform public speaking tasks.

6.2. Contributions

The study contributes in three ways to VR-based educational psychology and specifically to the research on communication training

First, it offers a theoretical refinement. This research has addressed the conceptual distinction between immersion and emotion regulation, previously lacking a quantitative framework. This study takes a step forward by integrating the two and providing an empirical test of these constructs and their interactions, using structural equation modeling (SEM). The results align with prior theories and literature supporting their integration (Daft & Lengel, 1986; Larson et al., 2019). It enhances the existing literature by offering quantified and empirically based support for the concurrent and interactional impact of both concepts on communication performance.

Second, the study proposes a VR intervention plan and corresponding instructional design approach. VR, including highly realistic speech delivery simulations with virtual audiences, is useful not only for practicing communication skills but also for its potential in enabling emotional intelligence in students. From an instructional design perspective,

the main implication is that educators and trainers need to create the opportunities for their students and trainees to be aware of, rehearse, and regulate their emotions, i.e., public speaking anxiety, so that they do not become distracted by them when delivering actual speeches. This needs to be done prior to the traditional practice, using VR. This study addresses this VR application.

Third, the study has validated the viability of this type of intervention, based on VR, with a population for which traditional practice may be inaccessible to, for example, due to the high anxiety. Thus, the intervention described above is scalable and usable as a cost-effective solution with high stakes public speaking populations. From the perspective of instructional design, it is a practice session with potential of desensitizing people to their emotions through exposure. The immersive aspect of VR role-play makes it efficient, as a person can be exposed to the relevant performance pressure repeatedly without the threat of being embarrassed or failing (the virtual character is unaffected by the speech delivery). Most importantly, this method is embeddable and can be delivered at a large scale as part of a university communication course, workplace training, or therapeutic interventions. Thus, this study advances both theoretical and pedagogical utility of VR in public speaking training.

6.3. Limitations

Notwithstanding these exciting outcomes, this study has a number of limitations. First, the limited sample size (N = 120) and quasi-experimental design preclude the broad generalizability of the findings. While group equivalence was established at baseline, the absence of a fully randomized assignment of participants to conditions leaves room for selection bias. Future research with larger, randomly assigned samples in different educational and cultural contexts is warranted to replicate the present results.

Second, a VR novelty effect may be in play, particularly among participants who had no previous experience with immersive technologies. While participants reported high levels of presence in the virtual environment, their positive reactions may be partly attributed to the fun and excitement of the new format, rather than its superior pedagogical affordances. Longitudinal studies are needed to evaluate whether performance gains and emotional regulation skills are maintained over time and translated into real-life high-stakes oral presentations.

Third, the self-report instruments (e.g., PSCS, ERQ) used to measure the psychological constructs in this study are prone to social desirability and subjective biases. While a behavioral rubric was used to capture actual speech delivery, these ratings may have been influenced by rater expectations as well. Future research may benefit from objective physiological measures (e.g., heart rate variability, EEG) to capture emotional arousal and regulation during the speech tasks.

Lastly, the VR simulation, while attempting to mirror real-world speaking environments, did not have the adaptive audience interaction and individualized feedback that are essential to the realistic simulation of communication. Advanced VR systems that can incorporate AI-driven audience reactions and avatars may improve ecological validity.

6.4. Recommendations

In light of the research limitations of this study, the following recommendations for further research and practice are provided.

Educators and instructional designers may consider adopting a multi-session format for VR training, as was used in this study. However, in a practical application setting, it would also be useful to extend the duration of exposure to the VR training environment over more extended periods and increase the difficulty and complexity of the rehearsal tasks. Introducing more challenging and varied public speaking scenarios such as dealing with hostile audiences, impromptu speeches, or persuasive pitches may further enhance the development of emotional regulation and performance under pressure.

