frontiers in Virtual Reality

TYPE Original Research
PUBLISHED 24 March 2026
DOI 10.3389/frvir.2026.1755571

Check for updates

OPEN ACCESS

EDITED BY
Nadia Magnenat Thalmann,
University of Geneva, Switzerland

REVIEWED BY
Ïo Valls-Ratés,
University of Southern Denmark, Denmark
Jordan Aiko Deja,
De La Salle University, Philippines

*CORRESPONDENCE
Lamia Bettahi,
lamia.bettahi@uliege.be

RECEIVED 27 November 2025
REVISED 22 January 2026
ACCEPTED 16 February 2026
PUBLISHED 24 March 2026

CITATION
Bettahi L, Remacle A, Schyns M, Etienne É, Etienne A-M and Leclercq A-L (2026) Validating virtual reality for public speaking research and intervention: comparing anxiety, voice, and fluency responses to real and virtual audiences. *Front. Virtual Real.* 7:1755571. doi: 10.3389/frvir.2026.1755571

Validating virtual reality for public speaking research and intervention: comparing anxiety, voice, and fluency responses to real and virtual audiences

Lamia Bettahi[1,2]*, Angélique Remacle[2,3], Michaël Schyns[4], Élodie Etienne[5], Anne-Marie Etienne[2] and Anne-Lise Leclercq[2]

[1]Fund for Scientific Research – F.R.S.-FNRS, Brussels, Belgium, [2]Research Unit for a Life-Course Perspective on Health and Education, University of Liège, Liège, Belgium, [3]Center for Research in Cognition and Neurosciences, Faculty of Psychological Science and Education, Université Libre de Bruxelles, Brussels, Belgium, [4]QuantOM, HEC Liège, N1 UER Opérations, University of Liège, Liège, Belgium, [5]ALMAnaCH, Centre Inria de Paris, Paris, France

Introduction: Public speaking (PS) is a widespread activity required in many personal and professional settings. This activity is known to elicit anxiety, subsequently affecting oral communication, especially voice and speech parameters. As mastering PS skills requires practice in situations that are as similar as possible to reality, virtual reality (VR) may represent a promising method for research, training and intervention in this domain. However, it is of paramount importance to first validate VR environments in their ability to reproduce authentic anxiety responses and communicative behaviors, which are often overlooked.

Methods: Therefore, this study examined university students (N = 60) anxiety responses (self-reported and heart rate) as well as voice and fluency adjustments to a PS task performed either in (1) a real meeting room in front of an audience, (2) a virtual meeting room in front of an audience, and (3) the same virtual meeting room without any audience. As this last condition contained no anxious stimulus, it was included to act as a control for the anxiety induced by VR immersion. The main objective of this study was to examine the influence of the real vs. virtual nature of the audience on anxiety, voice and fluency parameters.

Results: Our results showed that the virtual audience elicited changes in anticipatory anxiety (increased heart rate and self-reported anxiety) compared to the control condition. The participant's strong feeling of presence and lack of side effects such as cybersickness support the acceptability and usability of the virtual environment.

Discussion: Our results extend previous data and support the feasibility and relevance of using VR for PS. Additionally, we describe different VR immersion profiles among participants.

KEYWORDS
speech fluency, heart rate, oral communication, public speaking, virtual reality, voice

1 Introduction

Public speaking (PS) is a skill required in many personal and professional situations. It can be used to defend an opinion, convey an idea or convince others on a daily basis, whether at work or at school (Van Ginkel et al., 2015; Wörtwein et al., 2015). It is also a competence that university students must acquire, as 78.8% of employers consider oral communication skills to be among the five most important attributes in recent college graduates (Montes et al., 2019).

PS competence can be defined as 'the combination of knowledge, skills, and attitudes needed to speak in public in order to inform, self-express, relate and persuade' (De Grez et al., 2009). This ability is frequently regarded as comprising three underlying dimensions: knowledge, motivation, and skills (Morreale and Pearson, 2008). Knowledge is defined as the theoretical background a person deploys to plan and deliver a speech. Motivation refers to the willingness to communicate and depends on self-perceived communication competence. A repertoire of PS skills is also necessary, such as visual nonverbal behavior (e.g., eye contact, gestures, posture), nonverbal-auditory impressions (e.g., voice and speech parameters), language usage (e.g., linguistic expression, usage of rhetorical devices) and organization (e.g., structure of the speech) (Morreale and Pearson, 2008). Consequently, in order to deliver a competent speech, it is incumbent upon the speaker to adroitly amalgamate these dimensions within the context of the situation.

Nevertheless, most people do not experience a sense of ease when it comes to public speaking. This assertion is applicable to the general population (i.e., typical individuals) (Ferreira Marinho et al., 2017), as well as to individuals diagnosed with anxiety disorders (e.g., social anxiety) (Ebrahimi et al., 2019; Heeren et al., 2013), voice disorders (e.g., dysphonia) (Amir and Levine-Yundof, 2013), or speech disorders (e.g., stuttering) (Iverach et al., 2011). In higher education students, several studies demonstrate that PS is a highly anxiety-provoking situation (Baptista et al., 2012; Dwyer and Davidson, 2012; Ferreira Marinho et al., 2017; Rasouli et al., 2024; Russell and Topham, 2012).

PS anxiety, and anxiety in general, can be divided into two types: trait anxiety and state anxiety. The former refers to one's general tendency to feel tension, nervousness, and worry regardless of the situation whereas the latter refers to the anxiety felt in a specific moment (Leal et al., 2017; Spielberger, 1983). PS anxiety can also induce symptoms in three distinct dimensions: cognitive, physiological and behavioral (Lazarus and Opton, 1996). In the cognitive domain, PS anxiety can lead to catastrophic thoughts due to the fear of being judged negatively by others (e.g., "I will never succeed"). In the physiological domain, acute psychosocial stressor such as PS can initiate the autonomic stress response[1] of the nervous system (via Sympatho-Adrenal-Medullary axis) and increase blood flow, heart rate (*HR*), sweat and pupil dilation in a matter of minutes (Taylor et al., 2000). Although there are no universally recognized measures for stress evaluation, current neurobiological evidence supports the use of *HR* for objective assessment (Kim et al., 2018). As these devices are relatively low-cost and suitable for large samples, several studies use non-invasive monitors (wrist or chest worn) with a photoplethysmography technique (LED lights) to determine *HR* through optical blood flow (Bitsika et al., 2014; Tomasi et al., 2024; van Kraaij et al., 2020). In other words, in these studies, *HR* is calculated by accounting for the blood flow's inverse relationship with the amount of light refracted (Spierer et al., 2015). Therefore, high beats per minute (bpm) indicate high *HR* and high physiological stress, whereas low bpm reflect less *HR* and lower physiological stress. While *HR* of a normal resting adult is around 60–100 bpm, it can increase up to 150 bpm in anxiety-provoking situations (Kothgassner et al., 2016; Zhang and Zhang, 2009). In a PS context, an increase of 4.7 bpm in *HR* may already be considered as a meaningful indicator of anxiety (Schneider et al., 2021).

In the behavioral domain, PS anxiety impacts oral communication, particularly voice production and fluency. For example, voice fundamental frequency (f_0), meaning the number of times per second our vocal folds vibrate, correlates with the perception of a higher-pitched voice. A decrease in f_0 variation, corresponding to less intonation, can also be observed in anxious speech (Buchanan et al., 2014; Giddens et al., 2013; Hagenaars and Van Minnen, 2005; Van Puyvelde et al., 2018). Moreover, speech fluency, meaning the global impression of a smooth speech (De Jong, 2018), can be affected by anxiety. Disfluencies (interruptions in the flow of speech) seem to increase as anxiety rises. Specifically, an elevation of the number of filled pauses, the most common disfluencies in spontaneous speech (e.g., uh; Duez, 2001), is observed. Research also shows an increase in silent pauses frequency and length (Buchanan et al., 2014; Goberman et al., 2011; Hofmann et al., 1997; Laukka et al., 2008; Metz and James, 2019). According to Juslin and Scherer (2005), vocal response under emotional arousal involve adaptations at three different levels, recalling and refining the three dimensions of anxiety presented above. The first level is the physiological level, referring to changes in the autonomic nervous response system (e.g., increased HR; breathing patterns modification). The behavioral domain is then split into two different levels: the phonatory-articulatory level and the acoustical level. The phonatory-articulatory level refers to the change of muscular movement in the major structures involved in voice and speech production (i.e., increased tightness of oral motor structures influencing vocal folds vibration). Finally, the acoustic level refers to changes in the speech wave which could be objectively measured (i.e., f_0 elevation leading to a higher-pitched voice). Although these levels are interdependent their relationship remains unclear. Hence, changes at one level do not necessarily produce equivalent effects on another. Consequently, there is a clear need for studies that examine these interconnected levels of vocal behavior to better understand how anxiety shapes communicative expression and how these changes can be targeted through training. While physiological indicators such as HR provide valuable information about the autonomic stress response, they offer limited insight into how anxiety manifests through actual communicative behavior. In contrast, voice and fluency parameters provide a direct and objective window into the behavioral expression of anxiety during PS. These measures not

1 Although the boundaries between stress and anxiety remain an open discussion (Daviu et al., 2019), in this article, we will refer to stress only as a physiological emergency state of an organism in response to a challenge to its homeostasis (Chrousos, 2009).

only capture subtle vocal modulations linked to arousal but also reflect how anxiety interferes with speech delivery. Therefore, analyzing these parameters is essential for understanding the impact of anxiety on PS.

Moreover, these adaptations in oral communication during discourse can influence the audience's perception of the speaker and impact their credibility, reputation and career success (Niebuhr et al., 2018; Wörtwein et al., 2015). However, despite the importance of voice and fluency in PS, there is little information in the literature about interventions that allow their improvement. A recent scoping review (Menjot et al., 2023) highlighted substantial heterogeneity in the types of interventions proposed as well as inconsistencies in the measurements (e.g., quality of source recording, lack of statistical analysis). This scarcity of robust evidence underscores the need for scientifically grounded research focusing on measurable vocal and fluency outcomes. As PS skills can be acquired and improved through training (Morreale and Pearson, 2008), it is crucial to foster the development and implementation of scientifically grounded interventions (Robles, 2012).

In the treatment of anxiety-related disorders such as PS anxiety, the most common and validated approach is exposure therapy (Yang et al., 2019). The approach posits that cognitive, physiological and behavioral symptoms associated with feared situations (here PS) can be targeted by repeatedly confronting people with said situations in the absence of the anticipated harm (Craske et al., 2014). Nevertheless, due to logistical and temporal limitations, it is not always feasible to repeatedly expose individuals to particular scenarios, such as speaking to an audience. Virtual reality (VR), as a computer generated immersive three-dimensional environment, could therefore be a useful tool to expose people to feared situations in an accessible, customizable and safe way (Craske et al., 2014; Kaplan et al., 2021; Scheveneels et al., 2019). Nevertheless, there are some important characteristics that have to be assessed before using any virtual environment: sense of presence and cybersickness. Users must feel a sufficient sense of presence, i.e., the psychological perception of being within the virtual environment (Heeter, 1992; Slater et al., 1994), in order to enable them to feel, think and behave as if they were in reality. Moreover, cybersickness (i.e., negative side effects of VR like nausea, vertigo or headaches caused by a conflict between sensory information from real and virtual settings) has to be limited (Heeter, 1992; Slater et al., 1994). These two key concepts of VR can be influenced by multiple underlying determinants related to VR simulation (display resolution, head tracking quality) or psychological characteristics of the person immersed (immersive tendencies, personality) (Felton and Jackson, 2022). Presence, for example, is known to be positively correlated with anxiety meaning people with high anxiety level tend to feel more immersed and present in VR (Maymon et al., 2024). According to these authors, the sense of being situated in reality (presence) is a reflective feeling informed by the extent to which changes in our subjective emotional state are experienced (anxiety).

Overall, VR has considerable potential for skill development and offers the possibility of including key components, such as customization of the task and adjustment of the difficulty progression while fostering interest and motivation (Demers et al., 2021; Gruber and Kaplan-Rakowski, 2020). In the specific case of PS, VR interventions have been shown to significantly reduce anxiety in clinical (Lister, 2016; Yuen et al., 2019; Zacarin et al., 2019) and non-clinical individuals (Daniels et al., 2020). VR has also proven to be as efficacious as *in vivo* training to alleviate PS anxiety (see Reeves et al., 2022 for a meta-analysis). As PS is of paramount importance in educational context (Daniels et al., 2020), several intervention studies have focused on student population. Stupar-Rutenfrans et al. (2017) found a decrease in self-reported anxiety level in their 35 university students after three training sessions with mobile 360° video (empty classroom, small audience and large audience). LeFebvre et al. (2021) reported a significant reduction of PS anxiety following two sessions of VR training in a group of 17 university students. The same results were observed with 27 secondary school students and a 90-min therapist-led VR training in Kahlon et al. (2019). Harris et al. (2002) further strengthened these conclusions by showing a reduction of self-reported anxiety as well as physiological measures of anxiety (HR) in university students after four VR PS sessions. Takac et al. (2019) observed the same results after three VR immersion but, contrary to their hypotheses, no changes were found in speech anxiety (SUDS and HR). However, anxiety reduction is not observed in all VR PS training. For example, Davis et al. (2020) found a significantly higher self-reported PS anxiety in their 195 university students after one session of VR training compared to training in front of a peer.

The above studies main goals were to reduce PS anxiety but other studies also focused on PS skills development. For example, Van Ginkel et al. (2020) compared the effect of VR-assisted versus instructor-assisted training and feedback between two groups of 11 pre-university students. Authors concluded both types of training were effective in increasing eye contact and speech rate thus improving PS quality. Boetje and van Ginkel (2021) also reported that training in VR twice after feedback was more beneficial for oral skills than training once. More recently, Remacle et al. (2023) conducted a two-arm randomized controlled trial in teachers: while the first group attended an information session and three VR training sessions, the second group only attended an information session (control). Results indicated a greater positive effect on vocal self-efficacy for the VR group than for the control group but the study did not demonstrate significant changes in objective voice measures following training. As previously stated, research investigating PS voice and fluency skills in addition to anxiety reduction is scarce. Kryston et al. (2021) found an improvement of PS quality ratings after VR training in a sample of 140 university students but no difference was found for PS anxiety. In secondary school settings, Valls-Ratés et al. (2022) compared PS skills of 50 students after three training sessions either in VR or alone. Results showed a reduction of self-assessed anxiety for both groups but an improvement of voice quality (resonance) only for the VR group. Another study by the same authors (Valls-Ratés et al., 2023) compared VR PS training with or without encouraging the use of gestures. After three sessions, all 70 students had significantly reduced their self-assessed anxiety and improved in persuasiveness and charisma ratings. However, no difference between groups were found in objective prosodic and gestural features. In multiple studies, VR has also been shown to be positively accepted by participants (Monteiro et al., 2020; Palmas et al., 2019). Moreover, commercial VR tools such as Chiara, Ovation, Virtual Orator and VirtualSpeech use VR combined with AI to provide feedback on public speaking (e.g., volume, speech rate body language, eye contact). However, these tools generally do not disclose how their metrics are computed, which limits transparency and prevents users from fully understanding the rationale behind the feedback.

In summary, there are encouraging results regarding the improvement of global communication skills after VR training (Schmid Mast et al., 2018). However, despite these promising outcomes, most VR-based PS studies did not conduct a rigorous evaluation process of environments to ensure the simulation captured fundamental features of the real task and environment, and elicited realistic behaviors (Harris et al., 2020). Although validation of the VR environment using evidence-based methods is required to maximise the efficiency and effectiveness of VR for skill development, this is not yet common practice. VR environments assessments too often rely on visual features that are not key determinants of successful training outcomes (Harris et al., 2020). Pioneers work in PS VR studies, like Slater et al. (1999) and Pertaub et al. (2001); Pertaub et al. (2002), focused mainly on ensuring anxiety elicitation in front of different types of virtual audience (negative and positive). More recently, some studies have also shown VR PS tasks could provoke physiological anxiety (HR) in addition to self-reported anxiety (e.g., Felnhofer et al., 2014; Slater et al., 2006). Kothgassner et al. (2016) further expanded this validation process by comparing real and virtual environments. In this study, similar self-reported and physiological (HR and salivary cortisol secretion) anxiety responses were observed during a 5-min PS task performed either in front of a real audience, a virtual audience or an empty virtual lecture hall.

The above studies focused on self-reported or physiological indicators of anxiety, while the communicative dimensions—particularly objective voice and fluency parameters—remained overlooked. Yet, these behavioral indicators are fundamental to understanding whether virtual environments truly elicit realistic performance. To our knowledge, in the PS context, only three studies validated their VR environment by comparing oral communication in front of a real and a three-dimensional virtual audience before proposing it for training. In the first, Notaro et al. (2021) found that participants speaking in VR were evaluated as producing more intonation and a louder voice as well as to exhibit a reduction in objective speech rate compared to the same speech performed in front of a live audience. Remacle et al. (2021) demonstrated that a virtual classroom was able to induce teacher's vocal features similar to those used in a real classroom in terms of fundamental frequency (median and variation) and intensity level compared to a free speech condition. In youth, Moïse-Richard et al. (2021) showed that self-reported anxiety and fluency ratings (self- and hetero – assessed Likert scale) when talking in front of a virtual class did not differ from those observed when talking to a real class. Fluency ratings in the two talks were also significantly higher than when talking in an empty virtual apartment. Moreover, this study revealed real audience elicited higher anticipatory anxiety than the virtual class. However, the generalization of these study results should be made with caution as they are mostly based on subjective data (Moïse-Richard et al., 2021; Notaro et al., 2021) coming from a relatively small sample size ($N \leq 30$).

In this regard, the present validation study set out to provide additional results and a more comprehensive approach by examining anxiety indicators as well as objective oral communication indicators (voice and fluency). This multimodal approach constitutes a novel and necessary step toward validating VR environments that not only reproduce anxiety responses but also capture authentic communicative behaviors. Therefore, our study aims to validate a virtual audience in terms of anxiety in the first place and in terms voice and fluency parameters in the second place. By bridging physiological and communicative levels of analysis, the present work advances VR validation research and lays the groundwork for evidence-based PS training interventions. In order to achieve this goal, anxiety, voice and fluency parameters during a PS task were compared in three conditions: (1) a real meeting room with an audience, (2) a virtual meeting room with an audience, and (3) the same virtual meeting room without any audience. Our hypotheses are that in order to validate the virtual audience for PS, participants' anxiety, voice and fluency parameters should be more similar in the two audience conditions (real and virtual audience) compared to the virtual empty condition without anxious stimulus. Indeed, as VR technology may *per se* induce a significant amount of stress it is crucial to control its influence (Hartanto et al., 2014). We therefore introduced an additional condition – the virtual meeting room without an audience – to assess the specific effect of the audience by controlling for the effect of the virtual device and simulation. This condition will enable us to single out the effect of the audience on PS among other non-social stressors related to VR immersion (Kothgassner et al., 2014; Von der Pütten et al., 2010). The corresponding research questions are the following:

RQ1: Are anxiety, voice and fluency parameters comparable between PS in front of a real audience and PS in front of a virtual audience?
RQ2: Are there differences regarding anxiety, voice and fluency parameters between PS in front of a virtual audience and PS in front of an empty virtual meeting room?

Moreover, as these are crucial aspects of VR validation and implementation (Harris et al., 2020), another goal of the study was to document the extent to which the virtual audience elicited a strong feeling of presence without generating cybersickness. This verification will ensure that this VR environment is acceptable and can be later used in PS training and interventions.

RQ3: Is immersion in the virtual audience acceptable and usable for PS purposes?

Finally, as VR immersion can be mediated by multiple underlying determinants (Felton and Jackson, 2022; Maymon et al., 2024), we examined whether and how our participants' initial trait anxiety levels, presence and immersion could influence their reactions to the virtual environments. This research question might allow us to distinguish different profiles of participants regarding VR immersion and ultimately identify profiles that are most suited for future PS VR interventions using these environments.

RQ4: Are there different profiles regarding VR immersion in our participants?

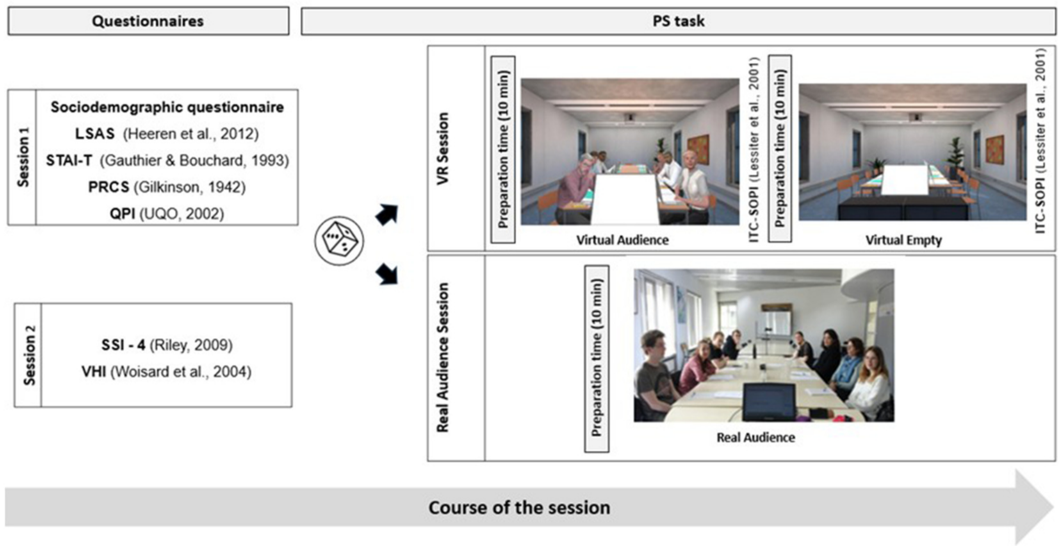

FIGURE 1
Experiment procedure.

2 Materials and methods

2.1 Ethical statement

This study was approved by the Ethics Committee of the Faculty of Psychology, Speech-Language Therapy and Educational Sciences of the University of Liège–Belgium in March 2023 (file number: 2223-059/5559). It was conducted following the ethical standards described in the Declaration of Helsinki (1964). All participants provided written informed consent after receiving a complete description of the study.

2.2 Participants

The sample included 60 university students (34 females and 26 males) with a mean age of 21.03 years (SD = 6.4 years). Inclusion criteria were being a native French speaker and being a first- or second-year university student. By recruiting native French speakers, we ensured that anxiety and oral communication changes (voice and fluency) in PS were not a result of a lack of language proficiency. Moreover, recruitment of newly enrolled university students prevented heterogeneity in PS experience due to academic requirements. As this study focuses on non-pathological individuals, exclusion criteria were reporting of a voice, fluency or hearing disorder at the time of the experiment. The absence of voice and fluency disorders was controlled using tasks described in the material and measures section of this article (see Section 2.4.).

2.3 Procedure

The experiment took place in two sessions 1–14 days apart. One of the two sessions was dedicated to the real meeting room with an audience condition, and the other one was dedicated to the immersions in the virtual environments (Virtual meeting room with an audience and Empty virtual meeting room). Each session lasted for approximately 2 hours. To control for potential order effects, participants were counterbalanced: half of the participants were exposed to the real audience first, while the other half were exposed to the virtual environments first (either the virtual meeting room with an audience or the empty virtual meeting room first, in a randomized order). At the beginning of their first session (whether real or virtual), besides sociodemographic information, all participants completed anxiety questionnaires (Personal Report of Confidence as a Speaker scale, Gilkinson, 1942; Liebowitz Social Anxiety Scale, Heeren et al., 2012; State-Trait Anxiety Index – Trait subscale, Gauthier and Bouchard, 1993) and a questionnaire about their tendencies to experience presence in common activities (Immersive Tendencies Questionnaire, Robillard et al., 2003). At the beginning of the second session, participants completed the assessment related to exclusion criteria (Stuttering Severity Instrument–4, Riley, 2009; Voice Handicap Index, Woisard et al., 2004) (Figure 1). In both sessions, the PS task took place in a quiet room measuring 10 by 8 m. In the real meeting room, participants stood facing the audience who was sitting 2 m away.