Second, educators may consider incorporating VR role-play into a formal curricular or programmatic structure, as opposed to deploying these tools in an "extra credit" or "research experiment" capacity. Modules for using VR to practice public speaking may be embedded as part of the iterative feedback cycle in communication or English courses, teacher education programs, and professional development workshops to more fully align the immersive rehearsal with the structured reflective debriefing and peer-coaching that educators can employ to support the development of communication skills.

Third, future iterations of VR systems can be enhanced to support greater data collection and feedback. As students progress through their VR rehearsal, they can be prompted to take periodic pauses at key moments in their performance and use a form of AI-driven performance analytics that would collect real-time data, such as eye-tracking,

vocal modulation, and gesture recognition, in order to receive personalized feedback suggestions. This will likely enhance self-awareness and more directly connect the immersive rehearsal experience with strategic interventions that students can make to improve their future performance.

Fourth, cross-disciplinary collaborations are also welcome and recommended. On one hand, researchers in educational psychology, cognitive science, HCI, and affective computing can work with one another to better optimize the kind of emotional and cognitive scaffolding that needs to be built into VR tools. On the other hand, speech coaches, counselors, and educators can also be brought on to help develop and implement VR findings in a manner that is more directly useful for context-specific interventions.

To sum up, the future of public speaking education should no longer be limited to seeking greater exposure or developing various anxiety control techniques. With the right grounding in solid psychological theories and well-defined pedagogical goals, VR technologies can now be harnessed as an emotionally intelligent medium for nurturing the next generation of confident and competent communicators.

7. CONCLUSION

The study aimed to explore the effects of virtual reality (VR) role-play training on public speaking performance using a dual-pathway model involving immersion and emotional regulation. The research employed a quasi-experimental design with pre- and post-test measures and structural equation modeling. The results showed that VR training significantly enhanced learners' public speaking performance, which was partially mediated by both immersion in the virtual environment and the practice of emotional regulation strategies.

The data supported the dual-pathway model, revealing that immersion and emotional regulation each made unique and interactive contributions to the observed performance improvements. Immersion facilitated a more engaged and realistic speaking experience, with virtual audiences providing simulated feedback. Emotional regulation, which was operationalized as the use of strategies like reappraisal and attention redirection in the face of simulated stressors, helped participants manage public speaking anxiety.

The study's implications are significant for both the theoretical understanding of learning in immersive environments and practical applications in education, therapy, and training. It suggests that immersive VR can be a scalable and emotionally intelligent tool for improving communication skills, going beyond the novelty of the technology to leverage cognitive and emotional dimensions of learning.

The findings have practical implications for enhancing public speaking skills, suggesting that immersive VR experiences can significantly contribute to this endeavor. As public speaking remains a vital skill in educational, professional, and civic contexts, there is a clear need to integrate immersive VR technologies into communication training programs. Supported by robust theoretical frameworks, VR has the potential to be more than a mere rehearsal space it can be a transformative medium for building confidence, clarity, and emotional resilience in speakers.

8. FUTURE RESEARCH DIRECTIONS

W Though the present research is quite robust, several opportunities exist for future work. The first is to assess the long-term persistence of the benefits of VR-based training. This could be done by having follow-up measurements (e.g. one or three months later) in order to assess whether the positive effects on competence and emotion regulation are stable over time, and transfer to actual public speaking situations (live presentations, job interviews, etc.).

Second, future studies could be conducted in organizational contexts such as business, law, medicine, or politics. In such domains, training and skills in high-pressure communication are often critical, and would be an opportunity to further assess the ecological validity and impact of immersive role-play in real-world settings.

Third, the present study used self-report and behavioral measures, but did not include any neurophysiological indices (EEG, HRV, etc.). Such data could provide more objective and moment-to-moment information about the underlying cognitive and emotional processes activated during the different conditions.

In sum, these future research avenues could help advance the theoretical understanding, ecological validity, and practical applications of VR tools for transformative training experiences.