2.3.1 Real audience session

The real audience consisted of 8 people (4 women and 4 men) sitting on chairs displayed around a rectangular table in the meeting room used to design the virtual environments. The audience was asked to be neutral, i.e., neither explicitly distractive nor supportive. After entering the meeting room, participants were asked to rate their anxiety level before the PS task with the Subjective Unit of Distress Scale (*SUDS–Before*; Benjamin et al., 2010) (Figure 2). They were then asked to perform their PS task (see below). After the speech, participants had to rate their anxiety level with the *SUDS* twice: once to give their anxiety level after the task (*SUDS–After*) and

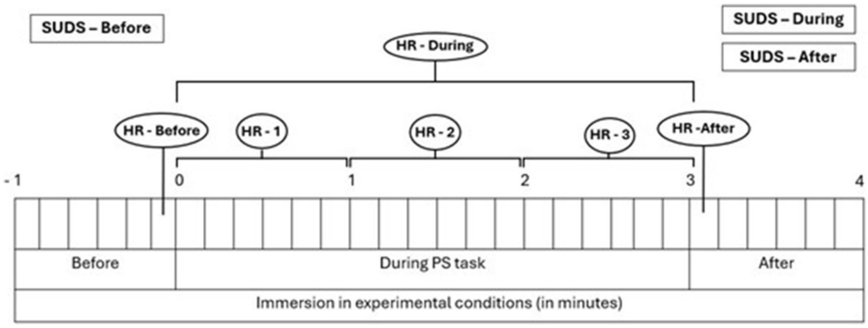

FIGURE 2
Moments of Evaluation of Anxiety Measures (in minutes).

FIGURE 3
VR equipment.

once to evaluate their mean anxiety level during the whole task (*SUDS–During*).

2.3.2 Virtual reality session

Participants were successively immersed in two virtual environments: a virtual meeting room with 8 animated virtual agents (Virtual Audience) and the same virtual meeting room without any audience (Virtual Empty). The virtual environments and agents were created by a team of three-dimensional animation specialists, technical artists, and three-dimensional game artists working in a research lab at our university (LabXR team-HEC Liège). The virtual audience condition was implemented to be as similar as possible to the real audience condition. In the virtual audience condition, 8 virtual agents (4 women and 4 men) were sitting on chairs around a rectangular table positioned in a meeting room. The participant stood in front of this rectangular table where we could find a screen displaying presentation slides and a timer. The slides were also visible on a projection screen behind the participant. Virtual agents' posture, facial expressions and head movement could be modified in order to express different levels of valence (i.e., positive or negative feelings) and arousal (i.e., alertness). The perception of these nonverbal behaviors was previously validated in Etienne et al. (2023). In the current study, virtual agents were calibrated to express medium levels of valence and arousal meaning they looked mainly at the participant and were neither explicitly distractive nor supportive.

The virtual environment software was installed on a PrimInfo laptop (GT73VR 7RF-491FR, Eindhoven). The PC was connected to an HTC Vive Pro Eye Office 1 head-mounted display (HMD) and to two controllers (HTC Controller Vive Pro; see Figure 3). The HMD enclosed the upper portion of the participant's face and was equipped with individual display screens for each eye, along with a tracking system to determine its orientation. No sound was included in the VR immersion. The participants viewed the environments through the HMD, while the researcher observed their perspective on a computer monitor.

After headset calibration at each VR immersion, participants were asked if they felt sick, nauseous or dizzy to control for cybersickness. Then, participants rated their anxiety level before the immersion with the *SUDS* (*SUDS–Before*). They were then asked to perform the PS task. After the PS task, participants had to rate their anxiety level with the *SUDS* twice: once to give their anxiety level after the task (*SUDS–After*) and once to evaluate their mean anxiety level during the task (*SUDS–During*) (Figure 2). Participants also completed the ITC-Sense of Presence Inventory (ITC–SOPI; Lessiter et al., 2001) after their immersion in the virtual environments. Finally, the participants rested for approximately 5 min and then prepared for the other PS task. *SUDS* and ITC–SOPI were completed according to the same procedure. After the last immersion, each participant was asked to provide feedback regarding their experience in the virtual meeting room and to provide suggestions for future developments.

2.3.3 PS tasks

There were three given themes for the three experimental conditions. Themes participants were asked to present were related to sports, ecology or cultural organizations from our university. The themes' prior knowledge was considered relatively weak and equivalent among newly enrolled students. The themes were counterbalanced between participants and conditions to prevent any impact on anxiety, voice and fluency measures. The preparation process was consistent across all conditions: 10 min before the PS task, the experimenter provided the speech theme, a set of slides, and a one-page informative text to the participant. These materials (text and slides) ensured all participants were put in the same conditions regarding themes' knowledge regardless of their memory capacity. Participants were informed that they could only refer to the slides while presenting and that no personal notes were allowed. After the preparation and just before the PS task, the theme mastery was self-rated by the participant on a 0 (absence of knowledge) – 100 (perfect mastery) scale. Although they were initially asked to prepare for a 5-min talk, most of the participants stopped speaking before this time threshold. In order to make the PS task comparable across participants and conditions, we only used the first 3 minutes of the talks in the latter analyses.

2.4 Material and measures

2.4.1 Anxiety questionnaires

Participants completed three self-report scales aimed at assessing their trait anxiety.

First, the French version of the short form of the Personal Report of Confidence as a Speaker scale (PRCS; Heeren et al., 2013) includes 12 true-or-false statements assessing participants' confidence in their PS abilities (e.g., "I am in constant fear of forgetting my speech"). Scores range from 0 to 12, with higher scores indicating lower confidence in PS skills.

Second, the French self-report version of the Liebowitz Social Anxiety Scale (LSAS-SR, Heeren et al., 2012) contains 24 items measuring fear and avoidance in social and performance situations. For each of the 24 situations, participants rate the intensity of fear they experience on a first 4-point Likert scale (0 = none; 3 = severe), and the frequency of avoidance of the situation on a second 4-point item scale (0 = never; 3 = usually). The total score is calculated by summing the ratings across all items, with higher scores reflecting greater levels of social anxiety.

Third, the French version of the State-Trait Anxiety Index – Trait subscale (STAI–T, Gauthier and Bouchard, 1993) was used to gain insight into the participant's trait anxiety regardless of the situation. Participants were asked to indicate how they usually feel (e.g., "I lack self-confidence") using a 4-point Likert scale (1 = never; 4 = almost always). The overall score is determined by the sum of the scores obtained for the 20 items with higher scores indicating greater anxiety.

2.4.2 Voice questionnaire

The French version of the Voice Handicap Index (VHI) was used to screen participants for any voice disorders. This index measures the biopsychosocial impact of voice problems (Woisard et al., 2004). It comprises 30 self-rated items on a 5-point Likert scale (0 = never; 4 = always). The total score ranges from 0 (no complaints) to 120 (many complaints). The maximum cut-off score for normality in a population without voice disorders is 20 (Woisard et al., 2004).

2.4.3 Fluency assessment

The Stuttering Severity Instrument-4 (SSI-4, Riley, 2009) was administered to exclude a speech fluency disorder (stuttering) in our participants. The assessment consists of a speaking and a reading task used to evaluate fluency in productions of at least 200 syllables. A total score is computed by the addition of three subscales: one measuring the percentage of stuttered disfluencies (i.e., repetitions of sounds, syllables, blocks and prolongations), one measuring the duration of the three longest stuttering episodes (in seconds) and one measuring the associated physical manifestations (i.e., distracting sounds, grimaces, head movements, body movements). The total score is then converted into percentile ranks to assess the severity of stuttering. The maximum cut-off score for normality in an adult population is 10 (less than first percentile). For the speaking task of this study, participants were asked to describe a typical day from morning to evening, as accurately as possible, in 5 min. For the reading task, they were asked to read aloud the phonetically balanced text called "La bise et le soleil" (Abercrombie, 2013).

2.4.4 Virtual reality questionnaires

A French version of the Immersive Tendencies Questionnaire (ITQ) was completed by participants. This questionnaire assesses individuals' general tendency to experience a sense of presence during everyday activities (Witmer and Singer, 1998). The ITQ includes 18 self-report items rated on a 7-point Likert scale (0 = never; 6 = often), yielding a total score as well as four subscale scores: (1) Focus, reflecting the ability to concentrate on current activities and ignore distractions; (2) Involvement, indicating the extent to which one becomes absorbed in stories and movies; (3) Emotions, measuring the intensity of emotional responses to stimuli such as films; and (4) Games, assessing the frequency of video game use (Robillard et al., 2003).

After their immersion in the virtual environments, the participants filled in a French version of the ITC-Sense of Presence Inventory (Lessiter et al., 2001). The ITC-SOPI is a self-report questionnaire about the user's experience in the virtual environment. It contains 44 items, scored on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree), and measures 4 dimensions: (a) Spatial Presence (sense of physical placement, interaction, and control over different parts of the environment; e.g., "I felt that the characters and/or objects could almost touch me"); (b) Engagement (feeling of being psychologically involved and enjoying the content; e.g., 'I felt myself being "drawn in"); (c) Ecological Validity/Naturalness (tendency to consider the mediated environment as lifelike or real; e.g., "The content seemed believable to me"); and (d) Negative Effects (displaying adverse physiological reactions; e.g., "I felt disorientated"). In order to compare negative symptoms felt (i.e., nausea or oculomotor problems) before and after VR immersion, participants were also asked to complete the Negative Effect subscale at the beginning of the VR session.

2.4.5 Anxiety measures

To gain insight into state anxiety, participants provided a Subjective Unit of Distress Scale (*SUDS*) rating at different times. The *SUDS* is a commonly used, reliable measure of self-reported anxiety (Benjamin et al., 2010). It typically ranges from 0 to 100 with higher ratings indicating greater anxiety. *SUDS* levels were taken at the beginning of each study session (*SUDS-Baseline*) and then three times for each condition. The first time to gain insight into anxiety levels a few seconds before the PS task (*SUDS–Before*) then twice after it: once to get anxiety levels after the task (*SUDS–After*) and once to get mean anxiety level felt during the task (*SUDS–During*).

Moreover, as acute stress responses are usually assessed with *HR* measures (Bitsika et al., 2014), we used the wrist-worn Polar Verity Sense with a photoplethysmography technique to gain insight into physiological anxiety. Polar Verity Sense demonstrates high reliability for laboratory experiments that do not involve excessive physical movement (Gil et al., 2021). *HR* (in bpm) was extracted every second during each PS task. We then calculated an average global *HR* bpm per PS task (*HR–During*). In order to track HR dynamics, we also calculated average *HR* bpm for each minute of the talks (*HR–1*; *HR–2*; *HR–3*) as well as 10 s before (*HR -Before*) and 10 s after the PS task (*HR–After*; see Figure 2). As a reminder, high *HR* values indicate high bpm and high physiological stress, whereas low *HR* values reflect less bpm and lower physiological stress.

All anxiety measures (*SUDS* and *HR*) were taken while facing the PS environment (whether real or virtual) with or without an audience.

2.4.6 Voice and fluency measures

Participants' PS tasks were recorded for subsequent analysis using a head-mounted condenser microphone (C 544 L, AKG Acoustics GmbH, Austria) positioned 5 cm from the mouth. The microphone was connected to a Priminfo laptop (GT73VR 7RF-491FR, Eindhoven) via an external sound card (iTrack Solo, Focusrite Audio Engineering Ltd., China). Audio was recorded in WAV format with a sampling frequency of 44.1 kHz and a 16-bit resolution. To ensure recording quality, the signal-to-noise ratio (SNR) was calculated by comparing the sound pressure levels of the voice signal to the background noise. All SNRs met the recommendations outlined by Barsties and de Bodt (2015). Acoustic analyses of the recorded speech were then conducted using Praat freeware (version 6.4.23).

In order to compare voice and fluency characteristics between PS tasks, we calculated an average value of parameters known to be impacted by PS anxiety for each task (see introduction Section 1.).

For voice production, two main measures were extracted from connected speech:

- Median of the fundamental frequency (f_0-in Hz) using an autocorrelation method,
- Variation of the fundamental frequency through the standard deviation of the fundamental frequency (f_0 SD–in Hz).

TABLE 1 Description of the sample of 60 university students enrolled in the study and comparison with the normative data.

	Study		Norms	Comparison	
	Range	M (SD)	M (SD)	T	*p*
PRCS	0–12	5.52 (3.38)	7.30 (3.58)	4.08	<0.01
LSAS-SR	9–99	55.38 (20.52)	52.81 (22.22)	0.97	0.33
STAI-T	33–61	30.5 (17.46)	40 (9)	13.13	<0.01
ITQ	51–107	78.12 (12.49)	64.11 (13.11)	8.69	<0.01
Focus	14–32	23.45 (4.53)	24.81 (7.54)	2.33	0.02
Involvement	10–35	22.9 (4.91)	15.33 (8.67)	11.95	<0.01
Emotions	8–26	16.97 (5.09)	14.25 (6.7)	4.13	<0.01
Games	3–21	10.03 (4.31)	6.56 (4.95)	6.24	<0.01
VHI	6–66	21.78 (12.47)			
SSI	6–8	6.18 (0.47)			

Speech transcription and disfluency coding followed standard procedure. A speech therapy master's student, in her final year of training, listened to the entire recordings of each PS task. The CLAN (Computerized Language Analysis) software, part of the CHILDES system (Child Language Data Exchange System; MacWhinney, 2000), was used to analyze disfluencies based on transcriptions of speech samples in CHAT (Codes for the Human Analysis for Transcripts) format. To determine inter-judge reliability, 20% of the speech samples were randomly selected and analyzed by the first author (L.B.). Spearman correlation coefficients between the scores of both judges were very high (rs = 0.96). Point-to-point percent reliability was 99.14%.

For speech fluency, two main measures were considered:

- The total percentage of disfluencies (*TD_Percentage*) meaning the total number of non-stuttering like disfluencies such as repetitions, revisions of words or sentences and filled pauses, (i.e., um, uh) calculated on the total number of syllables produced.
- The total percentage of filled pauses (*Filled pauses_Percentage*) calculated on the total number of syllables produced.

2.5 Statistical analyses

To answer our first two research questions, repeated measures ANOVAs were conducted with the three conditions as a within-subject factor. The Bonferroni correction was applied regarding the six main measures considered for the validation, two for each aspect of the PS task (anxiety, voice and fluency). When the p-value was lower than 0.008 ($\alpha = 0.05/6$), the null hypothesis of equal means was rejected and simple orthogonal contrast analyses were performed to test (1) the difference between the virtual audience and the real audience conditions, and (2) the difference between the virtual empty and the virtual audience conditions. The first contrast answers our first research question (RQ1). It aims to single out the effect of VR by assessing if anxiety, voice and fluency parameters are comparable when speaking in front of a real audience and a

TABLE 2 Group means, standard deviations, and range of the SUDS ratings for the different conditions and moments.

Condition	Real audience				Virtual audience				Virtual empty			
Moment	Baseline	Before	During	After	Baseline	Before	During	After	Baseline	Before	During	After
Mean	28.8	50.5	52.9	47	26.2	26.4	30.4	26	26.2	22.3	28.8	26.3
SD	19.5	24.4	26	26.6	21.3	20.6	21.1	22	21.3	19.1	24.1	23.8
Range	0–70	0–90	0–100	0–100	0–90	0–70	0–80	0–80	0–90	0–65	0–80	0–75

virtual audience. The second contrast answers our second research question (RQ2) and enables us to single out the effect of the audience on PS while controlling the effect induced by the virtual device and simulation. As changes in PS oral communication depend mainly on anxiety and related measures are extensively used to validate PS virtual environment (Kothgassner et al., 2016), we performed more detailed analyses of anxiety parameters in our ANOVAs. We therefore investigated the dynamics of psychological (*SUDS*) and physiological anxiety (*HR*) with measurements taken before, during and after PS task (see Figure 2). In regard to our third research question related to VR acceptability and usability (RQ3), we conducted paired t-tests on ITC–SOPI scores of the two virtual environments. Comparisons with the norms of the questionnaire were also performed. Finally, to answer our fourth research question and ultimately distinguish different profiles of participants (RQ4), we performed an ascending hierarchical classification on parameters known to influence VR immersion (trait anxiety: STAI-T, LSAS-SR, PRCS - presence: ITC-SOPI - Felnhofer et al., 2014; Ling et al., 2014) and on the main outcome measures of this study (*SUDS*, *HR*, *TD _Percentage*, *filled pauses _Percentage*, f_0 median and SD).

The applicability of the analyses was checked in advance, particularly normality using the Shapiro-Wilk test and the assumption of homogeneous variances using Mauchly's sphericity test. When the sphericity hypothesis was violated and epsilon value was below 0.75, the Greenhouse-Geisser correction adjusting the degrees of freedom is reported. Analyses were performed using the Jamovi software (version 2.3.28.) for the first three research questions and using the R software (version 4.3.3.) for the fourth research question.

3 Results

3.1 Description of the participants

In regard to the exclusion criteria (Table 1), none of the participants self-reported a voice disorder at the time of the study. Although 26 participants exceeded the VHI cut-off score for normality (=20; Woisard et al., 2004), after extensive analysis by a specialized speech and language pathologist (L.B.), it seemed that these participants did not differ from the other participants in their acoustic measures. We therefore decided to keep them in the study sample for the rest of the analyses. Regarding SSI–4 scores, as all participants were under the first percentile, none of the participants had a speech fluency disorder. Moreover, none of the participants self-reported a fluency disorder at the time of the study. Concerning anxiety questionnaire (PRCS, LSAS–SR, STAI–T), as our participants' scores were significantly lower than norms (Table 1), they did not have high social anxiety or poor confidence as speakers.

Regarding VR measures, our participants were not very familiar with VR considering only 35% of participants had tried an HMD more than three times. However, as shown in Table 1, participant's immersive tendencies were significantly higher compared to the normal range (Robillard et al., 2003). In other words, our participants seemed to experience more immersion than the general population. However, as norms of this questionnaire are

TABLE 3 Within-subject effect of repeated measures ANOVAs concerning anxiety scores during the talk (Greenhouse–Geisser correction applied).

		Sum of squares	df	Mean square	F	p	η^2_p
SUDS	Condition	53,309	1.27	42,098	48.1	<0.001	0.44
	Moment	9717	2.06	4707	12.5	<0.001	0.17
	Condition * moment	13,962	3.50	3993	21.0	<0.001	0.26
HR	Condition	7,183	1.49	4826.6	13.83	<0.001	0.19
	Moment	1,172	1.54	761.3	11.49	<0.001	0.16
	Condition * moment	101	2.96	34.2	0.63	0.59	0.01

TABLE 4 Contrast analysis concerning anxiety scores before, during and after the talk.

		Contrast	T	p
Before	SUDS	Real audience VS virtual audience	8.22	<0.001
		Virtual audience VS virtual empty	2.95	0.005
		Real audience VS virtual empty	8.71	<0.001
	HR	Real audience VS virtual audience	3.22	0.75
		Virtual audience VS virtual empty	0.28	1
		Real audience VS virtual empty	3.49	0.33
During	SUDS	Real audience VS virtual audience	6.74	0.001
		Virtual audience VS virtual empty	0.82	0.4
		Real audience VS virtual empty	6.60	0.001
	HR	Real audience VS virtual audience	3.84	0.011
		Virtual audience VS virtual empty	4.74	<0.001
		Real audience VS virtual empty	1.19	1
After	SUDS	Real audience VS virtual audience	6.13	0.001
		Virtual audience VS virtual empty	−0.14	0.88
		Real audience VS virtual empty	6.21	0.00
	HR	Real audience VS virtual audience	2.99	0.14
		Virtual audience VS virtual empty	3.24	0.07
		Real audience VS virtual empty	0.94	1

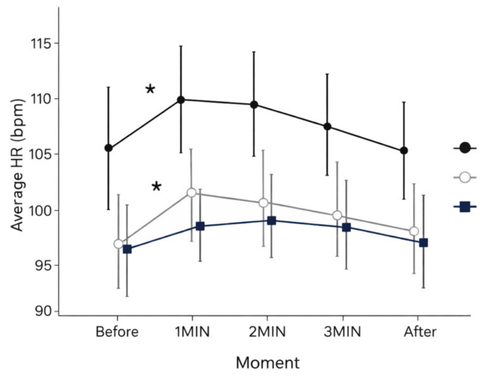

FIGURE 4
Temporal fluctuations of HR across conditions (mean and SD).

TABLE 5 Contrast analysis concerning HR scores before the talk and in the first minute of the talk.

	Condition	T	p
HR-before VS HR-1	Real audience	−5.17	<0.001
	Virtual audience	−5.39	<0.001
	Virtual empty	−2.99	0.41

relatively old (Robillard et al., 2003), they might not reflect actual immersive tendencies of the population.

3.2 Statistical analyses related to research questions

3.2.1 Anxiety

Regarding anxiety parameters (*SUDS* and *HR*), we conducted repeated measures ANOVA with two within subject factors: the moment of the evaluation and the condition. When significant differences were found in our ANOVA, we then performed contrast analyses. For the sake of readability, only significant differences of these contrast analyses are developed in the section below.

Prior to these analyses, we ensured participants' baseline *SUDS* ratings did not significantly differ from one session to the next by performing a paired t-test between sessions (T = 0.47, $p = 0.64$). A summary of *SUDS* ratings can be found in Table 2.

Repeated measures ANOVA conducted on *SUDS* ratings (Table 3) revealed significant effects of condition and moment. The interaction between condition and moment was also significant. Regarding *HR* (Table 3), significant effects of condition and moment were found. Nevertheless, no significant effect of interaction was observed.

Anxiety just before the talk (*SUDS-Before and HR-Before*; Table 4): Contrast analysis test showed that the *SUDS-Before* ratings in the real audience condition were significantly higher than in the virtual audience condition (RQ1). Moreover, *SUDS–Before* ratings for the virtual audience condition were also significantly higher than for the virtual empty condition (RQ2). Finally, the *SUDS-Before* ratings in the real audience condition were significantly higher than in the virtual empty condition. However, regarding *HR–Before* no significant differences were observed between audience conditions (RQ1- RQ2). Looking more closely at temporal fluctuations in *HR* within condition (Figure 4 – Table 5), we observe significant differences between *HR–Before* and *HR–1* in the two audience conditions: average *HR* rises at the beginning of PS in both the real and the virtual audiences. These results were not reproduced at the beginning of PS in the virtual empty condition.

Anxiety during and after the talk (*SUDS-Before*, *SUDS-After* and *HR-During*; Table 4): Contrast analysis revealed *SUDS–During* and *SUDS–After* ratings were significantly higher when talking in front of a real audience than when talking in front of a virtual audience (RQ1) or an empty virtual meeting room. Contrary to our expectations, the *SUDS–During* and *SUDS–After* ratings between the virtual audience and the virtual empty condition did not differ significantly (RQ2). Regarding *HR*, no significant differences were found in *HR-During* between the real and virtual audience conditions (RQ1). Contrast analysis revealed that *HR–During* was significantly higher when talking in front of a real audience than when talking in front of empty virtual meeting room but these conditions became similar again in *HR-After*. *HR–During* measures between the virtual audience and the virtual empty condition did not differ significantly (RQ2).

3.2.2 Fluency

In order to compare speech fluency between experimental conditions, two main measures were considered: the total percentage of disfluencies (*TD _Percentage*) and the total percentage of filled pauses (*Filled pauses_ Percentage*). We conducted our analyses on the average value of each of these fluency parameters over the total duration of the PS task (i.e., 3 min).

Regarding *TD_Percentage*, as Bonferroni correction applied, no significant differences were found between conditions (F = 3.36; p = 0.038, η^2_p = 0.05 – RQ1 – RQ2). Participants produced a little bit more than 10% of disfluencies regardless of the condition (real audience: M = 12.4; SD = 4.86; virtual audience: M = 11.4; SD = 4.53; virtual empty: M = 11.6; SD = 4.05).

Concerning *Filled pauses_Percentage*, significant differences were found between conditions (F = 7.25; p = 0.001, η^2_p = 0.11). We observed a significant difference between the two audience conditions (T = 3.64; p < 0.001 – RQ1) as well as between the real audience and the virtual empty condition (T = 2.79; p = 0.007). The performance in front of the real audience is the one that elicited the most filled pauses (M = 10.6; SD = 4.51) followed by the two virtual conditions for which no significant difference was found in the contrast analyses (virtual audience: M = 9.42; SD = 4.06 - virtual empty: M = 9.58; SD = 3.49 – RQ2).

3.2.3 Voice

In order to compare the voice production between the experimental conditions, two main measures were extracted from connected speech: median and standard deviation of the f_0 (in Hz). We conducted our analyses on the average value of each of these voice parameters over the total duration of the PS task (i.e., 3 min).

Regarding median f_0, as sphericity condition was violated and epsilon value was below 0.75, we used Greenhouse – Geisser correction. For this parameter, significant differences were found between conditions (F = 8.12; p = 0.002, η^2_p = 0.12). We observed a significant difference between the real audience and the virtual empty condition (T = −3.41; p = 0.004). However, no significant differences were found between the two audience conditions (T = -2.52; p = 0.044 – RQ1) and the two virtual conditions (T = 1.5; p = 0.419 – RQ2). The performance in front of the real audience is the one that elicited the highest-pitched voice followed by the virtual audience condition and then by the virtual empty condition.

Regarding standard deviation of f_0, as fundamental frequency variation was around 30 Hz for all experimental conditions, no significant differences were found (F = 2.19; p = 0.116, η^2_p = 0.04 - RQ1 – RQ2).