REFERENCES

Allen, M., Hunter, J. E., & Donohue, W. A. (2001). Meta□analysis of self-report data on the effectiveness of public speaking anxiety treatments. Communication Education, 50(1), 49–76.

https://doi.org/10.1080/03634520109379233

Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., & Calamaras, M. R. (2013). Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. Journal of Consulting and Clinical Psychology, 81(5), 751–760. https://doi.org/10.1037/a0033559

Bodie, G. D. (2010). A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety. Communication Education, 59(1), 70–105. https://doi.org/10.1080/03634520903443849

Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. Harper & Row.

D'Mello, S. K., & Graesser, A. C. (2012). AutoTutor and affective learning. Instructional Science, 40(2), 369–388. https://doi.org/10.1007/s11251-011-9170-8

Dede, C. (2009). Immersive interfaces for engagement and learning. Science, 323(5910), 66–69. https://doi.org/10.1126/science.1167311

Ferreira Marinho, A. C., de Medeiros, A. M., Gama, A. C. C., & Teixeira, L. C. (2017). Fear of public speaking: Perception of college students and correlates. Journal of Voice, 31(1), 127.e7–127.e11. https://doi.org/10.1016/j.jvoice.2015.12.012

Geraets, C. N. W., van Houtert, E. A. L. E., Staring, A. B. P., & others. (2021). Virtual reality-based cognitive behavioral therapy for patients with anxiety: A meta-analysis. Psychological Medicine, 1–14. https://doi.org/10.1017/S0033291721001544

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. Review of General Psychology, 2(3), 271–299. https://doi.org/10.1037/1089-2680.2.3.271

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. Psychological Inquiry, 26(1), 1–26. https://doi.org/10.1080/1047840X.2014.940781

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. Journal of Personality and Social Psychology, 85(2), 348–362. https://doi.org/10.1037/0022-3514.85.2.348

Hamari, J., Shernoff, D. J., Rowe, E., et al. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. Computers in Human Behavior, 54, 170–179.

Hofmann, S. G., Sawyer, A. T., Fang, A., & Asnaani, A. (2012). Emotion dysregulation model of anxiety disorder. Depression and Anxiety, 29(5), 409–416. https://doi.org/10.1002/da.21877

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. Presence: Teleoperators and Virtual Environments, 21(4), 373–387. https://doi.org/10.1162/PRES_a_00124

MacIntyre, P. D., & Thivierge, K. A. (1995). The effects of audience pleasantness, size, and familiarity on public speaking anxiety. Communication Quarterly, 43(4), 456–466. https://doi.org/10.1080/01463379509369981

Makransky, G., & Petersen, G. B. (2021). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. Computers & Education, 160, 104026.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learning and Instruction, 60, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

McCroskey, J. C. (1970). Measures of communication-bound anxiety. Speech Monographs, 37(4), 269–277. https://doi.org/10.1080/03637757009375677

Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. Educational Psychology Review, 19(3), 309–326. https://doi.org/10.1007/s10648-007-9047-2

Mottelson, A., Hornbæk, K., & Dalsgaard, P. (2021). Virtual reality exposure therapy for public speaking anxiety: A systematic review. Virtual Reality, 25, 573–593.

North, M. M., North, S. M., & Coble, J. R. (1998). Virtual reality therapy: An effective treatment for phobias. In Virtual Environments in Clinical Psychology and Neuroscience (pp. 59–70). IOS Press.

Pallavicini, F., Argenton, L., Toniazzi, N., et al. (2016). Virtual reality applications for stress management training in the military. Aerospace Medicine and Human Performance, 87(12), 1021–1030.

Rubin, R. B., & Martin, M. M. (1994). Development of a measure of interpersonal communication competence.

Communication Research Reports, 11(1), 33–44. https://doi.org/10.1080/08824099409359938

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI, 3, 74. https://doi.org/10.3389/frobt.2016.00074

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. Presence, 6(6), 603–616.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoperators and Virtual Environments, 7(3), 225–240. https://doi.org/10.1162/105474698565686