3.2.4 Virtual reality

As seen in Table 6, for both virtual environments, ITC–SOPI scores were significantly higher than mid-points for all subscales except for the negative effect subscale. Participants reported acceptable levels of presence and immersion, indicating that delivering a speech in the virtual meeting room—whether in the presence or absence of a virtual audience—was perceived as both believable and engaging. Additionally, they experienced low levels of negative effects, further suggesting that immersion in these virtual environments was acceptable (RQ3). In order to compare the virtual environments, paired t-tests between the ITC-SOPI dimensions of the virtual empty and the virtual audience condition were conducted (Table 6). Significant differences were found for spatial presence (T = 2.18; p = 0.03) and naturalness dimension (T = 2.56; p = 0.01) meaning that the virtual audience environment elicited a better spatial presence but less ecological validity than the virtual empty meeting room. No significant differences were found for engagement and negative effects (p > 0.05 - RQ3).

3.2.5 Group analysis

To answer our fourth research question (RQ4), we performed an ascending hierarchical classification from a combination of the participant's scores to the anxiety questionnaires (STAI-T, LSAS-SR, PRCS, mean *SUDS-During*) and, for each condition, the ITC-SOPI and mean *HR-During*, the *TD_Percentage* and *Filled pauses_ Percentage*, and the median and standard deviation of the f_0. To minimize bias, voice, fluency and *HR* measures were normalized by subtracting the mean values of the parameter in the three conditions and dividing by the standard deviation of these conditions. This data normalization is a pre-processing step preventing the domination of greater numeric feature values (such as median f_0) on smaller numeric features values (Singh and Singh, 2020). The Euclidean distance and Ward's aggregation index were used to cluster the 60 participants: a dendrogram represented the result of this classification. The number of groups in this classification was chosen considering the decrease in homogeneity between two successive clusters resulting from the algorithm, and more

TABLE 6 Group means, standard deviations, and comparison with the normative data for ITC–SOPI and paired t-tests comparing ITC–SOPI dimensions between the virtual audience and the virtual empty condition.

	Virtual audience				Virtual empty				Difference between environments	
	M study (SD)	M norms	T	*p*	M study (SD)	M norms	T	*p*	T	*p*
Spatial presence	58.65 (11.17)	47.5	7.73	<0.01	55.02 (14.52)	47.5	4.00	<0.01	2.18	0.03
Engagement	42.87 (7.05)	32.5	11.39	<0.01	41.57 (9.57)	32.5	7.34	<0.01	1.32	0.19
Naturalness	15.53 (3.77)	12.5	6.24	<0.01	16.73 (3.91)	12.5	8.37	<0.01	2.56	0.01
Negative effect	11.22 (4.85)	18	10.84	<0.01	11.55 (5.67)	18	8.81	<0.01	0.70	0.48

TABLE 7 Discriminating variables that significantly differed between the two groups of participants following the ascending hierarchical classification on participant's scores to the anxiety questionnaire (STAI-T, LSAS - SR, PRCS, mean SUDS scores during PS tasks) and, for each condition, the ITC-SOPI and mean HR scores, the TD_Percentage and Filled pauses_Percentage, and the median and standard deviation of the fundamental frequency.

Discriminating variables		p	Group 1 Mean	Group 2 Mean
Anxiety	STAI – T (/80)	<0.001	47.50	51.23
	LSAS (/144)	<0.001	45.50	65.16
	PRCS (/12)	<0.001	3.43	7.60
	Mean SUDS during presentation (/100)	<0.001	30.59	44.02
Presence	ITC–SOPI Naturalness subscale (/35)	0.02	15.5	17.90
	ITC–SOPI Engagement subscale (/65)	0.03	38.86	44.26
	ITC-SOPI Spatial Presence subscale (/95)	0.007	50.00	60.03
	ITC–SOPI Negative Effects subscale (/36)	<0.001	8.93	14.16
	ITC-SOPI Spatial Presence subscale (/95)	0.01	54.96	62.33
	ITC–SOPI Negative Effects subscale (/36)	<0.001	8.73	13.07
Speech fluency	TD_Percentage	<0.001	0.36 ^{n}	−0.34 ^{n}
	Filled pauses_Percentage	<0.001	0.36 ^{n}	−0.40 ^{n}
	TD_Percentage	<0.001	−0.42 ^{n}	0.20 ^{n}
	Filled pauses_Percentage	<0.001	−0.65 ^{n}	0.18 ^{n}

Virtual empty condition; Virtual empty condition; ^{n} = normalized values

concretely, looking at the plot of the gains of intra-category inertia. For this dataset, two groups appeared to be relevant to significantly distinguish between different types of profiles (RQ4).

Variables that allowed that distinction in two groups were the scores to the anxiety questionnaires (STAI-T, LSAS - SR, PRCS, mean *SUDS-During*), some subscales of the ITC-SOPI and *TD_Percentage* and *Filled pauses_Percentage* in the virtual conditions ($p < 0.001$; Table 7). The first group included 38 participants (18 females and 20 males) and was characterized by low anxiety level (STAI-T, LSAS - SR, PRCS, mean *SUDS-During*), low cybersickness (ITC–SOPI Negative effect subscale), globally low *TD_Percentage* (M = 11.54) and low *Filled pauses_Percentage* (M = 9.58) across conditions. Conversely, the second group included 22 participants (16 females and 6 males) and was characterized by significantly higher anxiety (STAI-T, LSAS - SR, PRCS, mean *SUDS-During*), more cybersickness (ITC–SOPI Negative effect subscale), a globally higher *TD_Percentage* (M = 12.25) and *Filled pauses_Percentage* (M = 10.37) across conditions. The feeling of presence in both virtual environments as well as the engagement and feeling of naturalness towards the virtual empty condition tended to be higher for this second group compared to the first group but didn't reach the significance level with Bonferroni correction applied ($p > 0.008$).

If we now focus on the interaction between conditions, based on the significant difference found with ascending hierarchical classification (Table 7), we can see that for the first group the virtual audience condition is the one that elicited the fewest disfluencies compared to empty virtual condition. For the second group, participants exhibited more disfluencies in the virtual audience condition than in the empty virtual condition. Possible explanations will be discussed in the discussion section.

10.3389/frvir.2026.1755571

4 Discussion

This study aimed to validate the use of a virtual audience for PS by comparing anxiety, voice and fluency parameters in three conditions: (1) a real meeting room with an audience, (2) a virtual meeting room with an audience, and (3) the same virtual meeting room without any audience. We wanted to find out whether the presence of a virtual audience, compared to an empty virtual meeting room (RQ2), triggered specific anxiety and oral communication changes (voice and fluency). We also examined whether these changes were comparable to those occurring with a real audience (RQ1). Another purpose of the study was to validate the virtual audience by ensuring it elicited a strong feeling of presence without generating unwanted negative side effects (i.e., cybersickness) (RQ3). Moreover, we investigated whether different profiles of participants regarding VR immersion could be identified (RQ4).

4.1 Anxiety

In this experiment, PS anxiety was assessed with *SUDS* and *HR* measures taken before, during and after the experimental conditions (Figure 2). Some of these measures enabled the validation of the virtual audience's capacity to elicit anxiety during PS.

First, self-report scores of anticipatory anxiety were higher before talking to the virtual audience than before talking in front of an empty virtual meeting room (*SUDS–Before*). As a reminder, participants rated their anxiety levels while facing the audience either real or virtual. The virtual audience itself thus seems to be able to generate levels of anxiety that cannot be attributed to the immersion in VR (Kothgassner et al., 2016 - RQ2). These findings suggest that the virtual audience itself effectively induces anxiety prior to public speaking. Second, the real and virtual audience successfully elicited anxiety during the first minute of the PS talk (*HR–1*) whereas the empty virtual meeting room did not generate such anxiety patterns. These results are concordant with previous literature that has shown self-reported and physiological PS anxiety could be elicited upon initial view of agent-based VR scenarios (e.g., Brundage et al., 2016; Felnhofer et al., 2014; Slater et al., 2006), even prior speech begins (Kothgassner et al., 2016; Takac et al. 2019).

Our results also show that self-reported anticipatory anxiety was higher when talking to a real audience in a real environment (RQ1) which is contrary to Kothgassner et al. (2016) but in accordance with Brundage et al. (2016) and Moïse-Richard et al. (2021) studies. This lower anticipatory anxiety in the virtual audience compared to the real audience could however be beneficial for PS training by making people more eager to accept to expose themselves to virtual PS scenarios. It may also help increase motivation and engagement in training programs, thereby supporting the integration of virtual audiences into intervention protocols (Lindner, 2021). However, these results regarding anticipatory anxiety were not completely reproduced with objective measures: no significant differences in HR 10 s before the talk were found between condition (*HR–Before*). Nevertheless, as an increase in HR between those 10 s (*HR–Before*) and the first minute of the PS talk (*HR–1*) was detected in both audience conditions (RQ1) but not in the virtual empty meeting room (RQ2), the pattern supports the assumption that the observed stress responses stem from the confrontation with an audience rather than from VR immersion alone.

Regarding anxiety during and after the talk, we did not consistently observe significant difference between condition both in self-report (*SUDS–During*; *SUDS-After*) and physiological response (*HR- During*; *HR–After*) (RQ1–RQ2). Real and virtual audience successfully elicited anxiety at the beginning of the PS talk (*HR–1*) but not in the latter minutes or after the task (*HR–2*; *HR–3*; *HR-After*). This attenuation may reflect the habituation phenomenon, a rapid adaptative stress response most non-anxious individuals can invoke while being repeatedly exposed to the same stimulus (Tsigos et al., 2020; Radley and Herman, 2023). This habituation phenomenon, reducing physiological and psychological impact of stressors, was also found in other studies involving psychosocial stressors (Finn and Sawyer, 2009; Kothgassner et al., 2021; Schommer et al., 2003). However, compared to its elicitation, anxiety habituation, particularly within speech, remains understudied although it could promote PS anxiety detection, intervention and prevention Takac et al., 2019. In PS research, Behnke and Sawyer (2001) identified two patterns. In subpattern one, anxiety peaks just before speaking and then declines during the PS task itself. In the second subpattern, anxiety rises just before speaking, peaks during the first minute of exposure to an audience, and then declines throughout the remainder of the talk. These authors also found that anxiety levels just before the beginning and in the first minute of the talk were significant predictors of overall PS anxiety whereas anxiety levels in the last minute of the talk were not (Behnke & Sawyer, 2004). As experimental conditions did not elicit high levels of anxiety, habituation might already impact our measurements during and after the presentation so that a return to baseline happened across conditions. Moreover, in order to avoid disrupting the PS task and impact presence, *SUDS–During* and *SUDS–After* ratings were taken both at the end of the PS task. This methodology might have impacted our results related to self-reported anxiety during and after the talk as the habituation phenomenon might already have occurred. Overall, the results from Behnke and Sawyer (2001); Behnke and Sawyer (2004). suggest the first minute of PS might be the most representative of PS anxiety and that this moment should benefit from more extensive examination in future studies. While not verified in every study (Brundage et al., 2016; Moïse-Richard et al., 2021; Takac et al. 2019), Behnke and Sawyer's habituation pattern theory within speech could explain why significant results were only found in the first moments of our PS task. All in all, our study fills a gap in literature by deepening our understanding of within speech habituation pattern and how they can impact oral communication (Takac et al. 2019).

4.2 Fluency

Concerning speech fluency, our results showed a significant impact of condition regarding the percentage of filled pauses, but not regarding the total percentage of disfluencies. More precisely, we found a significant difference in the percentage of filled pauses between the real audience and the two virtual environments (RQ1) but lacked distinguish the latter (RQ2). As results regarding fluency in VR have been heterogenous, interpretation of these findings remains complex. Indeed, although a significantly lower speech rate in VR, Notaro et al. (2021) observed similar filled pauses percentage between real and virtual audience. Moïse-Richard et al. (2021) found fluency ratings in PS were similar between real and virtual audiences but significantly higher than when talking in an empty virtual apartment. Conversely, Brundage et al. (2016) observed no significant difference in the total percentage of stuttered syllables among participants when speaking to a virtual audience compared to an empty virtual meeting room, despite significant differences in self-reported anxiety levels (SUDS).

This might suggest that more precise measures of speech fluency are needed to detect the impact of the audience on speech production. Relying on the broader context of PS research, we could argue for even more precise analyses of disfluencies as some authors found differential effect of filled pauses on the quality of PS according to their production mechanism (i.e., nasal "um" VS oral "err"; Niebuhr and Fischer, 2019; Voss and Niebuhr, 2024). While our study gives valuable insight about differential impact of PS anxiety in real and virtual settings, research in this domain remains scarce and would benefit from more detailed measures (i.e., distinction based on types of filled pauses).

4.3 Voice

In regard to pitch parameters, median f_0 was different between the real audience and the virtual empty condition. However, it was not statistically different in the two audience conditions (RQ1) and the two virtual conditions (RQ2). As observed in Remacle et al. (2021), voice frequency features in the virtual audience therefore seems to be more comparable to the real audience than the virtual empty condition (control). However, as observed differences in voice frequency between our study conditions are about 10 Hz for women and 5 Hz for men, they might not be clinically relevant. Same interpretation can be made for intonation (f_0 SD), where we did not find any significant difference between experimental conditions (as in Remacle et al., 2021).

4.4 Virtual reality

Regarding VR-related measures (RQ3), the mean ITC-SOPI scores fell within the normal range established by Lessiter et al. (2001), suggesting that participants experienced levels of immersion and presence comparable to those typically observed in the general population. Furthermore, no significant differences were found on the Negative Effects subscale of the ITC-SOPI before and after VR immersion, indicating that the virtual experience did not induce unwanted side effects such as nausea or oculomotor discomfort. This favorable user experience is an essential condition for validating our virtual audience and is expected to promote the improvement of oral communication skills practiced in VR (Witmer and Singer, 1998). The virtual audience therefore seems to be acceptable and usable for PS training and research.

4.5 General discussion and profile of participants

Overall, some results support the validation of our virtual audience, while others indicate aspects that require further investigation. Results that validated the virtual audience were: the increase in *HR* dynamics between *HR–Before* and *HR–1* as well as the absence of significant difference for *HR* before, during and after the talk, median f_0, f_0 SD, and *TD_Percentage* for the two audience conditions (RQ1), the significant difference between virtual conditions in *SUDS-Before* (RQ2), the strong sense of presence and very low cybersickness in the virtual audience (RQ3). The virtual audience thus seem to be efficient in inducing anticipatory anxiety but lacked inducing specific changes in voice parameters related to pitch. Despite the fact that our results show the ability of our virtual audience to generate PS-related anxiety in our participants, several results do not support our hypotheses: the significant difference between the two audience conditions in *Filled pauses_Percentage* and *SUDS* before, during and after the talk (RQ1) the absence of significant difference between the two virtual conditions in *TD_Percentage*, *Filled pauses_Percentage* median f_0, f_0 SD, *SUDS* and *HR* during and after the talk (RQ2). In other words, the virtual audience generates fewer filled pauses and less self-reported anxiety than the real audience (RQ1). Furthermore, it does not differ statistically from the empty virtual room in terms of disfluencies, vocal parameters, and measured or self-reported anxiety throughout the PS (RQ2).

These heterogenous results regarding the different aspects of PS (anxiety, voice and fluency) could be explained by the theory of vocal response under emotional arousal presented in our introduction. It posits voice adaptations rely on the interaction of different levels: physiological, phonatory-articulatory and acoustic (Juslin and Scherer, 2005). In the present study, the virtual audience is found to be efficient in inducing psychological stress (as shown by significant increase in self-reported anticipatory anxiety before speaking in front of a virtual audience as compared to before speaking in front of a virtual empty room) and leads to changes in response at physiological level (as reviewed by significant increase in *HR* in the first minute of PS). However, as indicated by the standardized anxiety questionnaires and physiological measures, our participants experienced globally low levels of anxiety. Combined with the lack of interaction with the virtual agents, this could explain why we failed to observe significant differences between the two VR conditions in all our acoustic level measures throughout the entire 3-min task. As the virtual audience did not answer questions and did not react to the talk in a positive or a negative way, valence and arousal expressed by our virtual agents might not have been sufficient to maintain the observed increase in anxiety levels after the first minute of the talk (as in Remacle et al., 2021). Once peak anxiety occurred from the initial view of the virtual audience, habituation pattern could have taken place similarly in both VR environments until a return to baseline happened (Takac et al. 2019). However, virtual agents' posture, facial expressions and head movement could be modified in following studies in order to increase the influence of social stressors in the virtual meeting room (Etienne et al., 2023; Slater et al., 1999). Moreover, these nonverbal behavior modifications could be implemented as real-time audience feedback to further heighten anxiety levels (El-Yamri et al., 2019; Schneider et al., 2019). We also assume our results could have been different with a sample of more anxious participants (Moïse-Richard et al., 2021; Kothgassner et al., 2016) or even a clinical population (e.g., Diemer et al., 2014; Robillard et al., 2003; Simon et al., 2020; Stupar-Rutenfrans et al., 2017). The group analysis conducted in this study supports this interpretation (RQ4). The second group, characterized by a majority of women with higher anxiety level, exhibited behaviors that were more comparable to those of the real audience. This group elicited a percentage of disfluencies closer to the real audience condition in the virtual audience condition than in the empty virtual condition and tended to feel more present in both virtual environments. Our results are concordant with the literature (Felnhofer et al., 2014; Ling et al., 2014) as women with high anxiety level are shown to be most immersed and to exhibit greater behavioral changes after VR intervention. This also suggests a sample of more anxious participants might be most suited in future PS VR purposes using these environments.

4.6 Limitations and perspectives

Some limitations may have affected this experiment. First, the recruitment of university students might reduce the representativeness of the sample and might undermine the external validity of the findings. As a result, the conclusions drawn may be limited in scope and difficult to generalize to a broader population. Moreover, regarding study procedure, participants were informed they would attend to two sessions: one dedicated to the real meeting room with an audience, and the other dedicated to the immersions in the virtual environments (virtual meeting room with an audience and empty virtual meeting room). Although these conditions were counterbalanced to control for potential order effects, participants still had to perform one PS task in the real session and two in the virtual session, one after another. This time proximity between the two VR immersion could reduce the likeliness to demonstrate a significant difference between environment and impact habituation pattern between immersion (Takac et al. 2019). However, due to practical constraints and to avoid dropouts, we did not ask participants to attend to a third session in order to have only one PS task by session. The knowledge of the study procedure might also have influenced PS anxiety, and more specifically anticipatory anxiety, as students were aware of the type of audience they would speak to. Once more, methodological precautions such as counterbalancing of the conditions limited this prior knowledge from impacting our results. Finally, restricting the analysis to the first 3 minutes of the planned 5-min talk may have influenced our results. As a reminder, we decided to cut the talks after 3 min in order to make PS comparable across participants and conditions. However, while the initial minutes of PS are widely regarded as the most indicative (Behnke and Sawyer, 2004), some participants anxiety and behavioral changes could still have occurred in the latter minute of the talk. Given the lack of statistical differences that was repeatedly observed, even in the beginning of PS, this interpretation remains to be rather unlikely. These limitations represent possible future improvements that could enhance the user experience and the effectiveness of the virtual environment.

As for now, our study provides practitioners involved in PS training (i.e., coaches, speech-language pathologists) with an innovative and scientifically validated tool: a virtual audience eliciting anticipatory anxiety. Beyond this practical contribution, our findings deepen our understanding of anxiety habituation pattern within speech and how it shapes voice and fluency. Importantly, this work stands among the first to jointly assess both anxiety responses and communicative behaviors in VR PS, and to directly compare these responses with those observed in real-life PS situations. Our work also reinforces the crucial role of the initial moments of PS in behavior change. This suggests the first minutes of PS should be prioritized in training thus reducing time and financial constraints without impacting efficacy (Takac et al. 2019).

5 Conclusion

VR simulations may represent a promising method to practice PS skills in situations similar to reality. Our results suggest that the virtual audience elicited an increase in anticipatory anxiety compared to a talk in front of an empty virtual meeting room. The strong feeling of presence and the lack of side effects such as cybersickness enhance the validation of our virtual audience and represent encouraging results for future use of this VR environment to train oral communication skills. Moreover, we found two different profiles regarding VR immersion in our participants. The profiles distinguished based on gender, state and trait anxiety, feeling of presence and percentage of disfluencies in the virtual environments. More specifically, women with higher anxiety levels and high percentage of disfluencies exhibited behaviors more comparable to those of the real audience. Therefore, our results extend previous data and support the feasibility and relevance of using virtual environments for PS research, intervention and training. Ultimately, it contributes to the scientific validation of VR for PS training, but also advances our understanding of how anxiety might shape vocal and fluency behavior in ecologically valid settings. Next step would be to analyze the effect of avatars' attitude and virtual audience's size on PS behavior in order to propose VR interventions and training considering the multiple variables influencing PS skills.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Faculty of Psychology, Speech-Language Therapy and Educational Sciences of the University of Liège – Belgium in March 2023 (file number: 2223-059/5559). It was conducted following the ethical standards described in the Declaration of Helsinki (1964). All participants provided written informed consent after receiving a complete description of the study. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any identifiable images or data included in this article.

Author contributions

LB: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing – original draft, Writing – review and editing. AR: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review and editing. MS: Software, Writing – review and editing. ÉE: Formal Analysis, Writing – review and editing. A-ME: Writing – review and editing. A-LL: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review and editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. Lamia Bettahi is a FRESH grantee of the Fonds de la Recherche Scientifique – FNRS (F.R.S.-FNRS, Brussels, Belgium; grant number: 40021892). All other authors work at the University of Liège and carried out this study as part of their research work.

Acknowledgements

We thank Sarah Van Kerckhove for her assistance with data collection and the AR/VR lab of HEC Liège, for their assistance throughout the conduct of the study.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abercrombie, D. (2013). The north wind and the sun, 1951–1978 [sound]. *Univ. Edinb. Sch. Philos. Psychol. Lang. Sci. Dept. Linguist. Engl. Lang.* doi:10.7488/ds/157

Amir, O., and Levine-Yundof, R. (2013). Listeners' attitude toward people with dysphonia. *J. Voice* 27, 524.e1–524.e7. doi:10.1016/j.jvoice.2013.01.015

Baptista, C. A., Loureiro, S. R., de Lima Osório, F., Zuardi, A. W., Magalhães, P. V., Kapczinski, F., et al. (2012). Social phobia in Brazilian university students: prevalence, under-recognition and academic impairment in women. *J. Affect. Disord.* 136, 857–861. doi:10.1016/j.jad.2011.09.022

Barsties, B., and De Bodt, M. (2015). Assessment of voice quality: current state-of-the-art. *Auris Nasus Larynx* 42, 183–188. doi:10.1016/j.anl.2014.11.001

Behnke, R. R., and Sawyer, C. R. (2001). Patterns of psychological state anxiety in public speaking as a function of anxiety sensitivity. *Commun.* 49, 84–94. doi:10.1080/01463370109385624

Behnke, R. R., and Sawyer, C. R. (2004). Public speaking anxiety as a function of sensitization and habituation processes. *Commun. Educ.* 53, 164–173. doi:10.1080/03634520410001682494

Benito, K. G., and Walther, M. (2015). Therapeutic process during exposure: habituation model. *J. Obsessive-Compuls. Relat. Disord.* 6, 147–157. doi:10.1016/j.jocrd.2015.01.006

Benjamin, C., O'Neil, K., Crawley, S., Beidas, R., Coles, M., and Kendall, P. (2010). Patterns and predictors of subjective units of distress in anxious youth. *Behav. Cogn. Psychother.* 38, 497–504. doi:10.1017/S1352465810000287

Bitsika, V., Sharpley, C. F., Sweeney, J. A., and McFarlane, J. R. (2014). HPA and SAM axis responses as correlates of self vs parental ratings of anxiety in boys with an autistic disorder. *Physiol. Behav.* 127, 1–7. doi:10.1016/j.physbeh.2013.12.011

Boetje, J., and van Ginkel, S. (2021). 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.* 37, 253–264. doi:10.1111/jcal.12484

Brundage, S. B., Brinton, J. M., and Hancock, A. B. (2016). Utility of virtual reality environments to examine physiological reactivity and subjective distress in adults who stutter. *J. Fluen. Disord.* 50, 85–95. doi:10.1016/j.jfludis.2016.10.001

Buchanan, T. W., Laures-Gore, J. S., and Duff, M. C. (2014). Acute stress reduces speech fluency. *Biol. Psychol.* 97, 60–66. doi:10.1016/j.biopsycho.2014.02.005

Chrousos, G. P. (2009). Stress and disorders of the stress system. *Nat. Rev. Endocrinol.* 5, 374–381. doi:10.1038/nrendo.2009.106

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., and Vervliet, B. (2014). Maximizingexposure therapy: an inhibitory learning approach. *Behav. Res. Ther.* 58, 10–23. doi:10.1016/j.brat.2014.04.006

Daniels, M. M., Palaoag, T., and Daniels, M. (2020). Efficacy of virtual reality in reducing fear of public speaking: a systematic review. *IOP Conf. Ser. Mater. Sci. Eng.* 803, 012003. doi:10.1088/1757-899X/803/1/012003

Davis, A., Linvill, D. L., Hodges, L. F., Da Costa, A. F., and Lee, A. (2020). Virtual reality *versus* face-to-face practice: a study into situational apprehension and performance. *Commun. Educ.* 69 (1), 70–84. doi:10.1080/03634523.2019.1684535

Daviu, N., Bruchas, M. R., Moghaddam, B., Sandi, C., and Beyeler, A. (2019). Neurobiological links between stress and anxiety. *Neurobiol. Stress* 11, 100191. doi:10.1016/j.ynstr.2019.100191

De Grez, L., Valcke, M., and Roozen, I. (2009). The impact of goal orientation, self-reflection and personal characteristics on the acquisition of oral presentation skills. *Eur. J. Psychol. Educ.* 24, 293–306. doi:10.1007/bf03174762

De Jong, N. H. (2018). Fluency in second language testing: insights from different disciplines. *Lang. Assess. Q.* 15, 237–254. doi:10.1080/15434303.2018.1477780

Demers, M., Fung, K., Subramanian, S. K., Lemay, M., and Robert, M. T. (2021). Integration of motor learning principles into virtual reality interventions for individuals with cerebral palsy: systematic review. *JMIR Serious Games* 9, e23822. doi:10.2196/23822

Diemer, J., Mühlberger, A., Pauli, P., and Zwanzger, P. (2014). Virtual reality exposure in anxiety disorders: impact on psychophysiological reactivity. *World J. Biol. Psychiatry* 15, 427–442. doi:10.3109/15622975.2014.892632

Duez, D. (2001). Significant des hésitations dans la parole spontanée. *Rev. Parole* 17, 113–138.

Dwyer, K. K., and Davidson, M. M. (2012). Is public speaking really more feared than death? *Commun. Res. Rep.* 29, 99–107. doi:10.1080/08824096.2012.667772

Ebrahimi, O. V., Pallesen, S., Kenter, R. M., and Nordgreen, T. (2019). Psychological interventions for the fear of public speaking: a meta-analysis. *Front. Psychol.* 10, 488. doi:10.3389/fpsyg.2019.00488

El-Yamri, M., Romero-Hernandez, A., Gonzalez-Riojo, M., and Manero, B. (2019). Emotions-responsive audiences for VR public speaking simulators based on the speakers' voice. *2019 IEEE 19th Int. Conf. Adv. Learn. Technol.* 2161, 349–353. doi:10.1109/ICALT.2019.00105

Etienne, E., Leclercq, A. L., Remacle, A., Dessart, L., and Schyns, M. (2023). Perception of avatars' nonverbal behaviors in virtual reality. *Psychol. Mark.* 40, 2464–2481. doi:10.1002/mar.21871

Felnhofer, A., Kothgassner, O. D., Hetterle, T., Beutl, L., Hlavacs, H., and Kryspin-Exner, I. (2014). Afraid to be there? Evaluating the relation between presence, self-reported anxiety, and heart rate in a virtual public speaking task. *Cyberpsychol. Behav. Soc. Netw.* 17, 310–316. doi:10.1089/cyber.2013.0472

Felton, W. M., and Jackson, R. E. (2022). Presence: a review. *Int. J. Hum. Comput. Interact.* 38, 1–18. doi:10.1080/10447318.2021

Ferreira Marinho, A. C., Mesquita de Medeiros, A., Côrtes Gama, A. C., and Caldas Teixeira, L. (2017). Fear of public speaking: perception of college students and correlates. *J. Voice* 31, 7–11. doi:10.1016/j.jvoice.2015.12.012

Finn, A. N., Sawyer, C. R., and Schrodt, P. (2009). Examining the effect of exposure therapy on public speaking state anxiety. *Commun. Educ.* 58 (1), 92–109. doi:10.1080/03634520802450549

Gauthier, J., and Bouchard, S. (1993). Adaptation Canadienne-Française de la forme révisée du State-Trait Anxiety Inventory de Spielberger. *Can. J. Behav. Sci.* 25, 559–578. doi:10.1037/h0078881

Giddens, C. L., Barron, K. W., Byrd-Craven, J., Clark, K. F., and Winter, A. S. (2013). Vocal indices of stress: a review. *J. Voice* 27, 21–29. doi:10.1016/j.jvoice.2012.12.010

Gil, D., Carrier, B., Fullmer, W., Cruz, K., Aguilar, C. D., Davis, D. W., et al. (2021). Validity of average heart rate and energy expenditure in polar OH1 and verity sense while self-paced running. *Int. J. Exerc. Sci. Conf. Proc.* 14, 27.

Goberman, A. M., Hughes, S., and Haydock, T. (2011). Acoustic characteristics of public speaking: anxiety and practice effects. *Speech Commun.* 53, 867–876. doi:10.1016/j.specom.2011.02.005

Gilkinson, H. (1942). Social fears as reported by students in college speech classes. *Speech Monogr.* 9, 141–160. doi:10.1080/03637754209390068

Gruber, A., and Kaplan-Rakowski, R. (2020). "User experience of public speaking practice in virtual reality," in *Cogn. Affect. Perspect. Immers. Technol. Educ.* Editor R. Zheng (Hershey, PA: IGI Global), 235–249. doi:10.4018/978-1-7998-3250-8.ch012

Hagenaars, M. A., and Van Minnen, A. (2005). The effect of fear on paralinguistic aspects of speech in patients with panic disorder with agoraphobia. *J. Anxiety Disord.* 19, 521–537. doi:10.1016/j.janxdis.2004.04.008

Harris, S. R., Kemmerling, R. L., and North, M. M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychol. Behav.* 5 (6), 543–550. doi:10.1089/109493102321018187

Harris, D. J., Bird, J. M., Smart, P. A., Wilson, M. R., and Vine, S. J. (2020). A framework for the testing and validation of simulated environments in experimentation and training. *Front. Psychol.* 11, 605. doi:10.3389/fpsyg.2020.00605

Hartanto, D., Kampmann, I. L., Morina, N., Emmelkamp, P. G., Neerincx, M. A., and Brinkman, W. P. (2014). Controlling social stress in virtual reality environments. *PLoS One* 9, e92804. doi:10.1371/journal.pone.0223988

Heeren, A., Maurage, P., Rossignol, M., Vanhaelen, M., Peschard, V., Eeckhout, C., et al. (2012). Self-report version of the liebowitz social anxiety scale: psychometric properties of the French version. *Can. J. Behav. Sci.* 44, 99–107. doi:10.1037/a0026249

Heeren, A., Ceschi, G., Valentiner, D. P., Dethier, V., and Philippot, P. (2013). Assessing public speaking fear with the short form of the personal report of confidence as a speaker scale: confirmatory factor analyses among a French-speaking community sample. *Neuropsychiatr. Dis. Treat.* 9, 609–618. doi:10.2147/NDT.S43097

Heeter, C. (1992). Being there: the subjective experience of presence. *Presence Teleoperators Virtual Environ.* 1, 262–271. doi:10.1162/pres.1992.1.2.262

Hofmann, S. G., Gerlach, A. L., Wender, A., and Roth, W. T. (1997). Speech disturbances and gaze behavior during public speaking in subtypes of social phobia. *J. Anxiety Disord.* 11, 573–585. doi:10.1016/S0887-6185(97)00040-6

Iverach, L., Menzies, R. G., O'Brian, S., Packman, A., and Onslow, M. (2011). Anxiety and stuttering: continuing to explore a complex relationship. *Am. J. Speech. Lang. Pathol.* 20, 221–232. doi:10.1044/1058-0360(2011/10-0091

Juslin, P. N., and Scherer, K. R. (2005). "Vocal expression of affect," in *The new handbook of methods in nonverbal behavior research*. Editors J. A. Harrigan, R. Rosenthal, and K. R. Scherer (Oxford University Press), 65–135.

Kahlon, S., Lindner, P., and Nordgreen, T. (2019). Virtual reality exposure therapy for adolescents with fear of public speaking: a non-randomized feasibility and pilot study. *Child. Adolesc. Psychiatry Ment. Health* 13, 1–10. doi:10.1186/s13034-019-0307-y

Kaplan, A. D., Cruit, J., Endsley, M., Beers, S. M., Sawyer, B. D., and Hancock, P. A. (2021). The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: a meta-analysis. *Hum. Factors* 63, 706–726. doi:10.1177/0018720820904229

Kappas, A., Hess, U., and Scherer, K. R. (1991). "Voice and emotion," in *Fundamentals of nonverbal behavior*. Editors B. Rim and R. S. Feldman (Cambridge University Press), 200–238.

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., and Koo, B. H. (2018). Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry Investig.* 15, 235–245. doi:10.30773/pi.2017.08.17

Kothgassner, O. D., Kafka, J., Rudyk, J., Beutl, L., Hlavacs, H., and Felnhofer, A. (2014). Does social exclusion hurt virtually like it hurts in real life? The role of agency and social presence in the perception and experience of social exclusion. *Proc. Int. Soc. Presence Res.*, 45–56.

Kothgassner, O. D., Felnhofer, A., Hlavacs, H., Beutl, L., Palme, R., Kryspin-Exner, I., et al. (2016). Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment. *Comput. Hum. Behav.* 62, 124–135. doi:10.1016/j.chb.2016.03.081

Kothgassner, O. D., Goreis, A., Glenk, L. M., Kafka, J. X., Pfeffer, B., Beutl, L., et al. (2021). Habituation of salivary cortisol and cardiovascular reactivity to a repeated real-life and virtual reality trier social stress test. *Physiol. Behav.* 242, 113618. doi:10.1016/j.physbeh.2021.113618

Kryston, K., Goble, H., and Eden, A. (2021). Incorporating virtual reality training in an introductory public speaking course. *J. Commun. Pedag.* 4, 133–151. doi:10.31446/JCP.2021.1.13

Laukka, P., Linnman, C., Åhs, F., Pissiota, A., Frans, Ö., Faria, V., et al. (2008). In a nervous voice: acoustic analysis and perception of anxiety in social phobics' speech. *J. Nonverbal Behav.* 32, 195–214. doi:10.1007/s10919-008-0055-9

Lazarus, R. S., and Opton, E. M. (1996). "The study of psychological stress: a summary of theoretical formulations and experimental findings," in *Anxiety and behavior*. Editor C. D. Spielberger (Academic Press), 225–262.

Leal, P. C., Goes, T. C., da Silva, L. C. F., and Teixeira-Silva, F. (2017). Trait vs. state anxiety in different threatening situations. *Trends Psychiatry Psychother.* 39, 147–157. doi:10.1590/2237-6089-2016-0044

LeFebvre, L. E., LeFebvre, L., and Allen, M. (2021). "imagine all the people": imagined interactions in virtual reality when public speaking. *Imagin. Cogn. Pers.* 40, 189–222. doi:10.1177/0276236620938310

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: the ITC-sense of presence inventory. *Presence Teleoperators Virtual Environ.* 10, 282–297. doi:10.1162/105474601300343612

Lindner, P. (2021). Better, virtually: the past, present, and future of virtual reality cognitive behavior therapy. *Int. J. Cognitive Ther.* 14 (1), 23–46. doi:10.1007/s41811-020-00090-7

Ling, Y., Nefs, H. T., Morina, N., Heynderickx, I., and Brinkman, W. P. (2014). A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. *PLoS One* 9, e96144. doi:10.1371/journal.pone.0096144

Lister, H. (2016). *The effect of virtual reality exposure on fear of public speaking using cloud-based software*. New Brunswick: University of New Brunswick.

MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk*. 3rd edn. Mahwah, NJ: Lawrence Erlbaum Associates.

Maymon, C. N., Crawford, M. T., Blackburne, K., Botes, A., Carnegie, K., Mehr, S. A., et al. (2024). The presence of fear: how subjective fear, not physiological changes, shapes the experience of presence. *J. Exp. Psychol. Gen.* 153, 1500–1516. doi:10.1037/xge0001576

Menjot, P., Bettahi, L., Leclercq, A. L., Durieux, N., and Remacle, A. (2023). Interventions that target or affect voice or speech production during public speaking: a scoping review. *J. Voice* 39, 1696.e1–1696.e15. doi:10.1016/j.jvoice.2023.06.021

Metz, M. J., and James, L. E. (2019). Specific effects of the trier social stress test on speech fluency in young and older adults. *Aging Neuropsychol. Cogn.* 26, 558–576. doi:10.1080/13825585.2018.1503639

Moïse-Richard, A., Ménard, L., Bouchard, S., and Leclercq, A. L. (2021). Real and virtual classrooms can trigger the same levels of stuttering severity ratings and anxiety in school-age children and adolescents who stutter. *J. Fluen. Disord.* 68, 105830. doi:10.1016/j.jfludis.2021.105830

Monteiro, D., Liang, H. N., Li, H., Fu, Y., and Wang, X. (2020). Evaluating the need and effect of an audience in a virtual reality presentation training tool. *Commun. Comput. Inf. Sci.* 1300, 62–70. doi:10.1007/978-3-030-63426-1_7

Montes, C. C., Heinicke, M. R., and Geierman, D. M. (2019). Awareness training reduces college students' speech disfluencies in public speaking. *J. Appl. Behav. Anal.* 52, 746–755. doi:10.1002/jaba.569

Morreale, S. P., and Pearson, J. C. (2008). Why communication education is important: the centrality of the discipline in the 21st century. *Commun. Educ.* 57, 224–240. doi:10.1080/03634520701861713

Niebuhr, O., and Fischer, K. (2019). "Do not hesitate!–unless you do it shortly or nasally: how the phonetics of filled pauses determine their subjective frequency and perceived speaker performance," in *Proc interspeech 2019*, 544–548.

Niebuhr, O., Skarnitzl, R., and Tylečková, L. (2018). "The acoustic fingerprint of a charismatic voice: initial evidence from correlations between long-term spectral features and listener ratings," in *Proc 9th int conf speech prosody*, 359–363. doi:10.21437/SpeechProsody.2018-73

Notaro, A., Capraro, F., Pesavento, M., Milani, S., and Busà, M. G. (2021). Effectiveness of VR immersive applications for public speaking enhancement. *Electron Imaging* 33, 1–7. doi:10.2352/ISSN.2470-1173.2021.9.IQSP-294

Palmas, F., Cichor, J., Plecher, D. A., and Klinker, G. (2019). "Acceptance and effectiveness of a virtual reality public speaking training," in *Proc ISMAR 2019*, 363–371. doi:10.1109/ISMAR.2019.00034

Pertaub, D. P., Slater, M., and Barker, C. (2001). An experiment on fear of public speaking in virtual reality. *Stud. Health Technol. Inf.* 81, 372–378. doi:10.1037/e705412011-025

Pertaub, D. P., Slater, M., and Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence* 11 (1), 68–78. doi:10.1162/105474602317343668

R Core Team (2023). R: a language and environment for statistical computing, version 4.3. Available online at: https://cran.r-project.org.

Radley, J. J., and Herman, J. P. (2023). Preclinical models of chronic stress: adaptation or pathology? *Biol. Psychiatry* 94, 194–202. doi:10.1016/j.biopsych.2022.11.004

Rasouli, S., Ghafurian, M., Nilsen, E. S., and Dautenhahn, K. (2024). University students' opinions on using intelligent agents to cope with stress and anxiety in social situations. *Comput. Hum. Behav.* 153, 108072. doi:10.1016/j.chb.2023.108072

Reeves, R., Curran, D., Gleeson, A., and Hanna, D. (2022). A meta-analysis of the efficacy of virtual reality and *in vivo* exposure therapy as psychological interventions for public speaking anxiety. *Behav. Modif.* 46, 937–965. doi:10.1177/0145445521991102

Remacle, A., Bouchard, S., Etienne, A. M., Rivard, M. C., and Morsomme, D. (2021). A virtual classroom can elicit teachers' speech characteristics: evidence from acoustic measurements during *in vivo* and in virtuo lessons, compared to a free speech control situation. *Virtual Real.* 25, 935–944. doi:10.1007/s10055-020-00491-1

Remacle, A., Bouchard, S., and Morsomme, D. (2023). Can teaching simulations in a virtual classroom help trainee teachers to develop oral communication skills and self-efficacy? A randomized controlled trial. *Comput. Educ.* 200, 104808. doi:10.1016/j.compedu.2023.104808

Riley, G. D. (2009). *SSI-4: stuttering severity instrument*. 4th edn. Pro-Ed.

Robillard, G., Bouchard, S., Fournier, T., and Renaud, P. (2003). Anxiety and presence during VR immersion: a comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games. *Cyberpsychol Behav.* 6, 467–476. doi:10.1089/109493103769710497

Robles, M. M. (2012). Executive perceptions of the top 10 soft skills needed in today's workplace. *Bus. Commun. Q.* 75, 453–465. doi:10.1177/1080569912460400

Russell, G., and Topham, P. (2012). The impact of social anxiety on student learning and well-being in higher education. *J. Ment. Health* 21, 375–385. doi:10.3109/09638237.2012.694505

Scheveneels, S., Boddez, Y., Van Daele, T., and Hermans, D. (2019). Virtually unexpected: no role for expectancy violation in virtual reality exposure for public speaking anxiety. *Front. Psychol.* 10, 2849. doi:10.3389/fpsyg.2019.02849

Schmid Mast, M., Kleinlogel, E. P., Tur, B., and Bachmann, M. (2018). The future of interpersonal skills development: immersive virtual reality training with virtual humans. *Hum. Resour. Dev. Q.* 29, 125–141. doi:10.1002/hrdq.21307

Schneider, J., Romano, G., and Drachsler, H. (2019). Beyond reality—Extending a presentation trainer with an immersive VR module. *Sensors* 19 (16), 3457. doi:10.3390/s19163457

Schneider, M., Kraemmer, M. M., Weber, B., and Schwerdtfeger, A. R. (2021). Life events are associated with elevated heart rate and reduced heart complexity to acute psychological stress. *Biol. Psychol.* 163, 108116. doi:10.1016/j.biopsycho.2021.108116

Schommer, N. C., Hellhammer, D. H., and Kirschbaum, C. (2003). Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress. *Psychosom. Med.* 65, 450–460. doi:10.1097/01.PSY.0000035721.12441.17

Simon, J., Etienne, A. M., Bouchard, S., and Quertemont, E. (2020). Alcohol craving in heavy and occasional alcohol drinkers after cue exposure in a virtual environment: the role of the sense of presence. *Front. Hum. Neurosci.* 14, 124. doi:10.3389/fnhum.2020.00124

Singh, D., and Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Appl. Soft Comput.* 97, 105524. doi:10.1016/j.asoc.2019.105524

Slater, M., Usoh, M., and Steed, A. (1994). Depth of presence in virtual environments. *Presence Teleoperators and Virtual Environ.* 3 (2), 130–144. doi:10.1162/pres.1994.3.2.130

Slater, M., Pertaub, D. P., and Steed, A. (1999). Public speaking in virtual reality: facing an audience of avatars. *IEEE Comput. Graph. Appl.* 19 (2), 6–9. doi:10.1109/38.749116

Slater, M., Pertaub, D. P., Barker, C., and Clark, D. M. (2006). An experimental study on fear of public speaking using a virtual environment. *Cyberpsychol. Behav.* 9 (5), 627–633. doi:10.1089/cpb.2006.9.627

Spielberger, C. D. (1983). *State-trait anxiety inventory for adults*. Mind Garden.

Spierer, D. K., Rosen, Z., Litman, L. L., and Fujii, K. (2015). Validation of photoplethysmography as a method to detect heart rate during rest and exercise. *J. Med. Eng. Technol.* 39, 264–271. doi:10.3109/03091902.2015.1047536

Stupar-Rutenfrans, S., Ketelaars, L. E., and van Gisbergen, M. S. (2017). Beat the fear of public speaking: mobile 360° video virtual reality exposure training in home environment reduces public speaking anxiety. *Cyberpsychol. Behav. Soc. Netw.* 20 (10), 624–633. doi:10.1089/cyber.2017.0174

Takac, M., Collett, J., Blom, K. J., Conduit, R., Rehm, I., and De Foe, A. (2019). Public speaking anxiety decreases within repeated virtual reality training sessions. *PLoS One* 14, e0216288. doi:10.1371/journal.pone.0216288

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., and Updegraff, J. A. (2000). Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychol. Rev.* 107, 411–429. doi:10.1037/0033-295X.107.3.411

The jamovi project (2024). Jamovi (version 2.5) Available online at: https://www.jamovi.org (Accessed November 11, 2025).

Tomasi, J., Zai, C. C., Zai, G., Herbert, D., Richter, M. A., Mohiuddin, A. G., et al. (2024). Investigating the association of anxiety disorders with heart rate variability measured using a wearable device. *J. Affect Disord.* 351, 569–578. doi:10.1016/j.jad.2024.01.137

Tsigos, C., Kyrou, I., Kassi, E., and Chrousos, G. P. (2020). *Stress: endocrine physiology and pathophysiology*. Endotext.

Valls-Ratés, Ï., Niebuhr, O., and Prieto, P. (2022). Unguided virtual-reality training can enhance the oral presentation skills of high-school students. *Front. Commun.* 7, 910952. doi:10.3389/fcomm.2022.910952

Valls-Ratés, I., Niebuhr, O., and Prieto, P. (2023). Encouraging participant embodiment during VR-assisted public speaking training improves persuasiveness and charisma and reduces anxiety in secondary school students. *Front. Virtual Real* 4, 1074062. doi:10.3389/frvir.2023.1074062

Van Ginkel, S., Gulikers, J., Biemans, H., and Mulder, M. (2015). Towards a set of design principles for developing oral presentation competence: a synthesis of research in higher education. *Educ. Res. Rev.* 14, 62–80. doi:10.1016/j.edurev.2015.02.002

Van Ginkel, S., Ruiz, D., Mononen, A., Karaman, C., De Keijzer, A., and Sitthiworachart, J. (2020). The impact of computer-mediated immediate feedback on developing oral presentation skills: an exploratory study in virtual reality. *J. Comput. Assist. Learn* 36, 412–422. doi:10.1111/jcal.12424

van Kraaij, A. W. J., Schiavone, G., Lutin, E., Claes, S., and Van Hoof, C. (2020). Relationship between chronic stress and heart rate over time modulated by gender in a cohort of office workers: cross-sectional study using wearable technologies. *J. Med. Internet Res.* 22, e18253. doi:10.2196/18253

Van Puyvelde, M., Neyt, X., McGlone, F., and Pattyn, N. (2018). Voice stress analysis: a new framework for voice and effort in human performance. *Front. Psychol.* 9, 1994. doi:10.3389/fpsyg.2018.01994

Von der Pütten, A., Kramer, N. C., Gratch, J., and Kang, S. H. (2010). "It doesn't matter what you are!" explaining social effects of agents and avatars. *Comput. Hum. Behav.* 26, 1641–1650. doi:10.1016/j.chb.2010.06.012

Voss, S., and Niebuhr, O. (2024). "Beautiful noise? The impact of filled pauses on the perception of speaker charisma," in *2024 IEEE international professional communication conference (ProComm)* (IEEE), 131–138.

Witmer, B. G., and Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence Teleoperators Virtual Environ.* 7, 225–240. doi:10.1162/105474698565686

Woisard, V., Bodin, S., and Puech, M. (2004). The voice handicap index: impact of the translation in French on the validation. *Rev. Laryngol. Otol. Rhinol.* 125, 307–312.

Wörtwein, T., Chollet, M., Schauerte, B., Morency, L.-P., Stiefelhagen, R., and Scherer, S. (2015). Multimodal public speaking performance assessment. *Proc. Int. Conf. Multimodal Interact.*, 43–50. doi:10.1145/2818346.2820762

Yang, L., Zhou, X., Pu, J., Liu, L., Cuijpers, P., Zhang, Y., et al. (2019). Efficacy and acceptability of psychological interventions for social anxiety disorder in children and adolescents: a meta-analysis of randomized controlled trials. *Eur. Child. Adolesc. Psychiatry* 28, 79–89. doi:10.1007/s00787-018-1189-x

Yuen, E. K., Goetter, E. M., Stasio, M. J., Ash, P., Mansour, B., McNally, E., et al. (2019). A pilot of acceptance and commitment therapy for public speaking anxiety delivered with group videoconferencing and virtual reality exposure. *J. Context. Behav. Sci.* 12, 47–54. doi:10.1016/j.jcbs.2019.01.006

Zacarin, M. R. J., Borloti, E., and Haydu, V. B. (2019). Behavioral therapy and virtual reality exposure for public speaking anxiety. *Trends Psychol.* 27, 491–507. doi:10.9788/TP2019.2-14

Zhang, G. Q., and Zhang, W. (2009). Heart rate, lifespan, and mortality risk. *Ageing Res. Rev.* 8, 52–60. doi:10.1016/j.arr.2008.10.001